RETRIEVING VERTICAL PROFILES AND TROPOSPHERIC COLUMNS OF FORMALDEHYDE FROM GLOBAL LONG-TERM MAX-DOAS MEASUREMENTS

Dissertation
for the award of the degree
"Doctor of Natural Sciences"
submitted to the
Department of Physics, Mathematics and Computer Science
of the Johannes Gutenberg University at Mainz

Sebastian Donner
born in Werneck

Mainz, 23 May 2024





MAX PLANCK INSTITUTE FOR CHEMISTRY
JOHANNES GUTENBERG UNIVERSITY MAINZ

1. Referee: Prof. Dr. Thomas Wagner
2. Referee: Prof. Dr. Peter Hoor
Supervisor: Prof. Dr. Thomas Wagner

Oral examination: 23 October 2024

ABSTRACT

In the presence of elevated abundances of nitrogen oxides (NO_x), the photooxidation of volatile organic compounds (VOCs) can result in increased O_3 levels. Consequently, the abundances of both NO_x and VOCs have to be considered when aiming to model or control O_3 abundances. It is therefore indispensable to understand and quantify the sources of VOCs, to characterise their distribution in the atmosphere and to identify chemical transformations they undergo. In this context, ground-based MAX-DOAS measurements at four measurement sites were used in this thesis to establish a consistent long-term data set of HCHO concentration profiles and vertical column densities in different parts of the world covering different climatic zones and environmental conditions. The HCHO data set was complemented by glyoxal, NO_2 and aerosol results.

To achieve a consistent analysis of the measured data, a comprehensive sensitivity study was performed using both simulated and measured spectra to systematically assess and optimise the DOAS analysis settings of HCHO resulting in improved uniform settings for multiple atmospheric conditions. Here, major improvement was achieved by the use of so-called Puķīte terms for NO_2 and O_3. Further, remaining systematic uncertainties were investigated. Improvements were also made for the analysis settings of the other species, in particular glyoxal (and NO_2), and the DOAS retrieval in general.

Based on the resulting consistent data set, different sources of HCHO and glyoxal under different environmental conditions were identified by investigating their annual and diurnal cycles (and those of aerosol), their dependencies on air temperature and wind direction as well as vertical concentration profiles, and combining the results of the different species. For ATTO, this was amended by direct comparisons of the results of two instruments installed at 80 m and 298 m from which small-scale concentration gradients were inferred.

The results showed that both biogenic and anthropogenic sources play a role at all stations whereby biogenic sources are often dominant. Their relative contributions vary significantly between the different sites and seasons. Furthermore, it was concluded that formation from precursors usually dominates, regardless of whether the precursors are of anthropogenic or biogenic origin. Nevertheless, also direct (anthropogenic) emissions can sometimes significantly contribute to the prevailing HCHO abundances especially during winter at mid-latitudes when biogenic activity is low. The results also showed that in several cases long-range transport or transport from more nearby sources can become important although often local effects dominate the abundances. One of the most evident findings was that HCHO and glyoxal have similar or even common sources but are chemically processed differently in the atmosphere affecting both formation from precursors and degradation. The comparisons of the MAX-DOAS HCHO results to TROPOMI satellite data showed overall reasonable to good agreement between both data sets for all stations.

ZUSAMMENFASSUNG

Unter dem Einfluss erhöhter Stickoxidkonzentrationen (NO_x) kann die Photooxidation von volatilen organischen Verbindungen (VOCs) zu erhöhten O_3 Konzentrationen führen. Folglich müssen sowohl die NO_x als auch die VOC Konzentrationen berücksichtigt werden, wenn man O_3 Konzentrationen modellieren oder regulieren möchte. Daher ist es unerlässlich, die Quellen von VOCs zu verstehen und zu quantifizieren, ihre Verteilung in der Atmosphäre zu charakterisieren und chemische Prozesse zu identifizieren, an denen sie beteiligt sind. In diesem Zusammenhang wurden in dieser Arbeit bodengestützte MAX-DOAS-Messungen an vier Messorten dazu verwendet, einen konsistenten Langzeitdatensatz von HCHO Konzentrationsprofilen und vertikalen Säulendichten in verschiedenen Teilen der Welt zu erstellen, der unterschiedliche Klimazonen und Umweltbedingungen abdeckt. Der HCHO Datensatz wurde außerdem durch NO_2, Glyoxal- und Aerosolergebnisse ergänzt.

Um einen konsistenten Datensatz zu erhalten, wurde eine umfassende Sensitivitätsstudie durchgeführt, für die sowohl simulierte als auch gemessene Spektren verwendet wurden, um die DOAS-Analyse für HCHO systematisch zu untersuchen und zu optimieren. Dies führte zu einer einheitlichen DOAS-Analyse für viele verschiedene atmosphärische Bedingungen. Dabei wurde eine wesentliche Verbesserung durch die Verwendung von sogenannten Puḳīte-Termen für NO_2 und O_3 erreicht. Außerdem wurden die verbleibenden systematischen Unsicherheiten untersucht. Verbesserungen wurden auch für die Analysen der anderen Stoffe, insbesondere Glyoxal (und NO_2), und für die DOAS-Analyse im Allgemeinen erzielt.

Basierend auf dem resultierenden konsistenten Datensatz wurden verschiedene HCHO und Glyoxalquellen unter verschiedenen Umweltbedingungen identifiziert, indem jahreszeitliche und tägliche Variationen der Spurengaskonzentrationen (und Aerosolkonzentrationen), deren Abhängigkeiten von Lufttemperatur und Windrichtung sowie vertikale Konzentrationsprofile untersucht und die Ergebnisse der verschiedenen Spurenstoffe kombiniert wurden. Für die Messungen an ATTO wurde dies durch direkte Vergleiche der Ergebnisse von zwei in 80 m und 298 m Höhe installierten Instrumenten ergänzt, aus denen kleinräumige Konzentrationsgradienten abgeleitet wurden.

Die Ergebnisse zeigten, dass sowohl biogene als auch anthropogene Quellen an allen Stationen eine Rolle spielen, wobei biogene Quellen oft dominieren. Ihre relativen Beiträge variieren erheblich zwischen den verschiedenen Messorten und Jahreszeiten. Generell überwiegt die Bildung aus Vorläufersubstanzen, unabhängig davon, ob die Vorläufersubstanzen anthropogenen oder biogenen Ursprungs sind. Dennoch können auch direkte (anthropogene) Emissionen erheblich zu den vorherrschenden HCHO Konzentrationen beitragen, insbesondere im Winter, wenn die biogene Aktivität in den mittleren Breiten gering ist. In mehreren Fällen zeigte sich, dass Ferntransport oder der Transport von näher gelegenen Quellen wichtig werden kann, obwohl oft lokale Effekte die Konzentrationen bestimmen.

Eines der wichtigsten Ergebnisse war, dass HCHO und Glyoxal ähnliche oder sogar gemeinsame Quellen haben, aber in der Atmosphäre unterschiedlich chemisch verarbeitet werden, was sowohl die Bildung aus Vorläuferstoffen als auch den Abbau betrifft. Die Vergleiche der MAX-DOAS HCHO Ergebnisse mit den TROPOMI-Satellitendaten zeigten eine überwiegend gute Übereinstimmung für alle Stationen.

CONTENTS

I MOTIVATION AND SCIENTIFIC BACKGROUND

1 INTRODUCTION 3

2 THE ROLE OF FORMALDEHYDE IN TROPOSPHERIC CHEMISTRY 11
- 2.1 The NO_x cycle . . . 12
- 2.2 Methane chemistry . . . 15
- 2.3 The HO_x cycle . . . 17
- 2.4 Relevance of glyoxal . . . 20
  - 2.4.1 Sources and sinks of glyoxal . . . 20
  - 2.4.2 The ratio glyoxal to formaldehyde (R_{GF}) . . . 22

II METHODS AND DATA

3 MAX-DOAS AND PROFILE RETRIEVAL 27
- 3.1 Differential Optical Absorption Spectroscopy (DOAS) . . . 27
  - 3.1.1 Absorption spectroscopy and the Beer-Lambert law . . . 27
  - 3.1.2 Absorption spectroscopy in the atmosphere . . . 28
  - 3.1.3 The classical DOAS approach . . . 29
  - 3.1.4 Passive DOAS in this work . . . 31
  - 3.1.5 Implications using passive DOAS . . . 33
    - 3.1.5.1 Ring effect . . . 34
    - 3.1.5.2 Puḳīte terms . . . 35
    - 3.1.5.3 Saturation and solar I_0-effect . . . 36
- 3.2 MultiAXis-DOAS (MAX-DOAS) . . . 39
  - 3.2.1 Measurement geometry . . . 39
  - 3.2.2 Quantities retrieved from MAX-DOAS measurements . . . 41
  - 3.2.3 Radiative transfer and air mass factors . . . 41
  - 3.2.4 O_4 absorptions . . . 44
- 3.3 The Mainz Profile Algorithm (MAPA) . . . 45

4 INSTRUMENTS AND MEASUREMENT SITES 51
- 4.1 Mainz, Germany . . . 51
  - 4.1.1 The 4-azimuth MAX-DOAS instrument . . . 54
    - 4.1.1.1 Instrumental setup . . . 54
    - 4.1.1.2 Characteristics . . . 55
- 4.2 Bayfordbury, United Kingdom . . . 56
  - 4.2.1 The TubeMAX-DOAS instrument . . . 59
    - 4.2.1.1 Instrumental setup . . . 59
    - 4.2.1.2 Characteristics . . . 61
- 4.3 Mohali, India . . . 64
- 4.4 ATTO, Brazil . . . 67
- 4.5 Research cruise MSM 68/2, North Atlantic . . . 69

5 ADDITIONAL DATA SETS 73
- 5.1 ERA5 data . . . 73
- 5.2 AERONET data . . . 74

5.3 TROPOMI formaldehyde data 76

III IMPROVEMENT OF MAX-DOAS RESULTS

6 SENSITIVITY STUDY FOR THE DOAS FIT SETTINGS OF FORMALDEHYDE 81
6.1 Used data sets and study approach 81
6.1.1 Synthetic and measured spectra 82
6.1.2 Study approach 84
6.2 Puḳīte terms and other effects for NO_2 and O_3 absorptions 86
6.2.1 Treatment of Puḳīte terms 87
6.2.2 NO_2 absorptions 93
6.2.2.1 Synthetic spectra 93
6.2.2.2 Measured spectra 106
6.2.3 O_3 absorptions 112
6.2.3.1 Synthetic spectra 113
6.2.3.2 Measured spectra 118
6.3 Choice of the fitting interval 124
6.3.1 Study approach 124
6.3.2 Analyses and investigations 126
6.4 Remaining deviations 142
6.4.1 Ring effect and noise 142
6.4.2 Intensity weighted convolution for other absorbers 146
6.4.3 Further higher order terms 149
6.4.4 Wavelength dependence and accuracy of the AMFs 153
6.4.5 Accuracy and precision of the simulated spectra 158
6.4.6 Low resolution simulations 162
6.4.7 Discussion of remaining deviations 167
6.5 Further fit parameters 169
6.5.1 Intensity offset, spectral shift and stretch 170
6.5.2 Polynomial degree 174
6.5.3 BrO absorption 177
6.6 Comparison with other common fit settings 179
6.7 Main findings and uncertainty estimation 185

7 IMPROVEMENT OF THE DOAS FIT SETTINGS OF FURTHER SPECIES 193
7.1 Water vapour absorptions in the O_4 fitting window 193
7.2 Improvements of the glyoxal fit 199
7.2.1 Treatment of water vapour residual structures 199
7.2.2 NO_2 Pukite terms under low NO_2 conditions 206
7.2.3 Choice of the fitting interval for glyoxal and NO_2 209
7.2.4 Spectral gaps for the treatment of a residual structure around 430 nm 212

8 FURTHER IMPROVEMENTS OF THE DOAS RETRIEVAL 217
8.1 Spectral calibration assuming a Super Gaussian slit function . . . 217
8.2 Influence of changing spectrometer temperature 220

IV MEASUREMENTS OF TROPOSPHERIC FORMALDEHYDE

9 DATA ANALYSIS 227
9.1 Spectral analysis 227

9.2 MAPA retrieval . . . 232
9.2.1 Retrieval settings . . . 232
9.2.2 Elevation biases in MAPA retrievals . . . 234
9.2.3 Near-surface concentrations and total columns . . . 241
9.3 Data selection . . . 243
9.4 ERA5 vs. local meteorological data . . . 247
10 BAYFORDBURY, UK 253
10.1 Overview and particularities of the analysis . . . 253
10.2 Annual and diurnal cycles . . . 254
10.3 Influence of meteorological conditions . . . 265
10.4 Trace gas correlations . . . 275
10.5 Vertical profiles . . . 283
10.6 Comparison to TROPOMI . . . 286
11 MAINZ, GERMANY 291
11.1 Overview and particularities of the analysis . . . 291
11.2 Annual and diurnal cycles . . . 293
11.3 Influence of meteorological conditions . . . 303
11.4 Trace gas correlations . . . 311
11.5 Vertical profiles . . . 314
11.6 Comparison to TROPOMI . . . 317
11.7 Further results: HONO . . . 319
12 MOHALI, INDIA 325
12.1 Overview and particularities of the analysis . . . 325
12.2 Annual and diurnal cycles . . . 326
12.3 Influence of meteorological conditions . . . 343
12.4 Trace gas correlations . . . 353
12.5 Vertical profiles . . . 361
12.6 Comparison to TROPOMI . . . 365
12.7 Further results: HONO . . . 368
13 ATTO, BRAZIL 375
13.1 Overview and particularities of the analysis . . . 375
13.2 Annual and diurnal cycles . . . 377
13.3 Influence of meteorological conditions . . . 398
13.4 Trace gas correlations . . . 411
13.5 Vertical profiles . . . 425
13.6 Comparison to TROPOMI . . . 430
13.7 Further results: water vapour . . . 434
13.8 Direct comparison of the two instruments . . . 437

V CONCLUSIONS AND OUTLOOK

14 CONCLUSIONS 445
15 OUTLOOK 451

VI APPENDIX

A SUPPLEMENTARY TABLES 455
B SUPPLEMENTARY FIGURES 459

B.1 HCHO sensitivity study 459
B.2 Improvement of the DOAS fit settings of further species 462
B.3 Data analysis 464
B.4 Bayfordbury 469
B.5 Mohali 471

LIST OF FIGURES 473
LIST OF TABLES 485
BIBLIOGRAPHY 487
PUBLICATIONS WITH OWN CONTRIBUTION 510

ACKNOWLEDGEMENTS 515
CURRICULUM VITAE 519

ACRONYMS

AERONET	aerosol robotic network
AMF	air mass factor
AOD	aerosol optical depth
ATTO	Amazon Tall Tower Observatory
boxAMF	box-air mass factor
DC	dark current
DOAS	Differential Optical Absorption Spectroscopy
dAMF	differential air mass factor
dSCD	differential slant column density
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	fifth generation ECMWF reanalysis
FOV	field of view
FP	first principles
FRS	Fraunhofer reference spectrum
FWHM	full width at half maximum
HR	high resolution
ITCZ	intertropical convergence zone
LR	low resolution
LT	local time
LUT	look-up table
MAPA	Mainz Profile Algorithm
MC	Monte Carlo
NMVOC	non-methane volatile organic compound
OVOC	oxygenated volatile organic compound
PSA	pseudo absorber
PW	precipitable water
RAA	relative azimuth angle
RMS	root mean square
RTM	radiative transfer model
SCD	slant column density
SF	scaling factor
SLF	slit function

SOA	secondary organic aerosol
SZA	solar zenith angle
TROPOMI	Tropospheric Monitoring Instrument
UT	universal time
UV	ultraviolet
VCD	vertical column density
VOC	volatile organic compound
XS	cross-section

Part I

MOTIVATION AND SCIENTIFIC BACKGROUND

1 INTRODUCTION

Earth's atmosphere, although relatively thin compared to the Earth radius (6378 km) with a height of only around 100 km (Kármán line), makes our planet habitable. This gaseous blanket is composed and structured in a way that it protects the inhabitants of Earth from dangerous short-wave UV radiation, gives oxygen to breathe and water essential for all species on Earth.

Only three gases, namely nitrogen (N_2, ∼78 %), oxygen (O_2, ∼21 %) and argon (Ar, ∼1 %), account for almost the entire volume of the atmosphere. The abundances of these gases are controlled over geologic time scales by the biosphere, uptake and release from crustal material and degassing of the interior. The next most abundant constituent is water vapour which mainly occurs in the lower atmosphere and whose abundances are controlled by evaporation and precipitation. Its atmospheric concentration is highly variable ranging from almost zero in very dry regions up to about 3 – 4 % in humid tropical regions. Besides these main constituents, myriads of gases, the so-called *trace gases*, are present in the atmosphere representing significantly less than 1 % of the atmosphere. Nevertheless, these trace gases are of central importance for Earth's radiative balance and the chemical properties of the atmosphere as well as for the organisms on Earth (Seinfeld and Pandis, 2016).

Amongst the most abundant trace gases are carbon dioxide (CO_2) and methane (CH_4). Together with water vapour (and other less abundant species like N_2O) they cause the natural greenhouse effect by "trapping" a certain fraction of the long-wave radiation emitted by the Earth's surface in the atmosphere. This effect is essential for life on Earth since it leads to a mean surface temperature of about 15 °C which would otherwise be equal to the Earth's effective radiative temperature of −18 °C (Seinfeld and Pandis, 2016). While the atmospheric abundances of CO_2 and CH_4 were practically unchanged from the end of the last ice age some 10 000 years ago until roughly 300 years ago at mixing ratios of 260 and 0.7 ppm (parts per million), respectively, they show a dramatic increase since the start of the industrialisation about 200 years ago (Seinfeld and Pandis, 2016). There is scientific consensus that this increase is caused by mankind due to the combustion of fossil fuels (coal and oil) for energy, transportation or industrial processes. Nowadays, the mixing ratios of CO_2 and CH_4 have already exceeded 420 and 1.9 ppm (NOAA – Global Monitoring Laboratory, 2024), respectively, and the resulting man-made greenhouse effect already led to an increase of the global mean temperature of 1.45 °C by 2023 compared to the pre-industrial baseline (World Meteorological Organization, 2024). This process is known as global heating and is an impressive example for the importance of these less abundant trace gases for life on Earth.

Another important atmospheric trace gas is ozone (O_3). Around 90 % of the atmospheric O_3 is accumulated in the stratospheric ozone layer in an altitude

range from 20 to 30 km and absorbs together with O_2 (being the precursor for O_3) almost all solar radiation with wavelengths below 290 nm. In addition, also radiation with wavelengths below 320 nm is significantly attenuated. In that way, ozone protects living organisms at Earth's surface from harmful UV radiation and makes life on Earth possible. Also the increasing temperature with increasing altitude within the stratosphere is a result of this absorption of UV radiation. Contrary to that, ozone in the troposphere, although accounting only for about 10 % of the total atmospheric O_3, is considered as an air pollutant since it is harmful to human health at high concentrations, e.g. affecting the respiratory system, and plants (Seinfeld and Pandis, 2016). Moreover, it acts as a greenhouse gas heavily influencing Earth's radiative budget especially in the upper troposphere (Lelieveld and Dorland, 1995; Iglesias-Suarez et al., 2018; Nussbaumer et al., 2023). Apart from these adverse effects, the photolysis of tropospheric ozone in the presence of water vapour is a major source of the hydroxyl radical (OH) which is the most important oxidising agent in the atmosphere being responsible for its self-cleaning capacity since the oxidation of atmospheric constituents by OH leads to the removal of pollutants of both natural and anthropogenic origin from the atmosphere (Atkinson, 2000). Despite minor natural sources, namely entrainment from the stratosphere, the photodissociation of molecular oxygen by remaining short-wave radiation, thunderstorms as well as the oxidation of carbon monoxide (CO) and methane (in the presence of nitrogen oxides) in the background atmosphere, dangerous tropospheric O_3 concentrations occur in polluted regions where the photooxidation of volatile organic compounds (VOCs) in the presence of elevated abundances of nitrogen oxides (NO_x) leads to a net production of ozone (Haagen-Smit, 1952; Williams and Koppmann, 2007). Typical O_3 surface mixing ratios range from 20 to 60 ppb (parts per billion) while values can exceed 100 ppb in polluted urban environments (Seinfeld and Pandis, 2016). Since ozone production is a non-linear function of the abundances of NO_x and VOCs (more precisely VOC reactivity), both have to be considered when aiming to control ozone abundances (Pusede and Cohen, 2012).

Hence, nitrogen oxides, i.e. nitric oxide (NO) and nitrogen dioxide (NO_2), are amongst the most important molecules in atmospheric chemistry. Apart from being strongly involved in ozone chemistry, NO_x itself is dangerous for humans, e.g. causing headache, breathlessness, vertigo or damage to the respiratory tract. Typical NO_x levels are 10 to 100 ppt (parts per trillion) in the remote atmosphere, around 1 ppb in rural areas and 5 to 20 ppb in urban environments (Seinfeld and Pandis, 2016).

While the situation is rather lucid for NO_x, it is estimated that as many as 10^5 different VOC species have already been measured in the atmosphere which is probably only a small fraction of the number actually present (Goldstein and Galbally, 2007; Kaiser et al., 2015). Excluding CH_4, around 90 % of the global VOC emissions are biogenic, e.g. by plants or soil (Seinfeld and Pandis, 2016), affecting the OH budget and thus the self-cleaning capacity of the atmosphere and chemistry in general also in remote locations. In the presence of NO_x, VOCs regardless of being of natural or anthropogenic origin are involved in O_3 production. Further, they are supposed to be involved in the formation of secondary

organic aerosol (SOA) and some of them act as greenhouse gases. Apart from that, the remaining 10 % originating from anthropogenic sources are of particular importance as they are typically emitted in areas with high NO_x abundances eventually leading to high O_3 levels affecting air quality. Moreover, especially under such polluted conditions where chemistry is more complex, the chemical oxidation of VOCs and thus the ozone-forming potential is often not well represented by models (Williams and Koppmann, 2007). In addition to these chemical aspects, some VOCs (especially oxygenated VOCs) are harmful to humans themselves as several of them affect the respiratory tract, irritate the eyes, are toxic or are suspected to be mutagenic and carcinogenic (Koppmann and Wildt, 2007). To model and control the impact of VOCs on tropospheric chemistry and air quality, it is indispensable to understand and quantify their sources, to characterise their distribution in the atmosphere (horizontally and vertically) and to identify chemical transformations they undergo. Although since the 1950's great progress was made in measuring the vast variety of different VOCs by different in-situ measurement approaches (for details and examples refer for instance to Williams and Koppmann, 2007, and references therein), it is impossible to assess each and every individual species. Further, most VOCs are highly reactive and thus the in-situ sampling might underlie measurement artefacts like inlet effects, where the gases stick to surfaces, thermal decomposition or chemical transformations, i.e. production or degradation (Williams and Koppmann, 2007). Since formaldehyde (HCHO) is formed as a stable intermediate from almost all VOCs at some stage during their oxidation, it can serve as a valuable proxy for VOC reactivity and chemical oxidation processes in the atmosphere which is for many applications sufficient to know (e.g. Williams and Koppmann, 2007; Pusede and Cohen, 2012; Seinfeld and Pandis, 2016). In a similar way, glyoxal (CHOCHO) is formed during the degradation of many VOCs. However, since it is a more complex molecule containing two carbon atoms, glyoxal is formed from a smaller subset of VOCs. Therefore, combining measurements of HCHO and glyoxal provides additional insights into VOC speciation and probably allows to discriminate different sources or chemical regimes (e.g. Vrekoussis et al., 2010; Kaiser et al., 2015). In addition, both species also have direct sources which have to be taken into account when considering VOC chemistry.

For all these reasons, accurate measurements of HCHO and glyoxal are crucial. Although these two oxygenated VOCs are even more reactive and sticky than other VOCs and thus more difficult to measure with sufficient accuracy (Koppmann and Wildt, 2007), especially in-situ measurements of HCHO have a long tradition using for instance in-situ spectroscopic (including DOAS-based long-path or White cell systems), colorimetric, fluorescence or chemiluminescent detection techniques. Detailed descriptions and inter-comparisons of the different methods can be found in existing literature such as Vairavamurthy et al. (1992), Hak et al. (2005) and Koppmann and Wildt (2007).

Since both HCHO and glyoxal show distinct spectral absorption features in the UV/visible spectral range, they are detectable by passive differential optical absorption spectroscopy (DOAS). More specifically, passive DOAS measurements make use of characteristic spectral absorption features of trace gases in spectra of

scattered sunlight. Major advantages of this remote sensing technique against in-situ measurements are that no "physical sampling" is necessary which avoids inlet or other effects described above, the possibility to measure several trace gases, e.g. HCHO, glyoxal and NO_2, with only one instrument as well as, compared to DOAS-based in-situ measurements, more compact and easier setups since no artificial light source nor mirrors are needed. Further, space-based DOAS measurements by satellites allow a global view on the distributions of both species which is not possible with the point measurements provided by in-situ techniques.

Such satellite observations of tropospheric HCHO have been performed for years from a variety of different UV/visible satellite instruments, e.g. the Global Ozone Monitoring Experiment (GOME; Thomas et al., 1998; Chance et al., 2000; De Smedt et al., 2008), the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY; Wittrock et al., 2006; De Smedt et al., 2008), the Ozone Monitoring Instrument (OMI; González Abad et al., 2015; De Smedt et al., 2015; Kaiser et al., 2018) and GOME-2 (Vrekoussis et al., 2010; De Smedt et al., 2012, 2015). More recently, the TROPOspheric Monitoring Instrument (TROPOMI; Veefkind et al., 2012; De Smedt et al., 2018, 2021) was launched onboard the Sentinel-5 Precursor (S5P) satellite in October 2017 providing an unprecedented spatial resolution of $3.5 \times 5.5\,\mathrm{km}^2$. The data obtained from these instruments are used in a vast amount of studies related to air quality and climate change (e.g. Stavrakou et al., 2009b; Duncan et al., 2010; Fortems-Cheiney et al., 2012; Barkley et al., 2013; Chan Miller et al., 2017; Alvarado et al., 2020). As an example, Fig. 1.1 shows the global HCHO distribution (April 2018 to April 2019) measured by the recent TROPOMI satellite instrument.

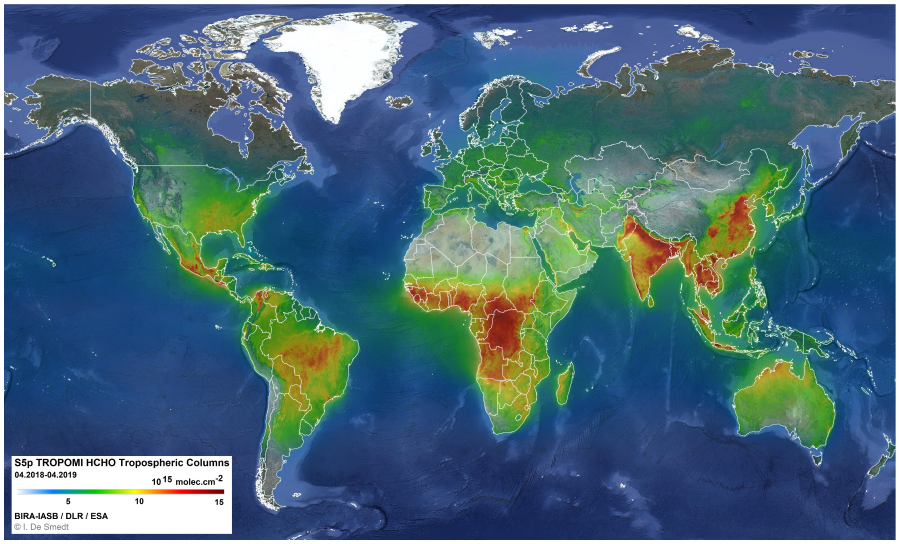


Figure 1.1: Global distribution of HCHO measured by TROPOMI. Shown are mean HCHO VCDs in the period from April 2018 to April 2019. Contains modified Copernicus data (2018/2019). Provided with permission by De Smedt (2021).

Satellite measurements are very valuable as they offer a global view on trace gas distributions. However, they usually have only one local overpass per day (if at all) and are thus not capable of capturing systematic diurnal variations of

the trace gas abundances. Although this might change with a new generation of geostationary satellite sensors such as the Geostationary Environment Monitoring Spectrometer (GEMS; Kim et al., 2020), the Tropospheric Emissions: Monitoring of POllution instrument (TEMPO; Zoogman et al., 2017), and Sentinel-4 (S4; Ingmann et al., 2012), further drawbacks of satellite measurements are that they cannot provide vertical distributions of trace gases (except for some limb sounders), their comparably large uncertainties and the decreasing sensitivity to trace gases towards the Earth's surface.

These limitations can be addressed by Multi-AXis (MAX-) DOAS measurements which make use of trace gas absorptions in spectra recorded at different elevation angles (Hönninger and Platt, 2002). The long atmospheric light paths for measurements at low elevation angles lead to high sensitivities to trace gases in the lowest tropospheric layers and thus to rather small uncertainties. MAX-DOAS measurements can be more than ten times more sensitive to trace gases located close to the ground compared to satellites. Combining the information from different elevations further allows the retrieval of vertical profiles of trace gas concentrations and aerosol extinction (Hönninger and Platt, 2002; Wagner et al., 2004; Frieß et al., 2006; Clémer et al., 2010). Since such MAX-DOAS systems have comparably easy and compact setups, measurements can be conducted automatically, for long time periods and with a good temporal sampling allowing long-term monitoring of tropospheric trace gases and aerosol including in particular also their diurnal cycles. Besides that, MAX-DOAS measurements can help to validate and complement satellite retrievals as they can be regarded as a link between satellite observations, model simulations and in-situ (point) measurements (Pinardi et al., 2013).

Amongst the first reported MAX-DOAS measurements of HCHO were measurements performed in the summers 2002 and 2003 in the Po Valley in Italy (Heckel et al., 2005; Wagner et al., 2011). Since then various ground-based MAX-DOAS measurements of HCHO have been reported focussing on different aspects such as air quality including for instance source identifications, comparison exercises between different MAX-DOAS instruments or different ways to measure HCHO and the validation of satellite measurements (e.g. Pinardi et al., 2013; Vlemmix et al., 2015; Behrens et al., 2019; Wang et al., 2019; Kumar et al., 2020; Schreier et al., 2020; De Smedt et al., 2021; Hong et al., 2022).

Measurements of glyoxal are more rare since its DOAS retrieval is more challenging compared to HCHO. On the one hand, this is caused by the smaller atmospheric abundances. On the other hand, strong spectral interferences caused by water vapour absorptions often affect the DOAS retrieval. In recent years, however, both the detectors and the retrievals have been improved so that also glyoxal can be retrieved from both ground- and space-based DOAS measurements. First open atmosphere measurements of glyoxal were performed by Volkamer et al. (2005a) in the metropolitan area of Mexico City in spring 2003 using an active DOAS setup. Shortly after that, Sinreich et al. (2007) reported on the first comprehensive analysis of glyoxal using a passive MAX-DOAS setup. After that, several studies based on MAX-DOAS measurements have been published (e.g. Sinreich et al., 2010; Behrens et al., 2019; Benavent et al., 2019; Hong et al., 2022).

Some of them not only focussed on the aspects already mentioned for HCHO but also attempted to gain benefit of the simultaneous measurement of glyoxal and HCHO abundances in a way to attribute them to specific sources or to identify different chemical regimes. Also satellite measurements of glyoxal have been reported in several publications (e.g. Wittrock et al., 2006; Lerot et al., 2010; Chan Miller et al., 2014; Lerot et al., 2021) and have been used in various studies (e.g. Stavrakou et al., 2009a; Vrekoussis et al., 2010; Kaiser et al., 2015; Chan Miller et al., 2016; Alvarado et al., 2020).

An aspect worth mentioning is that most of the past ground-based MAX-DOAS measurements of HCHO and glyoxal were performed at one specific measurement site having characteristic environmental conditions, i.e. meteorology and pollution levels. Further, the fit settings used for the DOAS analyses often differ notably between different studies. As the absorptions of HCHO and glyoxal are often rather weak, changes in the retrieval settings might influence the results significantly. Hence, a direct comparison of the different data sets is not always easily possible and a tricky task. In addition, several of the existing publications report on campaign-based rather than continuous data records over longer time periods.

In the context outlined above, this thesis aims to contribute to a better understanding of the sources and the processing of VOCs in the atmosphere. This objective is approached from a measurement perspective by using ground-based MAX-DOAS measurements at four different sites to establish a consistent long-term data set of HCHO concentration profiles and vertical column densities in different parts of the world covering different climatic zones and environmental conditions. The data set is complemented by glyoxal, NO_2 and aerosol results. Great emphasis is put on a consistent analysis of the data from the different instruments located at the different sites. Here, a major aspect is the systematic assessment of the DOAS analysis settings of HCHO and their optimisation with a strong focus on consistency using both simulated and measured spectra. Improvements are also made for the analysis settings of the other species. Based on the data set obtained using these optimised retrieval settings, an attempt is made to identify different sources of HCHO and to characterise how these depend on meteorological parameters, i.e. temperature and wind fields. For that, also the results of glyoxal, NO_2 and aerosol are considered. Further, it is investigated how the HCHO abundances at the different sites depend on the prevailing pollution levels, in particular the NO_2 abundances. Moreover, the effects of transport processes including nearby sources and long-range transport are examined. By studying the vertical concentration profiles of HCHO at the different sites, it is investigated how HCHO is photochemically formed from its precursors and how the vertical distribution is affected by photochemistry and vertical transport. For ATTO, where two instruments are operated at two different altitudes, this is amended by a direct comparison of the results of both instruments from which small-scale concentration gradients are inferred. Especially based on the comparison of the HCHO and glyoxal results, an attempt is made to identify different chemical regimes, i.e. different precursor VOC compositions, at the different sites and in the different seasons at a given site. A last aspect is the comparison of the

MAX-DOAS HCHO results to satellite data from the recent TROPOMI instrument to investigate in how far the satellite measurements are representative for the true tropospheric column densities or even surface concentrations under different environmental conditions.

The presented work is divided into five parts. The first part provides an overview on the scientific background and outlines the relevance of formaldehyde in tropospheric chemistry by introducing some basic concepts of tropospheric NO_x and VOC chemistry (Chapt. 1 and 2). This is followed by a presentation of the methods and data sets in the second part consisting of a description of the measurement principle, a presentation of the used instruments and the different measurement sites as well as an overview on the used external data sets (Chapt. 3 to 5). The third part presents the results of several studies aiming to improve the formaldehyde results (Chapt. 6), the results for other species (Chapt. 7) and the DOAS retrieval in general (Chapt. 8). In the main part, first the analysis procedure is summarised including descriptions of the standard retrieval settings, the data handling and the filter criteria (Chapt. 9). Then, detailed presentations of the trace gas and aerosol results obtained for the different measurement sites are given whereby the focus is put on formaldehyde (Chapt. 10 to 13). In the fifth part, a summary of the main findings is provided together with conclusions and an outlook (Chapt. 14 and 15).

2

THE ROLE OF FORMALDEHYDE IN TROPOSPHERIC CHEMISTRY

Formaldehyde (HCHO), also known by its systematic name *methanal*, belongs to the large group of oxygenated volatile organic compounds (OVOCs) where it is the smallest and most abundant aldehyde. As shown later in this chapter, HCHO is highly involved in tropospheric photochemistry affecting ozone (O_3) and NO_x chemistry through the HO_x cycle since it is a product (often amongst the first stable ones) in the oxidation pathways of almost all volatile organic compounds (VOCs) in the atmosphere (Koppmann and Wildt, 2007; Seinfeld and Pandis, 2016). Hence, HCHO can serve as a proxy for VOC reactivity providing valuable insights into various photochemical processes of the troposphere. In addition, formaldehyde is harmful for humans itself as it is toxic and carcinogenic (Salthammer, 2013).

In general, formaldehyde has complex primary and secondary sources which are often hardly distinguishable under ambient conditions (Koppmann and Wildt, 2007).

Direct emissions of HCHO can be of both anthropogenic and biogenic origin. Such primary sources are for instance (incomplete) fossil fuel combustion processes (e.g. exhaust especially from Diesel cars and lorries), natural and man-made biomass burning fires (e.g. bush fires as well as deforestation or crop residue fires), industrial processes (e.g. refining, petrochemistry or paint industry) and solvent evaporation. Also the biosphere itself plays an important role. Here, it is known that plants can synthesise and emit OVOCs but the actual origin of HCHO itself in plants is not yet fully understood (Koppmann and Wildt, 2007; Seco et al., 2007).

Although direct emissions can play a role locally under specific conditions, secondary formation from precursor VOCs due to oxidation by the reaction with OH radicals is globally the largest HCHO source in the atmosphere. Here, it should be stressed that the oxidation pathways of almost all hydrocarbons and VOCs of both anthropogenic and biogenic origin yield HCHO in significant amounts (Koppmann and Wildt, 2007; Fortems-Cheiney et al., 2012; Seinfeld and Pandis, 2016).

In the background atmosphere, methane oxidation (compare Sect. 2.2) is the major HCHO source leading to background abundances of about 50 ppt to 1 ppb. As methane has a relatively long lifetime against OH reaction of about 9 years, methane is a rather constant global source of HCHO (Seinfeld and Pandis, 2016). Over continental areas, the oxidation of non-methane VOCs (NMVOCs) dominates over methane oxidation and becomes the major HCHO source (Fortems-Cheiney et al., 2012). Further, the large variability in the observed atmospheric formaldehyde content over continents is mainly determined by the local HCHO production from such NMVOCs (Stavrakou et al., 2009b). Here, isoprene and the group of

monoterpenes which are both mainly of biogenic origin are the main contributors besides other NMVOCs of both anthropogenic and biogenic origin (e.g. methanol and acetaldehyde). Indeed, isoprene can be considered as the main precursor of near-surface HCHO globally. Under the presence of NO_x, isoprene oxidation is also a major contributor to ozone production (Wolfe et al., 2016). Altogether, in urban atmospheres HCHO mixing ratios can vary between around 1 and 40 ppb (Koppmann and Wildt, 2007) depending on the prevailing mixture of primary sources and precursor substances as well as the availability of solar radiation. As shown in this work, the variations of active HCHO sources in the course of the year lead to pronounced annual cycles of the measured HCHO abundances.

The two major sinks of HCHO are photolysis and reaction with OH radicals, whereby photolysis is particularly important (Koppmann and Wildt, 2007; Mellouki et al., 2015). For instance, Nussbaumer et al. (2021) showed that already at mid-latitudes photolysis is the dominant loss process. As outlined in detail in Sect. 2.2, both processes are net sources of HO_x radicals and thus important in tropospheric O_3 chemistry (Pusede and Cohen, 2012). Also dry (e.g. on vegetation surfaces) and wet (e.g. during rain events as HCHO is water-soluble) deposition lead to a loss of atmospheric HCHO. Such deposition processes are, however, typically not of great importance during day time.

The lifetimes of HCHO against reaction with OH and due to photolysis are around 1.5 days and 4 hours, respectively, leading to an overall chemical lifetime of around 3.6 hours[1] (Seinfeld and Pandis, 2016). Naturally, these lifetimes depend on the atmospheric OH concentration and the availability of solar radiation. The lifetime against NO_3 is of the order of several days and thus negligible.

After this description of the basic properties of HCHO, the remaining part of this chapter briefly outlines the role of HCHO (and glyoxal) in tropospheric (ozone) chemistry and explains its secondary formation pathways as well as chemical loss processes, which have already been mentioned above, in more detail. First, the tropospheric NO_x cycle is introduced in Sect. 2.1 followed by a description of tropospheric methane chemistry including its oxidation scheme in Sect. 2.2. This scheme is then generalised to describe the oxidation of volatile organic compounds in general and its relation to the tropospheric HO_x cycle in Sect. 2.3. Lastly, Sect. 2.4 provides an overview on the sources and sinks of glyoxal and introduces the concept of the ratio glyoxal to formaldehyde.

2.1 THE NO_x CYCLE

Nitrogen oxides (NO_x), i.e. the sum of nitric oxide (NO) and nitrogen dioxide (NO_2), are of major importance for tropospheric photochemistry. While in the free troposphere NO_x sources can be both of natural (lightning or transport from the stratosphere) and anthropogenic (aircraft) origin (Zhang et al., 2003), anthropogenic activities are typically the main contributor to NO_x emissions in the boundary layer. Here, NO_x is mainly emitted by combustion processes at high temperatures like for instance in car engines or fossil fuel-fired power plants but also biomass burning fires contribute. Under specific conditions, also

1 Hereby, the assumed 12-h daytime OH concentration is 1.5×10^6 molec cm^{-3}.

soil emissions of NO_x can play an important role (Zörner et al., 2016). Most of these direct NO_x emissions are in the form of NO while direct NO_2 emissions are typically supposed to be low (Finlayson-Pitts and Pitts Jr., 1999). In combustion processes, NO is formed at high temperatures above 2000 K from the chemical conversion of atmospheric nitrogen and oxygen at these high temperatures (Zeldovich, 1946; Seinfeld and Pandis, 2016). After emission, NO is quickly converted to NO_2 on a time scale of a few minutes during day time via the reaction with O_3 (reaction R1). The resulting NO_2 is then photolysed by solar radiation with wavelengths below 424 nm regenerating an NO molecule and yielding atomic oxygen ($O(^3P)$) (reaction R2; Seinfeld and Pandis, 2016). The latter finally reacts with O_2 to regenerate an O_3 molecule (reaction R3).

$$NO + O_3 \longrightarrow NO_2 + O_2 \tag{R1}$$

$$NO_2 + h\nu \xrightarrow{\lambda < 424\,\mathrm{nm}} NO + O(^3P) \tag{R2}$$

$$O(^3P) + O_2 + M \longrightarrow O_3 + M \tag{R3}$$

The first two reactions describe the interconversion between NO and NO_2 where $h\nu$ indicates the required solar radiation. Interestingly, reaction R3 is the only significant O_3 source in the atmosphere (Seinfeld and Pandis, 2016). Here, M represents a collision partner, i.e. mostly N_2 or O_2, needed to guarantee conservation of energy since the reaction is exothermic.

As long as a photostationary state is given, reactions R1 to R3 form a so-called null cycle as each of the involved species is consumed and regenerated (compare also Fig. 2.1). In consequence, neither the NO_x nor the O_3 budget is net affected and the interconversion of the different species can be described by the so-called *Leighton ratio* (*LR*; Leighton, 1961):

$$LR = \frac{[NO]}{[NO_2]} = \frac{j(NO_2)}{k[O_3]}. \tag{2.1}$$

Here, apart from the concentrations of the involved species, $j(NO_2)$ is the photolysis rate of NO_2 and k the (temperature dependent) rate constant of reaction R1. Under typical near-surface conditions NO and NO_2 establish an equilibrium within a few minutes and *LR* is about 0.3 that means most NO_x is present as NO_2[2]. In case NO_2 can be formed from NO without consuming O_3, net ozone production occurs. As described in the next sections, this indeed happens when VOCs and HO_x radicals come into play.

Before adding the latter two to the considerations, a brief overview on the NO_x sinks is given.

During day time, the major NO_x sink is the reaction of NO_2 with the OH radical (compare next section) yielding nitric acid (HNO_3)(Atkinson, 2000).

$$NO_2 + OH + M \longrightarrow HNO_3 + M \tag{R4}$$

As shown later, this reaction is also a termination reaction in the HO_x cycle. Since HNO_3 is highly water-soluble it is removed from the atmosphere during

2 This is a typical value around noon where $j(NO_2)$ and thus the relative contribution of NO is largest. A typical 12-h average *LR* is 0.1 in urban atmospheres (Seinfeld and Pandis, 2016).

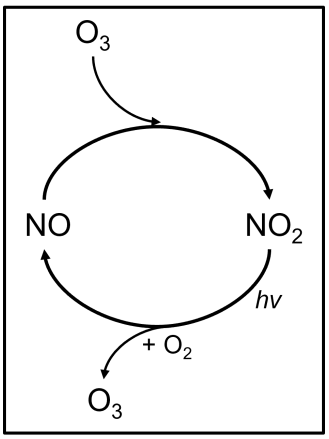


Figure 2.1: Illustration of the reactions involved in the NO_x cycle and the null cycle of O_3 as described in the text (self-created based on Atkinson, 2000).

rain events or by acid deposition. Further, HNO_3 contributes to the formation of nitrate aerosol (Bassett and Seinfeld, 1983). During day time, a typical near-surface lifetime of NO_x under background conditions is around one day (Seinfeld and Pandis, 2016). During night, *LR* is shifted even more towards NO_2 and approaches 0 since $j(NO_2)$ is 0. However, in the presence of O_3, NO_2 can then form NO_3 radicals[3].

$$NO_2 + O_3 \longrightarrow NO_3 + O_2 \tag{R5}$$

These can in turn react with NO_2 itself in a reversible reaction yielding N_2O_5 (Atkinson, 2000).

$$NO_2 + NO_3 \overset{M}{\rightleftharpoons} N_2O_5 \tag{R6}$$

In the course of the night as temperatures decrease and NO_2 levels increase, the equilibrium is shifted to the right. While gas-phase reactions of N_2O_5 are quite slow, it can be heterogeneously hydrolysed forming HNO_3 which ultimately removes NO_x from the atmosphere.

$$N_2O_5 + H_2O\,(s) \longrightarrow 2\,HNO_3 \tag{R7}$$

Together with reaction R4, this pathway is one of the major removal processes of NO_x from the atmosphere (Seinfeld and Pandis, 2016). In urban atmospheres, also the production of nitrous acid (HONO) through heterogeneous hydrolysis of NO_2 on aerosol surfaces has been observed[4]. Under the influence of solar radiation, HONO is rapidly photolysed leading to an early morning peak of OH radicals (Atkinson, 2000).

$$HONO + h\nu \longrightarrow NO + OH \tag{R8}$$

Lastly, also dry deposition on vegetation surfaces removes NO_x in the form of NO_2 from the atmosphere.

3 NO_3 is not important during day time as its noon lifetime is only around 5 s.

4 The day time reaction of NO with OH is in a rapid equilibrium with the photolysis of HONO (reaction R8).

Besides being highly involved in O_3 chemistry through reactions R1 to R3, NO_x is also harmful for humans itself causing for instance headache, breathlessness, vertigo or damage to the respiratory tract. In addition, nitrogen oxides are carcinogenic if frequently exposed to high concentrations.

2.2 METHANE CHEMISTRY

As outlined above, the main source of HCHO in the atmosphere is the oxidation of VOCs due to the reaction with hydroxyl radicals (OH). A major source of the latter in the troposphere is photolysis of O_3 by solar radiation with wavelengths between 290 and 335 nm in the presence of water vapour (H_2O) (Atkinson, 2000). The photolysis yields molecular oxygen and excited[5] singlet ($O(^1D)$) oxygen atoms (reaction R9). The latter can either lose their excess energy by collision with N_2 or O_2 and recombine with O_2 forming O_3 or, in fewer cases, react with H_2O yielding OH radicals (reaction R10) (Seinfeld and Pandis, 2016).

$$O_3 + h\nu \xrightarrow{290\,\text{nm} \leq \lambda \leq 335\,\text{nm}} O_2 + O(^1D) \tag{R9}$$

$$O(^1D) + H_2O \longrightarrow 2\,OH \tag{R10}$$

Based on this mechanism, highest OH concentrations are expected in the tropics where both water vapour concentrations and solar radiation are very high. Other sources contributing to the OH budget are the photolysis of HONO (reaction R8), HCHO or other carbonyls in the presence of NO (compare below) and the dark reaction of O_3 with alkenes offering a small night time source (Atkinson, 2000).

In the background atmosphere, the oxidation of methane (CH_4) is the dominant HCHO source[6] which is initiated by the reaction with OH yielding a methyl radical (CH_3) and H_2O.

$$CH_4 + OH \longrightarrow CH_3 + H_2O \tag{R11}$$

The methyl radical reacts quasi instantaneously with O_2 to form a methyl peroxy radical (CH_3O_2).

$$CH_3 + O_2 + M \longrightarrow CH_3O_2 + M \tag{R12}$$

In the presence of NO, the methyl peroxy radical reacts with NO yielding NO_2 and a methoxy radical (CH_3O)[7].

$$CH_3O_2 + NO \longrightarrow CH_3O + NO_2 \tag{R13}$$

The only relevant reaction for the methoxy radical under tropospheric conditions is the very rapid reaction with O_2 to produce HCHO and a hydroperoxyl radical (HO_2).

$$CH_3O + O_2 \longrightarrow HCHO + HO_2 \tag{R14}$$

5 Also a pathway via ground state atomic oxygen exists which, however, leads to a regeneration of O_3 and thus forms a null cycle.

6 If not stated otherwise, the following considerations are based on the textbook by Seinfeld and Pandis (2016) and the review paper by Atkinson (2000).

7 CH_3O_2 can also react with HO_2 eventually also yielding HCHO via methyl hydroperoxide (CH_3OOH). However, this reaction path is not detailed here in sake of conciseness.

The HO_2 radical regenerates OH via the reaction with NO yielding additionally NO_2.

$$HO_2 + NO \longrightarrow NO_2 + OH \tag{R15}$$

There are some noteworthy aspects about this reaction pathway. For instance, this oxidation chain recycles NO_2 from NO twice without consuming ozone. Also, net no OH radical is lost as it is recycled via HO_2 (reaction R15). Further, this scheme is the simplest way of secondary HCHO production via reaction R14. Methane can in principle be replaced by more complex NMVOCs leading to plenty of reaction pathways regenerating NO_2 from NO without consuming O_3 and to interconvert OH and HO_2 (compare next section).

As already described above, the two major sinks of HCHO are the reaction with OH and photolysis. The former path yields in a first step water vapour and a formyl radical (HCO) which reacts very rapidly with molecular oxygen resulting in carbon monoxide (CO) and a hydroperoxyl radical.

$$HCHO + OH \longrightarrow HCO + H_2O \tag{R16}$$
$$HCO + O_2 \longrightarrow HO_2 + CO \tag{R17}$$

The photolysis products depend on the wavelength of the inciting photon. For wavelengths above 334 nm (but below 370 nm), CO is directly formed in addition to molecular hydrogen (H_2) while for shorter wavelengths HCHO is split up into atomic hydrogen (H) and a formyl radical. The latter undergoes reaction R17 whereas H reacts with O_2 also yielding a hydroperoxyl radical.

$$HCHO + h\nu \xrightarrow{\lambda < 370\,\mathrm{nm}} CO + H_2 \tag{R18}$$
$$HCHO + h\nu \xrightarrow{\lambda < 334\,\mathrm{nm}} HCO + H \tag{R19}$$
$$H + O_2 + M \longrightarrow HO_2 + M \tag{R20}$$

Practically, every CH_4 molecule is converted to HCHO. Both reaction with OH and photolysis yield CO which is thus the principle product of methane oxidation. Without going into details, it should be noted that also the oxidation chemistry of CO, which ultimately yields CO_2 on a several-month timescale, includes reaction R15 and thus regenerates NO_2 without requiring an O_3 molecule.

The theoretical maximum yield of O_3 per methane molecule oxidised occurs under NO_x levels sufficiently high such that all involved peroxy radicals, i.e. HO_2 and CH_3O_2, react with NO yielding NO_2 and when HCHO is photolysed by reaction R19 (leading to two HO_2 molecules). In this case, NO_2 is formed from NO four times via reactions R13 and R15 as well as reaction R19 followed by reactions R17 and R20 as well as reaction R15. Thus, four O_3 molecules are formed for every CH_4 molecule oxidised via reactions R2 and R3. An illustration of this optimum yield path is provided in Fig. 2.2. In the real atmosphere, however, also reactions removing some of the involved species from the oxidation chain occur leading to a lower ozone yield. Details of these termination reactions are outlined in the next section.

Already the considerations in this section show that the production and degradation of formaldehyde is closely linked to ozone chemistry. While many of

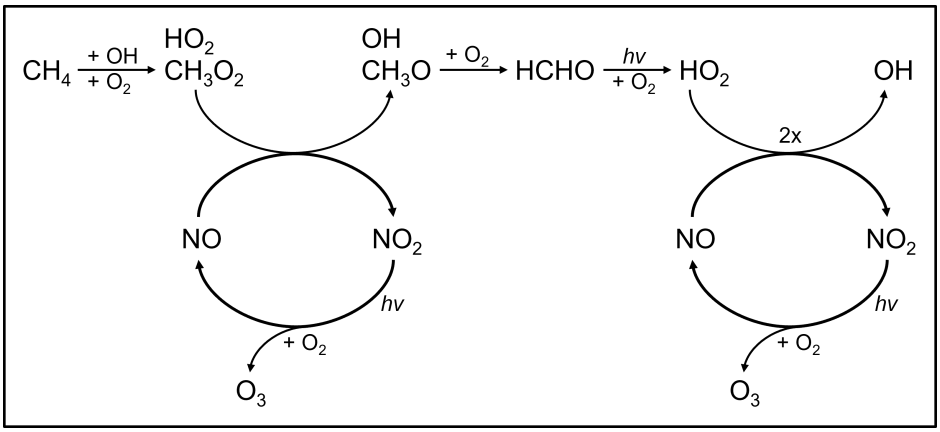


Figure 2.2: Illustration of the reactions involved in the methane oxidation (left) and the degradation of HCHO (right) as well as their respective links to the NO_x cycle. For the degradation of HCHO, the maximum O_3 yield path is shown (compare text). For the sake of readability, only relevant processes are included (self-created based on Atkinson, 2000).

the involved species/radicals almost instantaneously undergo reactions, HCHO is more stable and measurements of its abundances provide insights into the overall reactivity. In the next section, the oxidation scheme for CH_4 is generalised to describe the chemical degradation of various NMVOCs and more emphasis is put on the interconversion of OH and HO_2.

2.3 THE HO_X CYCLE

The considerations in the previous section showed that the OH radical initiates the oxidation of methane (reaction R11) but is also regenerated in the last step of the chain (R15). This recycling is known as the HO_x cycle which can be used to describe the oxidation chain not only of methane but of various NMVOCs being of both biogenic and anthropogenic origin (such as for instance alkanes and alkenes) and is described in many publications and textbooks (e.g. Atkinson, 2000; Pusede et al., 2015; Seinfeld and Pandis, 2016). The term "HO_x" summaries OH radicals and peroxy radicals, i.e. HO_2 and RO_2, which catalyse the cycle. Although for some classes of NMVOCs, also reactions with ozone and NO_3 radicals can play a role, e.g. for alkenes reactions with O_3 are important, the hydroxyl radical is the dominant oxidising species (Seinfeld and Pandis, 2016). Hence, the former processes are not detailed here and it is referred to existing textbooks like the one of Seinfeld and Pandis (2016) which is also the basis of the following discussions together with the publication by Pusede et al. (2015) if not stated otherwise.

In analogy to methane oxidation, the cycle is initiated by the reaction of a given VOC (RH) with an OH radical yielding an organic rest radical (R) and water vapour[8].

$$RH + OH \longrightarrow R + H_2O \tag{R21}$$

The organic radical reacts quasi instantaneously with the atmospheric oxygen forming an organic peroxy radical (RO_2).

$$R + O_2 + M \longrightarrow RO_2 + M \tag{R22}$$

RO_2 then oxidises NO to NO_2 while itself is reduced to RO. Here, also a second, less important, branch exists yielding an organic nitrate ($RONO_2$) which actually terminates the cycle. The importance of this second branch increases as the size of R increases.

$$RO_2 + NO \longrightarrow RO + NO_2 \tag{R23a}$$
$$RO_2 + NO + M \longrightarrow RONO_2 + M \tag{R23b}$$

Following the main path, RO reacts rapidly with O_2 generating HO_2 and a carbonyl (R'CHO)[9].

$$RO + O_2 \longrightarrow HO_2 + R'CHO \tag{R24}$$

Obviously, this reaction corresponds to the formation of HCHO from above (reaction R14) which is the carbonyl product of the methane oxidation chain. In the general scheme, the higher carbonyl (R'CHO) can in general be further oxidised if the atmospheric conditions, i.e. NO_x and HO_x levels, are suitable restarting the cycle at reaction R21. Here, it should be mentioned that this eventually also leads to the formation of HCHO as the organic rest (R) becomes smaller showing that HCHO is basically formed during the oxidation of any VOC. Finally, in the last step of the HO_x cycle, the HO_2 radical reacts with NO regenerating both NO_2 and OH (reaction R15).

Like for the methane oxidation chain, NO is oxidised to NO_2 twice for one completion of the HO_x cycle. In addition, the further chemical degradation of the carbonyl product might include this process as shown above for HCHO (reactions R16 to R20). Under the influence of sunlight, the two NO_2 molecules photolyse (reaction R2) and the resulting ground state oxygen atoms ($O(^3P)$) react with O_2 (reaction R3) yielding two O_3 molecules. An illustration of the HO_x cycle and its link to ozone formation via the NO_x cycle is provided in Fig. 2.3.

8 In the here presented example, the reaction occurs with H abstraction. In case RH contains a carbon-carbon double bond, the reaction happens by OH addition. This reaction pathway, however, is not detailed here.

9 For larger RO radicals, other reactions come into play which are not detailed here but it is referred to existing literature like the textbook of Seinfeld and Pandis (2016).

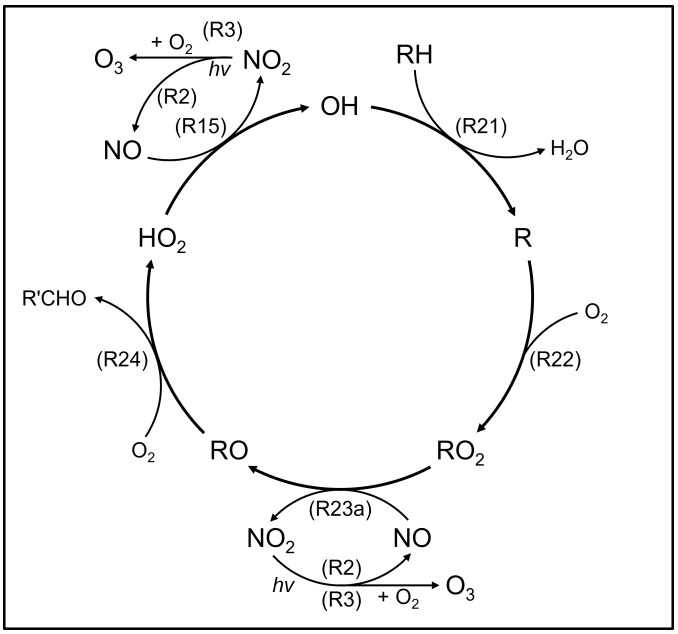


Figure 2.3: Illustration of the reactions involved in the HO_x cycle describing the oxidation of a VOC (RH) as well as its link to the NO_x cycle and thus ozone chemistry. Details are described in the text. For the sake of readability, only relevant processes are included and termination reactions are not shown. Reprinted (adapted) with permission from Nussbaumer and Cohen (2020). Copyright 2020 American Chemical Society.

In the real atmosphere, deviations from this idealised scheme occur as several termination reactions can interrupt the HO_x cycle. On the one hand, these termination reactions include self-reactions of HO_x.

$$OH + HO_2 \longrightarrow O_2 + H_2O \tag{R25}$$
$$HO_2 + HO_2 \longrightarrow H_2O_2 + O_2 \tag{R26}$$
$$RO_2 + HO_2 \longrightarrow ROOH + O_2 \tag{R27}$$
$$RO_2 + RO_2 \longrightarrow ROOR + O_2 \tag{R28}$$

On the other hand, also nitrogen oxides can terminate the cycle via reactions R4 and R23b as well as

$$RO_2 + NO_2 + M \longrightarrow RO_2NO_2 + M. \tag{R29}$$

The relevance of the different termination reactions depends on the relation between NO_x and VOC levels leading to a non-linear response of the ozone production to changes in either of the two groups. In particular, under low NO_x conditions, an increase of the NO_x abundance leads to increasing availability of NO which in turn enhances the probability of reactions R23a and R15. Hence, the oxidation rate of VOCs (RH) producing ozone is increased. The resulting response of the ozone production is almost linear with respect to NO_x and thus this regime is commonly referred to as "NO_x limited". Under such conditions, increases of the VOC abundances have only small effects since the corresponding increase of peroxy radicals (HO_2 and RO_2) mainly favours reactions R26 to R28

hindering the recycling of OH in the lack of sufficient NO to undergo reactions R23a and R15.

With further increasing NO_x levels, the reaction of NO_2 with OH radicals (R4) gains more importance net removing OH radicals from the cycle. This leads to a lower rate of VOC oxidation yielding less RO_2 as well as HO_2 radicals and thus decreases ozone production. Under high NO_x abundances, the production of HNO_3 (R4) finally dominates over ozone production and this regime is commonly know as "NO_x saturated" or "VOC limited". The latter is a reasonable term since ozone production is very sensitive to increases of the VOC abundance by increasing the probability of reaction R21 .

Assuming a constant atmospheric VOC content, maximum ozone production is found close to the transition point between both described regimes.

An implication of the "VOC limited" regime is that reductions of the NO_x abundances, e.g. on weekends in polluted cities, lead to enhanced ozone levels. This so-called weekend ozone effect was first observed in the 1970's by Levitt and Chock (1976) in the greater Los Angeles area (Nussbaumer and Cohen, 2020). Thus, air quality control measures should target both reduction of NO_x and (anthropogenic) VOC emissions. As becomes evident by the considerations in this and the previous sections, measurements of HCHO provide insights into ozone chemistry via VOC reactivity and thus provide an important tool of air quality monitoring (Pusede and Cohen, 2012).

To conclude the discussion of the HO_x cycle and its link to tropospheric ozone chemistry, it should be mentioned that ozone has several net loss processes. Besides the formation of OH radicals initiated by photolysis of O_3 in the presence of water vapour (reactions R9 and R10), under low NO_x conditions the reactions of O_3 with HO_x radicals are loss processes of tropospheric ozone (Atkinson, 2000).

$$O_3 + OH \longrightarrow HO_2 + O_2 \tag{R30}$$

$$O_3 + HO_2 \longrightarrow OH + 2\,O_2 \tag{R31}$$

The reaction of O_3 with NO yielding NO_2 (reaction R1) is typically not regarded as loss process since the reverse reaction (reaction R2 in combination with R3) is an ozone producing reaction.

2.4 RELEVANCE OF GLYOXAL

Apart from formaldehyde and nitrogen oxides whose relevance for tropospheric chemistry is outlined above, glyoxal (CHOCHO) is of importance for this work. In this section, its sources and sinks are briefly described (Sect. 2.4.1) and the concept of the ratio of glyoxal to formaldehyde abundances is introduced (Sect. 2.4.2).

2.4.1 *Sources and sinks of glyoxal*

Glyoxal (CHOCHO), also known by its systematic name *ethanedial*, is the smallest and most important α-dicarbonyl in the atmosphere and belongs, like HCHO, to the large group of OVOCs. As described below, also glyoxal plays an important role in tropospheric chemistry as it is a rather stable but short-lived intermediate

product of the oxidation pathways of many NMVOCs containing two or more carbon atoms (Calvert et al., 2002; Myriokefalitakis et al., 2008). Since glyoxal is highly water-soluble, it is supposed to play an important role in the formation of secondary organic aerosol (SOA) (Fu et al., 2008; Myriokefalitakis et al., 2008; Mellouki et al., 2015). Hence, also glyoxal allows valuable insights into tropospheric VOC chemistry. Like HCHO, glyoxal has both primary and secondary sources, whereby the primary sources are supposed to be ten times less important for glyoxal compared to HCHO (Volkamer et al., 2005a; Myriokefalitakis et al., 2008).

Direct emissions of glyoxal in considerable amounts are mainly related to biomass burning events and biofuel combustion (Fu et al., 2008; Stavrakou et al., 2009a). To smaller extents, also fossil fuel combustion and other anthropogenic activities such as chemical industries emit glyoxal (Stavrakou et al., 2009a).

Despite these primary sources, secondary formation from precursor VOCs is globally the major glyoxal source in the atmosphere since it is an intermediate product of the oxidation pathways of many NMVOCs. Modelling studies suggest that overall biogenic NMVOCs such as for example isoprene or terpenes (including monoterpenes) are the main precursors of glyoxal. Hereby, the oxidation of isoprene emitted by vegetation is commonly regarded as the largest individual source contributing with about 50 % to the global glyoxal abundance[10] (e.g. Fu et al., 2008; Myriokefalitakis et al., 2008).

Besides biogenic NMVOCs, also anthropogenic precursors can play an important role. Such mostly, but not solely, anthropogenic precursors are for instance acetylene, aromatics, alkenes or monoterpenes (e.g. Volkamer et al., 2001; Fu et al., 2008; Myriokefalitakis et al., 2008; Nishino et al., 2010; Chan Miller et al., 2016). The oxidation of acetylene has a high glyoxal yield and due to its rather long atmospheric lifetime of about 18 days it offers a glyoxal source also in the free troposphere or at more remote locations (Fu et al., 2008). Apart from direct glyoxal emissions, biomass burning fires, regardless of being of anthropogenic or natural origin, emit a large variety of glyoxal precursor species (Andreae, 2019; Kluge et al., 2020).

Similar to HCHO, the main sinks of glyoxal are photolysis and reaction with OH radicals while other reactions, e.g. with O_3 or NO_3, are not relevant (Volkamer et al., 2005a; Koppmann and Wildt, 2007; Mellouki et al., 2015). Also for glyoxal, photolysis is typically the dominant loss process (Atkinson, 2000; Fu et al., 2008; Mellouki et al., 2015). Since glyoxal is highly water-soluble, another important loss process which, under specific conditions, can even compete with photolysis is the uptake by aerosol particles (in polluted regions) or on water/cloud droplets (Volkamer et al., 2007; Stavrakou et al., 2009a). Hence, glyoxal can act as precursor for SOA. In addition, dry and wet deposition are minor sinks (Volkamer et al., 2005a; Myriokefalitakis et al., 2008).

The lifetimes of glyoxal against reaction with OH and due to photolysis are around 1 day and 2.8 hours, respectively[11]. Neglecting other loss processes, the

10 Also minor anthropogenic isoprene sources exist which are, however, not of major importance.

11 For the reaction against OH, an OH concentration of 1.5×10^6 molec cm^{-3} and a rate constant at 298 K of 8.76×10^{-12} cm^3 s^{-1} molec^{-1} was assumed (Feierabend et al., 2008; Myriokefalitakis et al., 2008). The photolysis rate was assumed as 1×10^{-4} s^{-1} representative for an SZA of 30° (Volkamer et al., 2005b).

photochemical lifetime of glyoxal can be estimated to be about 2.5 hours and thus is shorter than the one of HCHO (Mellouki et al., 2015; Seinfeld and Pandis, 2016). These lifetimes can vary under real conditions as the availability of solar radiation and OH radicals changes. As described above, also uptake on droplets and aerosol forming SOA might be competitive under specific conditions and thus can affect the atmospheric lifetime of glyoxal.

In the boundary layer, glyoxal mixing ratios range from a few ppt to low values of ppb, whereby higher abundances are typically observed in highly polluted urban regions and areas with high biogenic activity while lower abundances are found in more remote and pristine environments. For instance, mixing ratios have been reported to be of the order of a few ppt to a few tens of ppt in pristine marine environments (Mahajan et al., 2014; Lawson et al., 2015). In rather remote continental regions, glyoxal abundances are of the order of several tens of ppt to a few hundreds of ppt (Sinreich et al., 2007; Fu et al., 2008; Myriokefalitakis et al., 2008). Kluge et al. (2020) observed glyoxal mixing ratios of up to about 250 ppt in biomass burning regions over the Amazon rainforest. In a tropical rainforest region in south-east Asia with large isoprene emissions, MacDonald et al. (2012) found maximum mixing ratios reaching 1.6 ppb while the average values were in the range of several hundreds of ppt. In spring 2003, Volkamer et al. (2005a) measured mixing ratios of several hundreds of ppt reaching up to 1.82 ppb during day time under very polluted conditions in Mexico City.

2.4.2 *The ratio glyoxal to formaldehyde (R_{GF})*

A large number of different VOCs is emitted into the atmosphere by both natural and anthropogenic processes. Since it is not possible to measure the abundance of each and every individual species, smaller and more stable molecules formed during the degradation of the different VOCs are valuable proxies for the VOC composition and their different sources and degradation pathways (DiGangi et al., 2012). As described above, such molecules are for instance HCHO and glyoxal both being mainly formed from precursor VOCs of both biogenic and anthropogenic origin. In the lower troposphere, both species occur at sufficiently high concentrations such that they are detectable by in-situ and remote sensing measurement techniques. Further, their lifetimes are long enough to be measurable but short enough to link them to their emission sources. While HCHO is formed during the oxidation of practically every VOC (including the simplest one, i.e. methane), glyoxal originates from the oxidation pathways of a smaller subset of VOCs containing two or more carbon atoms, in particular alkenes, isoprene and monoterpenes (compare above), since it is a more complex molecule (Kaiser et al., 2015). Although their atmospheric lifetimes are comparable and in the order of a few hours, the lifetime of glyoxal is about 30 % shorter mainly caused by its shorter lifetime against photolysis. Altogether, both species have similar sources and sinks with slight differences especially in the sources allowing insights into VOC chemistry if both species are measured simultaneously. For that, typically the ratio of the abundances of glyoxal and HCHO is calculated and investigated.

This ratio is commonly referred to as "ratio glyoxal to formaldehyde" (R_{GF}). Several studies based on both in-situ and remote sensing measurement techniques conducted on ground-based, airborne (research aircraft), and space-borne (satellites) platforms used this ratio to discriminate between anthropogenic and biogenic VOC emissions, as indicator for the dominant VOC species or the prevailing chemical regime (i.e. VOC regime), to investigate the glyoxal yield from isoprene oxidation relative to the one of formaldehyde (i.e. branching ratio), to specify glyoxal precursors, to investigate the anthropogenic influence on photochemistry in more rural regions and to study the sources of glyoxal and the chemical evolution of VOCs in biomass burning plumes (e.g. Wittrock et al., 2006; Vrekoussis et al., 2010; DiGangi et al., 2012; Kaiser et al., 2015; Chan Miller et al., 2017; Alvarado et al., 2020; Kluge et al., 2020; Hong et al., 2022). Depending on the specific atmospheric conditions the R_{GF}-values can range from typically a few percent up to 40 % (MacDonald et al., 2012; Kaiser et al., 2015).

Although some studies suggested that under regimes dominated by biogenic VOCs higher R_{GF}-values are observed than under anthropogenic influence (e.g. by Vrekoussis et al. (2010) based on GOME-2 satellite measurements), other studies based on ground-based in-situ measurements showed higher ratios under the influence of anthropogenic VOC emissions (DiGangi et al., 2012). A more recent study by Kaiser et al. (2015) proposes that an easy attribution to biogenic and anthropogenic precursors based on solely the R_{GF} is probably not possible but rather the type of precursor VOC determines the R_{GF}. In line with satellite retrievals from the more recent OMI instrument (Chan Miller et al., 2014), Kaiser et al. (2015) suggest that in areas associated with high monoterpene emissions the R_{GF}-values are higher (> 3 %) than in regions associated with strong isoprene emissions (< 2.5 %)[12]. Further, the influence of anthropogenic emissions on the R_{GF} is supposed to depend on the specific (local) VOC composition, e.g. aromatics have a high glyoxal yield (Volkamer et al., 2001; Nishino et al., 2010). Overall, for these more recent studies, satellite and ground-based measurements are in better agreement but still slight discrepancies exist and the interpretation of the R_{GF} is still not fully understood. For an overview on existing studies on the R_{GF} and their corresponding values, it is referred to the publications of Kaiser et al. (2015) and Hong et al. (2022).

In this work, the R_{GF} is used to gain insights into the relation between the HCHO and glyoxal abundances on a seasonal scale. Further, attempts are made to identify differences in the chemical processing, i.e. production and degradation, of both species leading to different vertical distributions by comparing the R_{GF} calculated from vertical column density and near-surface concentrations. In addition, differences in the chemical regimes, i.e. different precursor compositions, between the different measurement sites and the different seasons at a given site (especially at ATTO) are addressed.

12 The threshold between both regimes is 4 % in Chan Miller et al. (2014).

Part II

METHODS AND DATA

3

MAX-DOAS AND PROFILE RETRIEVAL

In this work, Multi-AXis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements in the UV/visible spectral range are used to infer trace gas concentration and aerosol extinction vertical profiles as well as their respective total column quantities, namely the vertical column density (VCD) and the aerosol optical depth (AOD). This chapter introduces the basic concepts of this method starting with the DOAS principle itself (Sect. 3.1) including the extension of the classical DOAS approach to passive atmospheric applications relevant for this work. Further, the concept of MAX-DOAS is introduced (Sect. 3.2) followed by a brief description of the profile retrieval algorithm MAPA (Sect. 3.3) which is used to evaluate the former.

3.1 DIFFERENTIAL OPTICAL ABSORPTION SPECTROSCOPY (DOAS)

Introduced by Perner et al. (1976) and Platt et al. (1979), Differential Optical Absorption Spectroscopy (DOAS) has nowadays become a widely used remote sensing technique to retrieve the abundances of various atmospheric trace gases (e.g. NO_2, HCHO or glyoxal). For instance, several recent satellite instruments such as OMI (Levelt et al., 2006) and TROPOMI (Veefkind et al., 2012) use the DOAS method to provide an almost daily view on global trace gas distributions. Besides space-based applications, many ground-based applications of the DOAS method exist, e.g. Pinardi et al. (2013), Wang et al. (2017a), Dimitropoulou et al. (2020), Kumar et al. (2020), Wang et al. (2020), Lauster et al. (2021) and Hong et al. (2022). In this section, an overview on the underlying concepts of DOAS is provided mainly based on information taken from Platt and Stutz (2008) if not stated otherwise and following a similar structure as in Donner (2016).

3.1.1 *Absorption spectroscopy and the Beer-Lambert law*

Generally, absorption spectroscopy is based on the principle that electromagnetic radiation is attenuated when interacting with molecules while passing through a medium. More specifically, the attenuation of light while passing through gas is exploited for the applications in this work. For short light paths, scattering processes can be ignored and the attenuation is solely due to absorption of photons by the gas molecules which occurs as the molecules are exited into a state of higher energy. In the spectral, and thus energy, range relevant for this work, electronic, vibrational and rotational transitions as well as combinations of all three are relevant. The energy which is required to induce these transitions is quantised and characteristic for any molecule. Hence, a fingerprint-like absorption spectrum exists for each molecule – the so-called absorption cross-section $\sigma(\lambda)$. For a homogeneously distributed single absorber along the light path L, the

attenuation of the initial intensity $I_0(\lambda)$ (before passing through the absorber) is described by the Beer-Lambert law

$$I(\lambda, L) = I_0(\lambda) \cdot \exp\left[-\sigma(\lambda) \cdot c \cdot L\right], \tag{3.1}$$

where $I(\lambda, L)$ is the intensity of the light after passing through the gas with a fixed concentration c. For cases where the light path L through the gas is well known and $I(\lambda, L)$ as well as $I_0(\lambda)$ can be measured, e.g. under laboratory conditions, Eq. 3.1 can be rearranged allowing the easy calculation of the concentration of the absorber via

$$c = \frac{\ln\left(\frac{I_0(\lambda)}{I(\lambda,L)}\right)}{\sigma(\lambda) \cdot L} = \frac{\tau(\lambda)}{\sigma(\lambda) \cdot L} \tag{3.2}$$

with

$$\tau(\lambda) = \ln\left(\frac{I_0(\lambda)}{I(\lambda, L)}\right) \tag{3.3}$$

being the so-called optical depth. Figure 3.1 provides a schematic sketch to illustrate the quantities involved in the Beer-Lambert law.

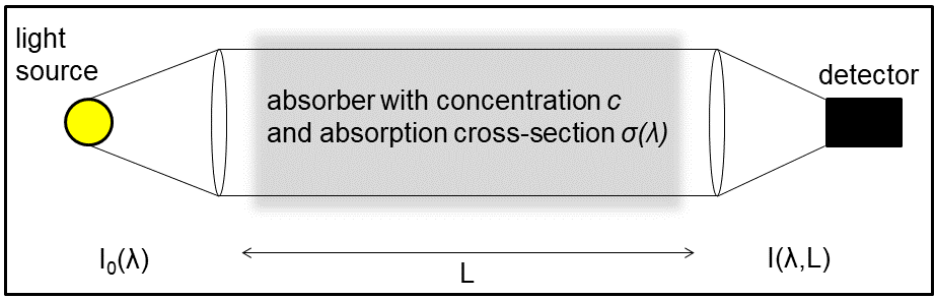


Figure 3.1: Sketch to illustrate the Beer-Lambert law. Light emitted by a light source passes through an absorber, i.e. gas, and the attenuated intensity is measured by a detector (self-created based on Platt and Stutz, 2008). Reproduced with permission from Springer Nature.

3.1.2 *Absorption spectroscopy in the atmosphere*

The application of absorption spectroscopy in the open atmosphere is associated with several implications which have to be considered and several assumptions made in Eq. 3.1 are not valid anymore. Hence, this simple form of the Beer-Lambert law has to be extended.

In particular, the atmosphere is composed of more than one trace gas, all having different concentrations c_i and absorption cross-sections $\sigma_i(\lambda)$. Hence, the exponent in Eq. 3.1 has to be modified accordingly yielding

$$\sigma(\lambda) \cdot c \cdot L \longrightarrow L \cdot \sum_i \sigma_i(\lambda) \cdot c_i. \tag{3.4}$$

Further, in the atmosphere the light paths are significantly longer than for laboratory applications and scattering processes can significantly attenuate the

light intensity besides absorption. Here, two important scattering types have to be considered. Rayleigh scattering describes elastic scattering processes on molecules and partly very small particles whose sizes are much smaller, i.e. less than about one-tenth (Lockwood, 2016), compared to the wavelength of the incident light. For the spectral range relevant for this work, this corresponds to molecules and particles smaller than ca. 50 nm, e.g. molecular nitrogen and oxygen or ultra-fine particles. The scattering probability of Rayleigh scattering and thus the extinction coefficient $\epsilon_R(\lambda)$ is proportional to λ^{-4} (Strutt Lord Rayleigh, 1899). Mie scattering, on the other hand, describes elastic scattering on particles of similar or larger size compared to the wavelength of the incident light, i.e. particles in the order of about 1 µm for sunlight being for instance aerosol particles or small cloud droplets. The corresponding extinction coefficient $\epsilon_M(\lambda)$ shows a weaker wavelength dependence described by $\lambda^{-\gamma}$ with the Ångström exponent γ (compare Sect. 5.2) which depends on the size and shape of the particle (Mie, 1908). The latter is estimated to be 1.3 for the general atmosphere (Hulst, 1981).

In addition, all these atmospheric effects are not constant along the atmospheric light path, e.g. the trace gases are not homogeneously distributed in the atmosphere. Therefore, all effects are functions of the increment along the light path s.

Taking all this together, the Beer-Lambert law (compare Eq. 3.1) for atmospheric applications becomes

$$I(\lambda, L) = I_0(\lambda) \cdot \exp\left[-\int_0^L \left(\epsilon_R(\lambda, s) + \epsilon_M(\lambda, s) + \sum_i \sigma_i(\lambda) \cdot c_i(s)\right) \mathrm{d}s\right] \cdot A(\lambda), \tag{3.5}$$

where the factor $A(\lambda)$ additionally accounts for potential instrumental effects on the measured intensity as well as for the efficiency of scattering processes which scatter light into the line of sight of the instrument and is thus needed for scattered light applications (Richter and Wagner, 2011).

In theory, all these described processes attenuating the incident light which are included in Eq. 3.5 have to be accurately characterised in order to determine the concentration of a specific trace gas of interest. In practice, however, this is almost impossible in the open atmosphere since it is nearly not possible to distinguish the broad-band features caused by scattering processing from the broad-band absorptions of quasi continuum bands of absorbing species. Further, the relative contributions of the different scattering processes amongst each other are hardly discriminable. Lastly, the initial intensity $I_0(\lambda)$ not affected by any atmospheric processes can hardly be determined especially for ground-based applications using (scattered) sunlight.

3.1.3 *The classical DOAS approach*

The DOAS method offers a possibility to overcome the challenges described before by exploiting the fact that scattering processes as well as continuous absorption features and instrumental effects show broad-band or even smooth

wavelength dependencies, while many trace gases exhibit highly structured absorption features. Hence, the fundamental idea of DOAS is to separate the broad- and narrow-band contributions in Eq. 3.5 and to apply the Beer-Lambert law to the narrow-band absorptions only. In the classical DOAS approach, this also includes a separation of the trace gas absorption cross-sections $\sigma_i(\lambda)$ into a broad-band part $\sigma_{i,b}(\lambda)$ and a narrow-band part $\sigma'_i(\lambda)$ as illustrated in Fig. 3.2, yielding

$$\sigma_i(\lambda) = \sigma_{i,b}(\lambda) + \sigma'_i(\lambda). \tag{3.6}$$

This separation is typically achieved by applying a high-pass filter. Naturally, the definition of broad- and narrow-band features depends on the wavelength range of interest and the width of the absorption bands of the trace gas of interest.

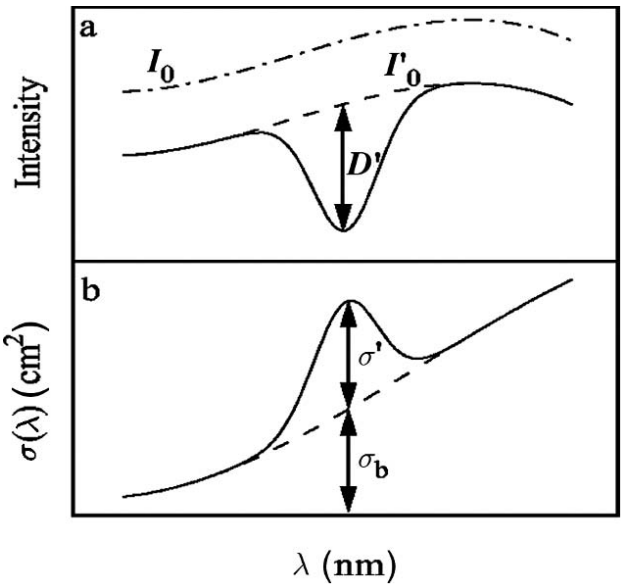


Figure 3.2: The basic idea of the classical DOAS approach: The intensity I_0 and the cross-section σ are separated into a broad- and a narrow-band part (taken from Platt and Stutz, 2008). Reproduced with permission from Springer Nature.

Using these separations in Eq. 3.5 yields

$$I(\lambda, L) = I'_0(\lambda, L) \cdot \exp\left[-\int_0^L \sum_i \sigma'_i(\lambda) \cdot c_i(s) \mathrm{d}s\right], \tag{3.7}$$

where

$$I'_0(\lambda, L) = I_0(\lambda) \cdot \exp\left[-\int_0^L \left(\epsilon_R(\lambda, s) + \epsilon_M(\lambda, s) + \sum_i \sigma_{i,b}(\lambda) \cdot c_i(s)\right) \mathrm{d}s\right] \cdot A(\lambda) \tag{3.8}$$

is the "broad-band intensity" which describes the attenuation of the initial intensity I_0 solely caused by broad-band effects. This "broad-band intensity" is approximated by a low order closure polynomial in most practical applications.

Similar to classical absorption spectroscopy (compare Eq. 3.3), a differential optical depth $\tau'(\lambda)$ (compare D' in upper part of Fig. 3.2) can be defined as

$$\tau'(\lambda) = \ln\left(\frac{I'_0(\lambda, L)}{I(\lambda, L)}\right) = \int_0^L \sum_i \sigma'_i(\lambda) \cdot c_i(s)\mathrm{d}s. \tag{3.9}$$

A key quantity in DOAS applications is the so-called slant column density (SCD) defined as the concentration of a given absorber c_i integrated along the effective light path L:

$$SCD_i = \int_0^L c_i(s)\mathrm{d}s. \tag{3.10}$$

With this quantity, Eq. 3.9 becomes

$$\tau'(\lambda) = \ln\left(\frac{I'_0(\lambda, L)}{I(\lambda, L)}\right) = \sum_i \sigma'_i(\lambda) \cdot SCD_i \tag{3.11}$$

which shows that the measurement of light that got partly absorbed by atmospheric trace gases and subsequently comparing its intensity with $I'_0(\lambda, L)$ allows the retrieval of the SCDs of the involved trace gases. Another advantage of this approach is the possibility to detect trace gases with rather weak absorptions as long as the resulting differential optical depths $\tau'(\lambda)$ are sufficiently high, i.e in the order of 10^{-4} to 10^{-3}.

3.1.4 *Passive DOAS in this work*

There are two basic application types of the DOAS method for atmospheric measurements. The first one are active DOAS measurements which use artificial light sources such as halogen-arc lamps or, in more recent applications (e.g. Nasse et al., 2019), laser driven light sources. Actually, most of the early atmospheric applications of the DOAS method have been active measurements in terms of Long-Path (LP-) DOAS measurements (e.g. Perner et al., 1976). Also in recent literature several examples of LP-DOAS applications exist (e.g. Liao et al., 2011; Stutz et al., 2011; Nasse et al., 2019). Typically, the DOAS retrievals of such measurements still use the classical DOAS concept described above by applying high-pass filters to both measured spectra and absorption cross-sections. The major advantages of such active DOAS measurements are the well characterised spectral properties of the light source (intensity without atmospheric effects can be determined easily), the well constrained atmospheric light paths and the fact that measurements are also possible during night. However, the light paths are often limited to a few kilometres (if the instrument's size is kept at a reasonable level) and thus very weak absorbers might not yield optical depths that are sufficiently large. Further, such active setups are often rather large, complex and expensive, since for example retro-reflectors are used at one end of the light beam to double the length of the light paths. In many cases, e.g. for mobile applications or for measurements at several partly remote locations (like in this work) more compact and cheap setups are desirable. This can be achieved by the second major application type of the DOAS method, namely passive DOAS measurements.

Such passive DOAS applications mostly use the sun as light source[1], whereby direct sun observations (e.g. Herman et al., 2009) and scattered sunlight measurements are possible. The latter are relevant for this work. For instance, to achieve long tropospheric light paths (reaching up to more than 10 km under low aerosol and clear conditions) scattered sunlight can be measured under very slant elevation angles. Besides the limitation that these measurements are only possible during day as they rely on sunlight, there are several other implications related to passive DOAS which are further elaborated in Sect. 3.1.5. The major aspect to be considered, however, is that the solar spectrum is not spectrally smooth but highly structured. The importance of this circumstance is further assessed below mainly based on the study of Hönninger et al. (2004) while the discussion is again structured similarly to Donner (2016).

The light emitted by the sun can in a reasonable approximation be described as the emission of a black body radiator with a temperature of about 5800 K. This continuous spectrum, however, is superposed by a vast number of strong absorption lines, the so-called Fraunhofer lines, originating from selective absorptions by atoms within the solar photosphere. These lines are especially prominent in the UV/visible spectral range which is relevant for this work and notably stronger than most trace gas absorptions occurring in Earth's atmosphere. The latter have optical depths in the order of 10^{-4} to 10^{-3}, while Fraunhofer lines have absorptions of up to 30 % at typical DOAS spectral resolution. Hence, the Fraunhofer lines have to be removed during the DOAS analysis to retrieve the much weaker absorptions of atmospheric trace gases. This is achieved by including a so-called Fraunhofer Reference Spectrum (FRS) in the DOAS analysis. For off-axis DOAS applications (compare Sect. 3.2), typically a zenith spectrum (i.e. measured at 90°) is used as FRS to minimise the atmospheric trace gas absorptions affecting the reference spectrum. This spectrum might be a carefully selected fixed single spectrum measured around noon (small SZA) to further reduce the trace gas absorptions in the reference spectrum or a new FRS for each elevation sequence of a MAX-DOAS application (compare Sect. 3.2). The latter is called sequential FRS (or sequential reference) in the following and used in this work (compare for instance Sect. 9.1). Since also the FRS was measured at Earth's surface, the light detected by the instrument underwent scattering processes and was partly absorbed by trace gases although the choice of the FRS intends to minimise the latter. However, the FRS is measured under a different viewing direction than the spectrum to be analysed which in consequence leads to differences in the attenuation by both absorption and scattering. Using the SCD introduced in Eq. 3.10 and taking into account all considerations made here, Eq. 3.5 can be reformulated as

$$I(\lambda, L) = I_{\mathrm{FRS}}(\lambda, L) \cdot \exp\left[-\sum_i \sigma_i(\lambda) \cdot (SCD_i - SCD_{\mathrm{FRS},i}) + P(\lambda)\right]. \quad (3.12)$$

Here, $I_{\mathrm{FRS}}(\lambda, L)$ is the intensity of the FRS while SCD_i and $SCD_{\mathrm{FRS},i}$ describe the slant column densities, i.e. trace gas absorptions, affecting the measured spectrum and the FRS, respectively. $P(\lambda)$ accounts for broad-band differences between

1 A few exceptions exist where moonlight or starlight is used (e.g. Wagner et al., 2000).

measured spectrum and FRS caused by differences in the scattering processes as well as instrumental effects. Note that here the trace gas absorptions are again described by full absorption cross-sections $\sigma_i(\lambda)$ which is the state of the art approach in modern DOAS applications and different to the classical DOAS approach described above. Hence, $P(\lambda)$ also accounts for broad-band differences in the trace gas absorptions. Similar to above, again a differential optical depth can be introduced as

$$\tau'(\lambda) = \ln\left(\frac{I_{\mathrm{FRS}}(\lambda, L)}{I(\lambda, L)}\right) = \sum_i \sigma_i(\lambda) \cdot (SCD_i - SCD_{\mathrm{FRS},i}) + P_m(\lambda). \tag{3.13}$$

This optical depth $\tau'(\lambda)$ is differential in a way as it describes the differential absorption with respect to the FRS. More precisely, $SCD_i - SCD_{\mathrm{FRS},i}$ represents the difference in the integrated trace gas concentrations along the effective light paths of the measured spectrum and the FRS. This quantity is commonly referred to as differential slant column density (dSCD) and the main quantity derived from passive DOAS applications using scattered sunlight. Applying this to Eq. 3.13 yields

$$\tau'(\lambda) = \ln\left(\frac{I_{\mathrm{FRS}}(\lambda, L)}{I(\lambda, L)}\right) = \sum_i \sigma_i(\lambda) \cdot dSCD_i + P_m(\lambda), \tag{3.14}$$

where $P_m(\lambda)$ is a low order closure polynomial of order m which accounts for all broad-band effects regardless of being of atmospheric or instrumental nature. Equation 3.14 has to be solved to infer atmospheric trace gas abundances from the measured quantities $I_{\mathrm{FRS}}(\lambda, L)$ and $I(\lambda, L)$, while the absorption cross-sections σ_i are taken from literature. In practice, additional effects such as possible intensity offsets, caused for instance by spectrometer stray-light as well as detector non-linearities, or slight inaccuracies in the spectral calibrations of the utilised spectrometers (shifts and stretches) have to be considered (for details refer for instance to Sect. 6.5 and 9.1). All this is accounted for by the DOAS algorithm used to solve Eq. 3.14 which is based on a combination of a linear-least squares fit and a Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) which is a non-linear least-squares fitting approach to minimise the difference between measurement and simulation. In this work, version 3.4.4 of the *QDOAS* software package (Danckaert et al., 2017) is used to perform the DOAS fit. Finally, the brief description here is closed with a reference to existing literature, e.g. Kraus (2006), Platt and Stutz (2008) and Danckaert et al. (2017), providing more in-depth discussions of the mathematical implementation of the DOAS fit.

3.1.5 *Implications using passive DOAS*

The use of (scattered) sunlight as light source for atmospheric DOAS applications is associated with several implications besides the Fraunhofer lines discussed before. In particular, only elastic scattering processes such as Mie and Rayleigh scattering have been considered so far, however, also inelastic scattering processes affect the measured intensities. Moreover, the implicit assumption was that the SCDs are independent of the wavelength. Lastly, all considerations so far have

been performed in high resolution (HR) space and did not account for the fact that the intensities in Eq. 3.14 are measured by spectrometers with spectral resolutions much coarser than the highly structured Fraunhofer lines and atmospheric absorptions. Although none of these simplifications affects the DOAS method in a fundamental way, they have to be accounted for in order to perform an accurate and precise DOAS retrieval. This section aims to briefly describe the implications of the mentioned aspects and to outline how these are accounted for in the DOAS retrievals in this work.

3.1.5.1 *Ring effect*

First observed by Grainger and Ring (1962), the so-called Ring effect describes the "filling-in" of the solar Fraunhofer lines. According to today's understanding, it is caused by inelastic rotational Raman scattering mainly on molecular nitrogen (N_2) and oxygen (O_2) (Solomon et al., 1987b; Bussemer, 1993; Wagner et al., 2009). Since inelastic scattering alters the energy and thus the wavelength of a photon, the scattered photons of a given initial wavelength are transferred to a different wavelength adding an additional intensity to the latter. As more photons (higher intensity) exist at wavelengths outside major Fraunhofer lines it is more likely that a photon will be transferred to a wavelength of a Fraunhofer line than the opposite. Effectively, this leads to a "filling-in" of the solar Fraunhofer lines and consequently their optical depths are changed (decreased). Naturally, the strength of this effect depends on the atmospheric light path and thus on the viewing geometry (i.e. SZA and RAA as described below). Therefore, it does not cancel out by the usage of an FRS and the complete removal of the Fraunhofer lines is not possible. Since the optical depth changes of the Fraunhofer lines caused by the Ring effect are of the order of a few percent, a very accurate correction of this effect is required as the absorptions caused by atmospheric trace gases can be more than one order of magnitude smaller (10^{-4} to 10^{-3}) (Hönninger et al., 2004). To achieve this, typically so-called Ring spectra are included in the DOAS evaluations of scattered sunlight spectra (Solomon et al., 1987b; Bussemer, 1993; Wagner et al., 2009), whereby different approaches exist.

Some groups in the DOAS community include the ratio of a synthetic simulated Raman spectrum and a simulated solar spectrum after convolving both spectra with the instrumental slit function as pseudo absorber (PSA) in the DOAS fit (Alvarado et al., 2020; Pinardi, 2020). In this work, however, two PSAs are derived from the so-called universal FRS based on Bussemer (1993) and Wagner et al. (2009) exploiting the known Raman response functions of N_2 and O_2 and included in the DOAS retrievals (compare for example Tables 9.1 to 9.5). Here, the universal FRS is the measured solar spectrum (typically at 90° elevation) used for the spectral calibration (compare Sect. 8.1 and 9.1). The first PSA is, in a broad-band sense, spectrally flat and contains only highly structured Ring contributions reflecting the probability for rotational Raman scattering, while the second Ring spectrum has a different broad-band wavelength dependence as suggested by Wagner et al. (2009). For further details on the Ring correction, refer to Wagner et al. (2009).

There are two further aspects which should be briefly mentioned. Firstly, also absorption features of atmospheric trace gases can be "filled-in" if the inelastic scattering event happens after the absorption (Lampel et al., 2015). This is mainly relevant for trace gases with large optical depths, such as O_3 at short wavelengths or NO_2 in very polluted scenes, in the retrieval of trace gases with very small optical depths. A way to account for this are Raman or Ring cross-sections for such absorbers. In Chapt. 6, such Raman cross-sections for NO_2 and O_3 are tested for the DOAS retrieval of HCHO. Secondly, Lampel et al. (2015) reported on spectrally shifted and "filled-in" Fraunhofer structures caused by vibrational Raman scattering. So far, vibrational Raman scattering on N_2 and O_2 molecules has been neglected[2] due to its small amplitude (Wagner et al., 2009). However, Lampel et al. (2015) showed that for the DOAS retrievals of species retrieved in wavelength intervals above ca. 400 nm vibrational Raman scattering might be relevant. Since the spectral shift towards longer wavelengths is in the order of tens of nanometres, this effect is not relevant for retrievals in the UV (e.g. HCHO) using scattered sunlight. Therefore, vibrational Raman scattering is not further considered in this work.

3.1.5.2 *Puķīte terms*

The so-called Puķīte terms correct for non-linear effects which play a significant role for medium and strong trace gas absorptions (Puķīte et al., 2010). These non-linearities are caused by the fact that the trace gas SCDs are wavelength dependent even within a given fitting interval as also the effective light path through the atmosphere is wavelength dependent which was implicitly neglected for the derivation of Eq. 3.14. There are two effects causing the wavelength dependence of the SCD. Firstly, the length of the effective atmospheric light path is wavelength dependent itself, since the probability for Rayleigh scattering (and thus extinction) is proportional to λ^{-4} (compare Sect. 3.1.2) leading to increasing scattering probabilities for photons with decreasing wavelengths. But also other scattering, such as Mie scattering, and reflection processes show broad-band wavelength dependencies (Puķīte et al., 2010). Secondly, the length of the effective light path depends on the absorption strengths of trace gases. Here, longer light paths lead to stronger absorptions and consequently to lower intensities, while shorter light paths have weaker absorptions leading to higher intensities. Consequently, the average light path becomes more representative for these shorter light paths which contribute more to the measured intensities (Puķīte et al., 2010). Since this second effect is caused by absorptions, it can also lead to narrow-band spectral features (Puķīte et al., 2010). Hence, the SCD also depends on the trace gas abundance itself.

While these non-linearities are often not accounted for in the retrievals of very abundant and/or strong absorbers, they are of particular importance for the retrievals of less abundant and/or weak absorbers where the wavelength dependencies caused mainly by the strong absorbers lead to high detection limits or poor results if they absorb in the same wavelength range. This is especially

2 For liquid water over clear ocean water, this effect is already sometimes considered.

important in the UV spectral range, where also HCHO is retrieved, as the Rayleigh scattering probability increases rapidly with decreasing wavelength. Further, the O_3 cross-section shows strong absorption features in that spectral region. The latter two aspects lead to a strong wavelength dependence of the SCDs, especially for O_3.

Puķīte et al. (2010) introduced a method to account for these effects in the DOAS retrieval itself. For this, a wavelength (λ) and vertical optical depth (ν) dependent representation of the SCD is expanded as Taylor series up to the first-order. Using the relation that the optical depth τ_a of a given absorber a is the product of its SCD and cross-section σ_a (compare Eq. 3.11), this yields

$$\tau_a(\lambda, \nu) \approx S_0 \cdot \sigma_a(\lambda) + S_\lambda \cdot \lambda\sigma_a(\lambda) + S_S \cdot \sigma_a^2(\lambda). \tag{3.15}$$

Here, S_0, S_λ and S_S are the "SCDs" of the contributing Taylor terms, while

$$\sigma_{\text{lin},a}(\lambda) = \lambda \cdot \sigma_a(\lambda) \tag{3.16}$$

and

$$\sigma_{\text{sq},a}(\lambda) = \sigma_a^2(\lambda), \tag{3.17}$$

are the so-called first-order Puķīte terms which can be included in the DOAS fit as PSAs. In the following, these two terms are referred to as linear and square "higher order" or Puķīte terms, respectively, and calculated according to Eq. 3.16 and 3.17 from the standard absorption cross-section σ_a of the respective absorber a, where λ is the wavelength. Note that considering more than the first-order Taylor terms for Eq. 3.15 allows in principle the derivation of further (higher order) Puķīte terms (compare Sect. 6.4.3). More details on the theory behind these terms are summarised in Puķīte et al. (2010) and Puķīte and Wagner (2016).

3.1.5.3 *Saturation and solar I_0-effect*

There are two subtle implications with the definition of the optical depth in Eq. 3.14. Both are related to the fact that the considerations above were performed in HR space, while the measurement of scattered sunlight spectra corresponds to a convolution (spectral smoothing) of the natural line widths of both the solar and atmospheric absorptions with the instrumental slit function of the form

$$I^*(\lambda) = F * I(\lambda) = \int I(\lambda) K(\lambda_0 - \lambda)\, \mathrm{d}\lambda. \tag{3.18}$$

$I^*(\lambda)$ describes the convolved solar spectrum, $F(\lambda)$ the convolution function and K is the convolution kernel, i.e. the instrumental slit function. Hence, for the spectral DOAS retrieval also the trace gas absorption cross-sections have to be convolved with this slit function. However, convolution and logarithm do not commute, i.e.

$$\tau'_{\text{meas}}(\lambda) = \ln\left(\frac{I_{\text{FRS}}(\lambda, L)}{I(\lambda, L)} * F(\lambda)\right) \neq \ln\left(\frac{I_{\text{FRS}}(\lambda, L)}{I(\lambda, L)}\right) * F(\lambda) = \tau'_{\text{model}}(\lambda) \tag{3.19}$$

with

$$\tau'_{\text{model}}(\lambda) = \sum_i \sigma_i(\lambda) * F(\lambda) \cdot dSCD_i + P_m(\lambda) \tag{3.20}$$

and a non-linearity between the measured (differential) optical depth $\tau'_{\text{meas}}(\lambda)$ and the modelled one $\tau'_{\text{model}}(\lambda)$ exists (Wagner et al., 2002a). This effect is known as saturation effect and the inaccuracies related to this effect increase with absorption strength, i.e. optical depth. Further, this saturation effect shows that the DOAS principle is mostly valid for small absorptions and is a limitation of the DOAS method itself as discussed in Sect. 6.4.

The second effect is commonly known as solar I_0-effect and occurs because the reference absorption cross-sections are measured in the laboratory with a smooth light source, while the atmospheric spectra are measured with an instrument with low spectral resolution and a highly structured light source, i.e. the sun. In scattered sunlight DOAS applications, the strong Fraunhofer lines are removed by calculating the (logarithmic) ratio of a measured spectrum and an FRS (compare Sect. 3.1.4) in order to be able to retrieve the comparably weak trace gas absorptions. As outlined above, both spectra have been measured at the rather low spectral resolution of the instrument and therefore have been spectrally smoothed (convolved) by its slit function before calculating their ratio. In that way, the spectra and not their (logarithmic) ratio are smoothed (Aliwell et al., 2002). Since also calculating the ratio and convolution do not commute, a complete removal of the Fraunhofer lines is not possible. Thus, the fitting of convolved laboratory cross-sections, which do not include this effect, to the optical depths measured by a low resolving instrument will cause systematic residual structures in the spectral fit (Danckaert et al., 2017). The related inaccuracies depend on the absorption strength and the relation between the width of the instrumental slit function and the width of the structured features in the light source and the absorption cross-section (Danckaert et al., 2017). For many atmospheric absorbers this effect can be neglected, but especially for strong absorbers like NO_2 or O_3 it can have a considerable influence on the results of weak and/or less abundant absorbers like HCHO.

The I_0-effect was described by Aliwell et al. (2002) and a correction including the calculation of synthetic spectra during the convolution process is proposed and commonly used. Following Aliwell et al. (2002), the I_0-corrected low resolution (LR) cross-section of an absorber can be expressed as

$$\sigma_{\text{LR,I}_0} = -\frac{1}{SCD} \ln\left(\frac{\int I_0(\lambda) e^{-SCD \cdot \sigma_{\text{HR}}(\lambda)} K(\lambda_0 - \lambda)\, \mathrm{d}\lambda}{\int I_0(\lambda) K(\lambda_0 - \lambda)\, \mathrm{d}\lambda}\right) \tag{3.21}$$

where I_0 is the HR solar spectrum, K is the convolution kernel, i.e. the instrumental slit function, and SCD is a typical slant column density of the given absorber. Obviously, the correction depends slightly on the choice of the typical SCD, however, several studies, including this work (compare Sect. 6.2.1), suggest that this choice is not very critical (e.g. Wagner, 1999). The described approach has two effects. Firstly, it weights the absorptions at the individual wavelengths with the corresponding intensities of I_0. Secondly, it considers non-linear effects

of the absorptions by the exponential term in the numerator. In that way, LR absorption cross-sections that take into account the solar I_0-effect are generated (Aliwell et al., 2002).

As discussed by Puķīte et al. (2021), Eq. 3.21 is only valid for one single absorber with constant SCD since otherwise the absorptions of the individual trace gases lead to multiplicative effects in the exponent. Thus, in principle for cases with more than one absorber, the different cross-sections cannot be convolved separately. A similar implication exists for absorbers for which Puķīte terms (compare Sect. 3.1.5.2) are used. Following Puķīte et al. (2021) and inserting the Taylor series expansion of the wavelength and vertical optical depth dependent representation of the SCD (compare Sect. 3.1.5.2) in Eq. 3.21, one obtains

$$\sigma_{\mathrm{LR,I_o}} = -\frac{1}{S}\ln\left(\frac{\int I_0(\lambda)e^{-\left(S_0\sigma_{\mathrm{HR}}(\lambda)+S_\lambda\lambda\sigma_{\mathrm{HR}}(\lambda)+S_\sigma\sigma^2_{\mathrm{HR}}(\lambda)\right)}K(\lambda_0-\lambda)\,\mathrm{d}\lambda}{\int I_0(\lambda)K(\lambda_0-\lambda)\,\mathrm{d}\lambda}\right), \quad (3.22)$$

with

$$S = S_0 + S_\lambda\lambda + S_\sigma\sigma_{\mathrm{HR}} \quad (3.23)$$

being the Taylor series representation of the wavelength and vertical optical depth dependent SCD. These three Taylor series terms in the exponential term in the numerator have a multiplicative effect similar to the case with more than one absorber and the integral cannot be split up. Consequently, the "normal" cross-section of an absorber and the corresponding higher order terms can in principle not be convolved separately.

Therefore, Puķīte et al. (2021) propose a simplified convolution for absorbers for which Puķīte terms are used. In case of an absorption weak enough to justify a Taylor series expansion of the exponent in Eq. 3.22 of the form

$$e^x \approx 1 + x \text{ (for } x \text{ close to 0)}, \quad (3.24)$$

the summands in the resulting equation can be integrated separately. Hence, the I_0-corrected terms for a weak absorption limit are defined as

$$\psi_{\mathrm{LR,weighted}} = \frac{\int I_0(\lambda)\psi_{\mathrm{HR}}(\lambda)K(\lambda_0-\lambda)\,\mathrm{d}\lambda}{\int I_0(\lambda)K(\lambda_0-\lambda)\,\mathrm{d}\lambda}, \quad (3.25)$$

where ψ can be either a "normal" cross-section σ or a Puķīte term according to Eq. 3.16 and 3.17. This convolution type is called "intensity weighted convolution" in the following and basically satisfies the first aim of the I_0-correction as it weights the absorptions at the individual wavelengths with the corresponding intensities of I_0. Parts of non-linearities introduced by strong absorptions might be accounted for by the usage of the square Puķīte term for strong absorbers (Puķīte et al., 2021). As described in Sect. 6.2.1, this convolution type is used for the cross-sections of O_3 and NO_2 as well as their respective Puķīte terms in this work. For further details on both saturation and solar I_0-effect and the correction of the latter, it is referred to existing literature, for example, Wagner et al. (2002a), Aliwell et al. (2002), Danckaert et al. (2017) and Puķīte et al. (2021).

3.2 MULTIAXIS-DOAS (MAX-DOAS)

Historically, the first applications of passive DOAS measurements using scattered sunlight have been zenith sky observations, also referred to as zenith scattered sunlight DOAS (e.g. Noxon, 1975; Noxon et al., 1978; Solomon et al., 1987a,b; Kreher et al., 1997; Richter et al., 1999; Aliwell et al., 2002). These measurements are mainly sensitive to stratospheric absorbers, while for studies on tropospheric absorbers mostly active DOAS measurements were used in those days (e.g. Perner et al., 1976; Perner and Platt, 1979; Platt et al., 1979; Mount, 1992; Harder et al., 1997; a comprehensive overview is provided in Hönninger et al., 2004). Further, the radiative transfer modelling needed for the determination of vertical column densities (compare below) was back then best understood for zenith scattered sunlight (Solomon et al., 1987b). Only in the early 90's, Sanders et al. (1993) first introduced measurements of scattered sunlight in off-axis geometry (i.e. measurements at viewing directions other than zenith) to study (still stratospheric) OClO over Antarctica. This gave rise to various studies using off-axis measurements of scattered sunlight to also study the tropospheric composition (e.g. Miller et al., 1997; Hönninger and Platt, 2002; Bobrowski et al., 2003; Hönninger et al., 2004; Heckel et al., 2005). Nowadays, passive DOAS applications using spectra of scattered sunlight measured at several off-axis viewing directions are a well-established and state of the art technique in ground-based remote sensing (e.g Wagner et al., 2011; Ma et al., 2013; Pinardi et al., 2013; Wang et al., 2017a; Kumar et al., 2020; Wang et al., 2020; Hong et al., 2022) and are known as Multi-AXis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements. Since the results in this work are obtained using MAX-DOAS observations, this section briefly introduces the measurement geometry, relevant quantities and the concept of air mass factors (AMFs).

3.2.1 *Measurement geometry*

In contrast to zenith sky DOAS measurements, MAX-DOAS measurements make use of scattered sunlight spectra recorded at multiple, typically between 1° and 90°, off-axis viewing directions (Hönninger et al., 2004). The viewing geometry of such measurements is described by three angles which are visualised in Fig. 3.3:

- The elevation angle α is defined as the instrument's viewing direction relative to the horizontal line, i.e. at $\alpha = 0°$ the instrument points to the (true) horizon and at $\alpha = 90°$ towards zenith (ca. 15° in Fig. 3.3).
- The solar zenith angle θ (SZA) is defined as the sun position relative to the zenith, i.e. at $\theta = 90°$ the sun is at the horizon (ca. 70° in Fig. 3.3).
- The relative azimuth angle ϕ (RAA) is defined as the azimuthal viewing direction of the instrument relative to the azimuthal sun position, i.e. at $\phi = 0°$ the instrument points in the direction of the sun. In Fig. 3.3, for simplicity reasons an RAA of 180° was chosen, i.e. the sun is located "behind" the instrument.

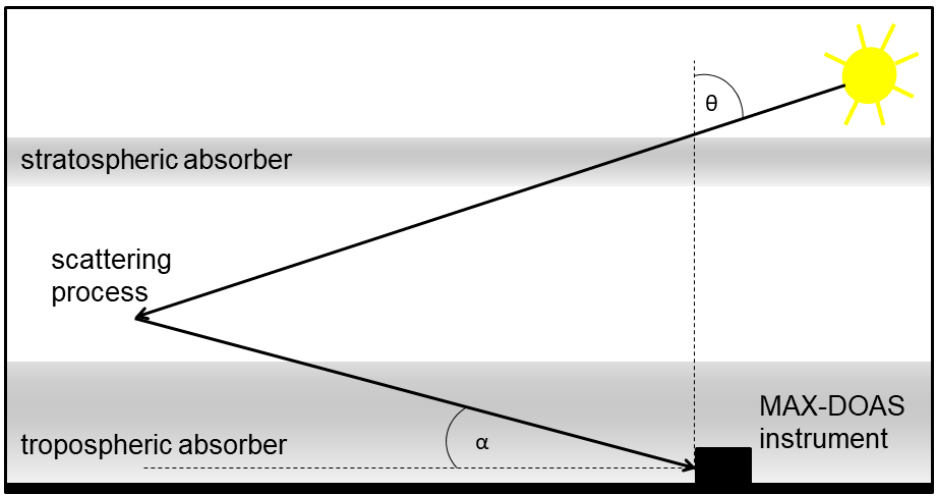


Figure 3.3: Viewing geometry of MAX-DOAS measurements and relevant angles as described in the text (self-created based on Hönninger et al., 2004).

Since the relevant scattering process which scatters the photon into the instrument's telescope typically takes place in the troposphere (below stratospheric absorbers), the sensitivity of MAX-DOAS measurements to stratospheric absorbers is nearly independent of the elevation angle, while the sensitivity to tropospheric absorbers strongly depends on the elevation angle (Hönninger et al., 2004). As can be seen in Fig. 3.3, this is because the light path through the tropospheric trace gas layer strongly changes as the elevation angle changes with lower elevations leading to longer light paths. At the same time, the light path through the stratosphere is almost unchanged. As a consequence of Eq. 3.1, a longer light path through an absorber with a given concentration leads to a stronger absorption and thus higher sensitivity. The sensitivity to stratospheric absorbers, on the other hand, mainly depends on the SZA, whereby higher sensitivity is given for large SZAs for the same reasons as before.

Given the strong elevation dependence of the sensitivity to tropospheric absorbers, naturally, the accurate and precise knowledge of the elevation angle is crucial. For instance, measurements during the CINDI-2 campaign (e.g. Kreher et al., 2020; Donner et al., 2020; Wang et al., 2020) showed that an error of only 1° in the telescope's elevation close to 0° elevation can lead to an error of around 20 % in the retrieved NO_2 dSCD (Donner et al., 2020). While state of the art mechanics controlling the elevation angle of a MAX-DOAS instrument provide high accuracy and precision of the elevations relative to each other, it is more difficult to absolutely calibrate the elevation angles, i.e. to determine the true horizontal viewing direction $\alpha = 0°$. Donner et al. (2020) present and inter-compare different methods for the calibration of the elevation angles of MAX-DOAS instruments to notably improve the results of MAX-DOAS measurements.

3.2.2 *Quantities retrieved from MAX-DOAS measurements*

The measurement strategy for such MAX-DOAS measurements is straight forward. Typically, measurements of scattered sunlight are repetitively taken for a fixed set of elevation angles between 1° and 90°. Such a set of elevation angles is referred to as (elevation) sequence in the following. As discussed in the previous section, the sensitivity to tropospheric absorbers or, more particularly, absorbers close to the surface increases with decreasing elevation angle. Therefore, the elevation sequences typically include many measurements at low elevations leading to long light paths and thus high sensitivities, while only a few measurements at elevations above 10° are included which help to further constrain the profile retrieval (compare Sect. 3.3). The measurements at 90° usually serve as FRS for the DOAS retrieval as outlined in Sect. 3.1.4, whereby two approaches are possible. The first option is to carefully select a single fixed 90° measurement around noon (small SZA) or under specifically clean conditions (low trace gas absorptions) and is typically used for car-MAX-DOAS applications (e.g. Wagner et al., 2010; Ibrahim et al., 2010; Donner, 2016; Shaiganfar et al., 2017; Lange et al., 2023) or cloud classification approaches from MAX-DOAS measurements (Wagner et al., 2014, 2016). The second option is to use the 90° measurement of a given sequence as FRS for the DOAS analysis of the off-axis measurements of this sequence. This option is called sequential FRS and the advantage of this approach is that stratospheric effects cancel out as the FRS and the off-axis spectra are recorded at almost equal SZAs. The latter can be further improved when the FRS at the beginning and at the end of a sequence are interpolated to the measurement time (and thus SZA) of an off-axis spectrum in between. The latter approach is applied in this work (compare Sect. 9.1).

The spectra obtained using the outlined measurement scheme are then analysed using the DOAS principle adapted to passive measurements (compare Sect. 3.1.4). Hence, the basic quantity retrieved from the MAX-DOAS measurements in this work are sequences of dSCDs as function of elevation, i.e. $dSCD(\alpha)$, which describe the differential integrated trace gas concentration with respect to the respective FRS along the effective light path at a given elevation. Since in the atmosphere the concentration of a specific trace gas varies with altitude, this information is also reflected by the relation of the dSCDs measured at different elevations amongst each other. Thus, combining the trace gas information of different elevation angles, allows to retrieve information on the vertical distributions (profiles) of tropospheric absorbers up to ca. 4 km including their total abundances (Hönninger and Platt, 2002; Hönninger et al., 2004). However, for this also information on the effective light paths of the measurements under different elevation angles is needed which will be elaborated in the next section.

3.2.3 *Radiative transfer and air mass factors*

As described in Sect. 3.1 (and corresponding subsections), the basic quantity obtained from scattered sunlight DOAS applications are SCDs (as shown below, the following considerations are also valid for dSCDs introduced in Sect. 3.1.4

and 3.2.2) describing the integrated trace gas concentration along the effective light path (compare Eq. 3.10). Contrary to active DOAS applications for which the light path is well defined, photons can reach the detector on a variety of different light paths and the trace gas absorptions and thus the SCDs strongly depend on the specific light paths the photons travel through the atmosphere besides the actual trace gas concentration. Note that this implies that for a single SCD the individual photons registered by the detector have travelled different paths through the atmosphere before being scattered towards the MAX-DOAS instrument (Hönninger et al., 2004). As discussed in Sect. 3.2.1, the light path strongly depends, amongst others, on the viewing geometry of a measurement. Hence, a comparison of SCDs measured under different viewing geometries is not straightforward and a quantity is desired that is independent of the atmospheric light path and hence the viewing geometry. This can be achieved by the vertical column density (VCD) which is defined as the vertically integrated concentration c_i of a given trace gas i from the ground to the top of the atmosphere (TOA):

$$VCD_i = \int_0^{TOA} c_i(z)\mathrm{d}z. \tag{3.26}$$

Obviously, this definition is independent of the light path, the wavelength or any variable other than the vertical trace gas concentration profile $c_i(z)$. However, for scattered sunlight DOAS applications this quantity is not directly measured but has to be inferred from the corresponding SCD via the so-called air mass factor (AMF) (Noxon et al., 1979; Solomon et al., 1987b) which is defined as the ratio of SCD and VCD:

$$AMF = \frac{SCD}{VCD}. \tag{3.27}$$

The AMF is a measure for the measurement sensitivity as it contains all information on the optical properties of the atmosphere and thus on the radiative transfer along the effective light path, whereby a high sensitivity is associated with a large AMF. Therefore, the AMF generally depends on the viewing geometry (α, θ and ϕ) and the optical properties of the probed air mass. The latter includes for instance the ground albedo, aerosol and trace gas profiles or cloud properties. Another important aspect is that the light path and thus the AMF depends on the probability for atmospheric scattering processes which itself, besides the aerosol abundance, depends on the wavelength of the incident photons leading to a (broad-band) wavelength dependence. Consequently, also the AMF shows a broad-band wavelength dependence which can even become important within a small fitting interval[3]. For instance, the average light path length decreases with decreasing wavelength as the Rayleigh scattering probability increases strongly towards shorter wavelengths (compare Sect. 3.1.2) leading to smaller AMFs in the UV. The wavelength dependence of AMFs and its influence on the comparison results between retrieved and simulated SCDs are discussed in more detail in Sect. 6.4.4.

3 Note that this implies that also the SCDs are wavelength dependent even within a given fitting interval which can have an effect on the DOAS retrieval (compare Sect. 3.1.5.2).

For very simple considerations neglecting multiple scattering, at elevations above ca. 10° and comparably small SZAs, the AMF can be geometrically approximated by the so-called geometric AMF (AMF_{geo}) (Hönninger and Platt, 2002; Hönninger et al., 2004; Wagner et al., 2010). As can be easily seen in Fig. 3.3, for a tropospheric absorber the ratio between vertical and slant light path is $\sin(\alpha)$ and thus AMF_{geo} is defined as

$$AMF_{\text{geo}} = \frac{1}{\sin(\alpha)} \tag{3.28}$$

which can be used for simple analyses of ground-based DOAS applications in particular car-MAX-DOAS measurements (e.g. Ibrahim et al., 2010; Wagner et al., 2010; Shaiganfar et al., 2017).

For more accurate analyses, especially if multiple scattering cannot be ignored, AMFs have to be calculated using radiative transfer models (RTMs) such as McArtim (Deutschmann et al., 2011) or SCIATRAN (Rozanov et al., 2014) requiring information on the atmospheric profiles of multiple parameters (e.g. temperature, pressure, aerosol and trace gases) as well as information on clouds, aerosol optical depth and ground albedo. With these information, radiances can be modelled with and without absorber included in the RTM simulations for a given viewing geometry and consequently the SCD can be computed. Since the VCD is also an input to the model, finally the AMF can be determined using Eq. 3.27.

So far, SCDs were considered, i.e. integrated trace gas concentrations along the effective light path independent of how they are actually obtained. As explained above, in this work the DOAS analysis is performed using a sequential FRS which is to some extent also affected by trace gas absorptions and thus contains a certain SCD (SCD_{FRS}). Consequently, the resulting quantity is a dSCD at the off-axis elevation α ($dSCD(\alpha)$; compare Sect. 3.1.4 and 3.2.2). Naturally, the VCD is the same for an off-axis spectrum and FRS assuming constant atmospheric conditions during one elevation sequence. Rearranging Eq. 3.27 for SCD and inserting in the relation $dSCD(\alpha) = SCD_\alpha - SCD_{\text{FRS}}$ yields

$$dSCD(\alpha) = VCD \cdot (AMF_\alpha - AMF_{\text{FRS}}) = VCD \cdot dAMF(\alpha). \tag{3.29}$$

Here, AMF_α and AMF_{FRS} describe the air mass factors of off-axis and reference measurement, respectively, while the differential air mass factor (dAMF) was introduced as

$$dAMF(\alpha) = AMF_\alpha - AMF_{\text{FRS}} = \frac{dSCD(\alpha)}{VCD}. \tag{3.30}$$

The latter describes the difference in the measurement sensitivities of off-axis and reference measurement and allows to convert the retrieved dSCDs to VCDs. Since in this work the FRS are measured at 90°, the AMFs of the FRS are referred to as $AMF_{90°}$ in the following.

All considerations in this section deal with one pair of SCD (or dSCD) and AMF (or dAMF) yielding a VCD. The output of MAX-DOAS measurements, however, are sequences of dSCDs. Naturally, all dSCDs should describe the same VCD assuming that no temporal changes of the VCD occur within a given sequence.

As in the real atmosphere, not only the sensitivity but also the concentration of a given trace gas varies with height, the combination of the dSCDs obtained at different elevations allows to retrieve the vertical distribution of that trace gas. This, however, is rather complex and requires, besides RTM simulations yielding AMFs, a well-suited profile algorithm (compare Sect. 3.3). Lastly, a short excursion to the height dependence of the measurement sensitivity is provided.

The vertical dependence of the measurement sensitivity is described by a height-resolved AMF, namely the so-called box-air mass factor (boxAMF). The boxAMF describes the ratio of the partial SCD (SCD_j) and the partial VCD (VCD_j) inside an atmospheric layer j (Wagner et al., 2007):

$$boxAMF_j = \frac{SCD_j}{VCD_j}. \tag{3.31}$$

Similar to AMFs, boxAMFs are calculated using RTM simulations by averaging simulated light path lengths through a layer weighted by their modelled radiances. Actually, the total AMF can be calculated from the boxAMFs via

$$AMF = \frac{\sum_{j}^{TOA} boxAMF_j \cdot VCD_j}{\sum_{j}^{TOA} VCD_j} = \sum_{j}^{TOA} boxAMF_j \cdot \frac{VCD_j}{VCD}. \tag{3.32}$$

From this equation, an advantage of simulating boxAMFs becomes evident as the total AMF for any species with a given vertical profile can be calculated from simulated (universal) boxAMFs by the sum over all $boxAMF_j$ weighted by the relative trace gas concentration profile (Wagner et al., 2007).

Since these boxAMFs describe the measurement sensitivities within different atmospheric layers, they allow amongst other applications to infer information on the altitude at which the atmospheric scattering events take place. This information is exploited in Sect. 6.2.3.1.

3.2.4 *O_4 absorptions*

The oxygen dimer O_4, or more precisely $(O_2)_2$ to stress that it is not a real bound molecule, is a collision induced oxygen dimer exhibiting several prominent absorption features in the UV/visible spectral range (Greenblatt et al., 1990). Since the atmospheric O_4 concentration is proportional to the square of the O_2 concentration (with an unknown equilibrium constant; Greenblatt et al., 1990), the O_4 concentration profile is well known (expressed as the square of the O_2 concentration) and mainly depends on the profiles of temperature and pressure. Therefore, the measured O_4 dSCDs mainly depend on the light path length and the altitude distribution of the atmospheric light paths. Thus, measurements of O_4 dSCDs allow to infer information on changes of the radiative transfer in the atmosphere, especially of the light path length, or to characterise the effects of multiple scattering, e.g. in clouds (Wagner et al., 2002b; Hönninger et al., 2004). As the atmospheric aerosol content strongly affects the atmospheric light path of a photon caused by scattering processes (compare above), measurements of

O_4 dSCDs at different elevation angles can be further used to retrieve aerosol extinction profiles (Hönninger et al., 2004; Wagner et al., 2004). The latter is also used by the profile algorithm applied in this work which is described in the next section. Another application of the properties of the atmospheric O_4 absorptions is the classification of the sky conditions, also known as cloud classification (Wagner et al., 2014, 2016; Wang et al., 2017b; Kumar et al., 2020).

3.3 THE MAINZ PROFILE ALGORITHM (MAPA)

The MAinz Profile Algorithm (MAPA) derives vertical profiles of aerosol extinction and trace gas concentrations from MAX-DOAS measurements of trace gas dSCDs. This section briefly introduces the concept of profile inversions and provides a description of the main concepts of MAPA, mainly based on the publication of Beirle et al. (2019)[4] if not stated otherwise.

The considerations in Sect. 3.2.3 showed that for a given atmospheric setting, i.e. amongst others viewing geometry as well as trace gas and aerosol profiles, RTMs can model trace gas SCDs (and consequently dSCDs) for a sequence of elevation angles in a forward model approach. On the other hand, MAX-DOAS measurements yield elevation sequences of trace gas dSCDs containing information on the trace gas concentration and aerosol extinction profiles (compare Sect. 3.2.2). To extract this profile information, the described forward model has to be inverted, i.e. the task is to find the trace gas concentration (or aerosol extinction) profile for which forward modelled and measured sequences of dSCDs agree best. This procedure is known as profile inversion and two basic approaches can be discriminated.

The first approach is known as optimal estimation and is based on a regularised matrix inversion (Rodgers, 2000) yielding concentration and extinction profiles as well as the corresponding averaging kernels describing the information content of a measurement. Moreover, this approach also provides information on the errors of the resulting profiles. Although this is a widely used method (e.g. Frieß et al., 2006; Clémer et al., 2010), there are some drawbacks. For instance, the profile results depend on the a-priori assumptions on the profile shape and its uncertainties.

Contrary to that, parametrised profile inversions try to parametrise the atmospheric profiles with a few parameters (e.g. Li et al., 2010; Wagner et al., 2011). The inversion is then based on the determination of the best matching parameter combination, whereby the main advantage of this approach is that no a-priori assumptions on the profile shape have to be made. However, only profile shapes represented by the specific parametrisation can be reproduced. In particular, the parametrisation used in MAPA cannot reproduce purely exponential profiles (compare below). Further, parameter-based profile retrievals do

4 Note that this publication describes MAPA v0.98 and not v1.01 which is applied in this work. The MAPA algorithm itself, however, is unchanged. Thus, the results are overall independent of the used MAPA version. The major changes are related to the internal integration of the resulting profiles (improvement of runtime) and bug fixes, e.g. affecting the flagging.

not provide averaging kernels nor a clear measure for the uncertainties of the retrieved profiles.

MAPA uses this second approach and is based on a profile parametrisation combining different box profiles which can be simple box profiles starting at the ground, elevated boxes or a combination of a box profile at the ground and an exponential decay with increasing altitude. In particular, the vertical profiles in MAPA are parametrised by three parameters:

- the total column, i.e. AOD for aerosol and VCD for trace gases,
- the layer height h and
- the shape parameter s.

A shape parameter of 1 describes a box profile where the entire column is concentrated below h (Fig. 3.4b). For shape parameters between 0 and 1, the fraction s of the total column is concentrated within a box which vertically extents to h while the remaining fraction $(1 - s)$ declines exponentially with altitude (Fig. 3.4a). Lastly, a shape parameter above 1 but below 2 represents an elevated box profile (Fig. 3.4c). It should be stressed again that this parametrisation is not capable of reproducing purely exponential profiles since the parametrisation is not defined for a layer height of 0 which will become important in the analysis part of this work.

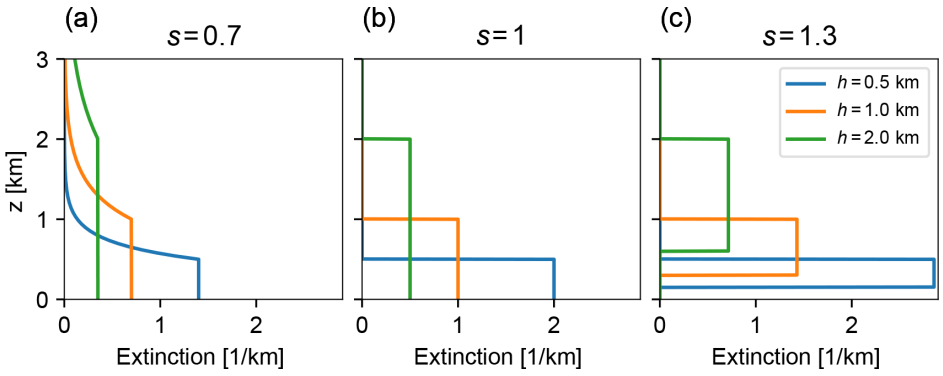


Figure 3.4: Visualisation of the profile parametrisation used in MAPA. Shown are relative profiles, i.e. the total column equals one. Different layer heights are colour-coded, while the different panels show profiles for shape parameters of 0.7, 1.0 and 1.3, respectively (taken from Beirle et al., 2019, used under CC BY 4.0).

To run the MAPA retrieval, the MAX-DOAS data have to be preprocessed to link all required information for the profile inversion to the sequences of dSCDs. In particular this necessary input includes

- the viewing geometry of each measurement consisting of the elevation angle α, SZA θ and RAA ϕ,
- the elevation sequences of the trace gas dSCDs and their corresponding DOAS fit errors as well as

- the a-priori O_4 VCD. This quantity is needed for each sequence/profile to be inverted since first an aerosol profile (based on O_4 dSCDs) has to be retrieved that is then used as input for the trace gas profile inversion (compare below). Here, several options exist. In this work, full profiles of pressure, temperature and humidity (water vapour concentration) obtained from ERA5 (compare Sect. 5.1) are provided and MAPA subsequently calculates the a-priori O_4 VCD itself as described in Beirle et al. (2022).

Instead of determining the best matching set of profile parameters, i.e. minimising the difference between MAX-DOAS measurement and forward model, using a non-linear least-squares Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) like done in previous parameter-based inversion algorithms (e.g. Li et al., 2010; Wagner et al., 2011), MAPA follows a different approach using the Monte Carlo (MC) method. This approach is faster and yields only physically meaningful parameters and uncertainty estimates, e.g. no height parameters below 0 (for details refer to Beirle et al., 2019).

Using this MC approach, the inversion steps are performed as follows. A predefined number of random parameter sets within the physically limited finite parameter space[5] are drawn. In the default configuration, 50 random numbers per parameter are drawn leading to 125 000 and 2500 parameter sets for the aerosol and trace gas retrieval, respectively. Next, using the forward model[6], i.e. basically Eq. 3.29, dSCDs within the given elevation sequence are calculated for each set of random parameters and compared to the measured dSCDs by calculating the root mean square (RMS) within this sequence. The dAMFs needed for this step are linearly interpolated from look-up tables (LUTs) which have been calculated offline using the RTM McArtim (Deutschmann et al., 2011) for a large set of parameters (i.e. profile shapes), viewing geometries (α, θ, ϕ), AODs and wavelengths (details can be found in Beirle et al., 2019). The lowest RMS (RMS_{min}) is identified as best match and an ensemble of up to a predefined number of parameter sets (100 in the default configuration) which perform similarly well within a predefined RMS range is kept ($RMS/RMS_{min} < 1.3$ in the default configuration). These steps are iterated three times to narrow down the parameter space for the next iteration, i.e. the parameter space is getting finer with each iteration. Hence, MAPA finally yields a best matching parameter set (i.e. smallest difference to the measurements) from which the vertical profile can be calculated using the profile parametrisation plus an ensemble of up to 100 (in the default configuration) parameter sets which perform slightly worse but within the predefined limits similarly well.

This ensemble allows to estimate the uncertainties of the best matching parameter set by calculating certain percentiles or standard deviations. Per default, MAPA provides the possibility to calculate the standard deviation, the 25th and 75th percentiles as well as the absolute minimum and maximum. For the applica-

5 This is easily possible since several unphysical parameter combinations can be excluded beforehand. Further, very thin layers which are not resolved by the RTM simulations are excluded.

6 Note that in practice the forward models for aerosol and trace gas retrieval differ because of the meaning of the total column, i.e. AOD and VCD for aerosol and trace gas retrieval, respectively. For further details, refer to Beirle et al. (2019)

tion of the MAPA results to study the atmospheric composition, naturally one result is required. Here, MAPA provides the mean, the weighted mean and the best match. Comparisons in this thesis showed that this choice is not critical and thus the mean of the MC ensemble[7] is used here (compare Sect. 9). Another aspect in this regard is that the usage of means leads to smeared out profiles, i.e. strong vertical gradients are smoothed, and no simple box profiles are obtained by using means, whereby the degree of smearing depends on the specific composition of the MC ensemble. This effect is particularly prominent for cases where the ensemble includes parameter sets with shape parameters larger than 1 (elevated boxes as shown in Fig. 3.4c) leading to steep concentration gradients towards the ground for the mean profiles which in turn lead to artefacts in the profiles. This effect will become important in the analysis part of this work (for instance in Sect. 9.2.2 and 9.2.3).

As outlined in Sect. 3.2.4, the aerosol profile inversion is based on O_4 dSCDs which provide a proxy for changes in the atmospheric light path caused by aerosol scattering. However, several publications on MAX-DOAS measurements and profile retrievals report on the need to scale the measured O_4 dSCDs to achieve good agreement between measurements and the forward model simulations (for details refer to Wagner et al., 2019 and Beirle et al., 2019 and references therein). Hence, MAPA offers the option to perform a scaling of the O_4 dSCDs, whereby in MAPA not the measured but the modelled dSCDs are scaled by a division. Here, two options exist. The first is to prescribe a fixed O_4 scaling factor (SF) for which typically values between 0.8 and 1 are chosen. Given the profile inversion scheme used in MAPA, the second option is to obtain a best matching variable SF from the profile retrieval itself. Details of the latter are outlined in Beirle et al. (2019). Note that inversions yielding unreasonable SFs are flagged later on. In this work, MAPA retrievals are performed with a fixed O_4 SF of 0.8 and a variable one. A discussion on why this choice is justified is provided in Sect. 9.2.1 and 9.3.

For each sequence of input dSCDs, MAPA will yield a best match in terms of an RMS_{min} regardless of the absolute value of this RMS, i.e. agreement between measurements and forward model. In order to provide information on the quality of the inversion result of a given sequence, MAPA provides a detailed flagging of questionable or poor results. Here, three categories exist based on the performance of the profile retrieval. Valid retrievals are sequences for which no flags are raised at all. For warnings, MAPA follows a conservative approach, i.e. low thresholds to raise a warning. Errors, on the other hand, are only raised for inversions which yield completely unreasonable profiles, i.e. comparably high thresholds. Therefore, profiles flagged with an error should be entirely excluded from atmospheric studies, while profiles flagged with a warning might still contain useful information, especially for the total column which is a rather robust measure. The total flag is a warning or an error if any of the individual flags yields a warning or an error, respectively. Without going into too much detail, the flags provided by MAPA can be categorised into four groups, namely

- flags based on the agreement between modelled and measured dSCDs,

7 Note that not the parameters but the resulting profiles are used to calculate the mean due to non-linearities.

- flags based on the consistency of the MC ensemble of parameters,
- flags based on the profile shape itself and
- flags related to external circumstances.

The last group for instance flags results with too high AODs for which the radiative transfer becomes rather insensitive to changes of the AOD. Further, inversions yielding a high AOD for geometries with low RAA, i.e. the instrument points close to the sun, are flagged. This is justified because for such scenarios the forward peak of aerosol scattering becomes important which is not well captured by the Henyey-Greenstein parametrisation used to parametrise the scattering phase-function in the RTM. These two flags will become important in the analysis part of this work since the AOD flag leads to a limited range of AODs that can be compared to AERONET results (compare Chapt. 13). Also, artefacts in the diurnal cycles of the AOD are found for measurements at low RAAs (compare Chapt. 12 and 13). Further, this group contains a dedicated flag for retrievals with a variable O_4 SF to avoid unreasonable SFs since the occasional need of an SF is still not understood (Wagner et al., 2019, 2021). Details on the different flag categories and the corresponding default thresholds leading to warnings and errors can be found in Beirle et al. (2019). In this work, these default thresholds are used. A last aspect here is that MAPA does not provide an explicit cloud flagging but allows to provide external cloud flags, e.g. obtained from cloud classification schemes (Wagner et al., 2014, 2016). Nevertheless, several applications of MAPA in this work show that the implemented MAPA flags are particularly sensitive to cloudy scenes which are associated with complex radiative transfer and thus implicitly provide a rather strict cloud flagging (compare for instance Sect. 9.3 or 10.6).

Finally, further clarifications which are relevant for this work are provided. In particular, MAPA yields vertical profiles of trace gas concentrations and aerosol extinction above instrument altitude which is not necessarily the height above ground. While this difference is negligible for most measurement sites, it plays an important role for ATTO where the measurements are performed on a tower (compare Sect. 4.4 and Chapt. 13). Further, the MAPA output is sampled on the default output grid having a resolution of 20 m up to a height of 200 m and 200 m above this altitude up to a height of 4 km. Above that height, a few more layers exist with an uneven spacing. Besides vertical profiles, the MAPA output provides integrated columns, i.e. trace gas VCDs and AODs. While only physically reasonable values are possible for the AODs, the trace gas VCDs can become negative caused by the different approach to calculate them (compare above). However, these negative VCDs must not be discarded or set to 0 but can be understood to represent noise for almost zero VCDs. It should be stressed that for each sequence to be inverted, first an aerosol retrieval has to be performed to constrain the radiative transfer, i.e. to extract the trace gas dAMFs from the LUTs which depend on the aerosol profile. To perform the consecutive trace gas retrieval, the best matching parameter combination of the aerosol retrieval is used since due to non-linearities, i.e. the mean of the ensemble profiles does not equal the profile corresponding to the mean parameters, it is not possible to use mean parameters for this. If not stated otherwise, default settings as described in Beirle

et al. (2019) were used for the MAPA algorithm itself. However, LUTs interpolated to the effective wavelengths of the respective fitting intervals of the DOAS fit were used instead of LUTs at wavelengths closest to the centre wavelength like in the default MAPA setup (compare Sect. 9.2.1). Finally, MAPA offers the possibility to apply an elevation angle offset to the reported elevation angles to correct for inaccuracies in the elevation calibration. Also, this option was used in this work (compare Sect. 9.2.2).

4

INSTRUMENTS AND MEASUREMENT SITES

Data from five self-built MAX-DOAS instruments which were operated at four different measurement sites were used for the studies presented in this work. In addition to these stationary data, spectra acquired during a research ship cruise were used in the sensitivity study for the fit settings of formaldehyde. A map showing the locations of the four stationary measurement sites and the track of the research ship cruise is provided in Fig. 4.1. This chapter introduces the measurement sites including their locations and specific environmental conditions (Sect. 4.1 to 4.5) and provides descriptions of the instruments used (Sect. 4.1.1 and 4.2.1).

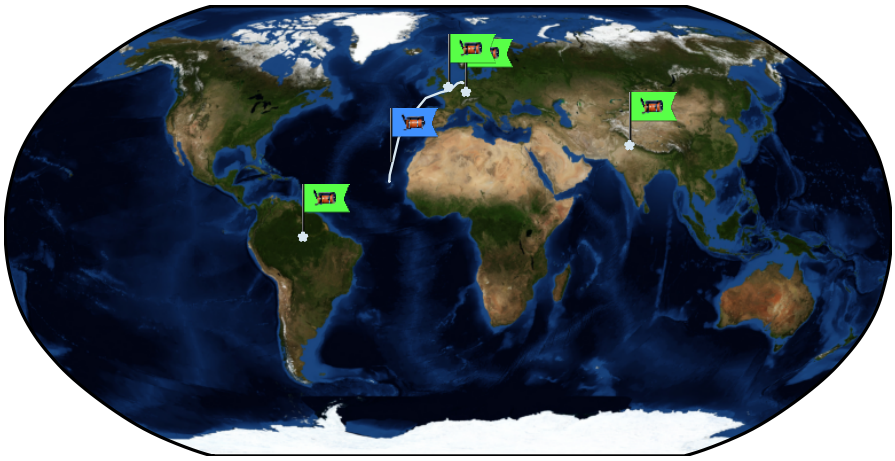


Figure 4.1: Overview of the locations of the four stationary measurement sites (green flags) and the track of the research ship cruise (blue flag). The map was created using Basemap (MIT License) with NASA Blue Marble data (http://visibleearth.nasa.gov).

4.1 MAINZ, GERMANY

The city of Mainz, itself having a population of around 225 000 inhabitants (Landeshauptstadt Mainz, 2023), is located on the Rhine river in the south-western part of Germany at the western edge of the densely populated Rhine-Main region. Besides several medium-sized towns, situated mainly to the north-east, east and south-east of Mainz, the city is also surrounded by three larger cities and one airport. Wiesbaden, having a population of ca. 295 000 inhabitants (Stadt Wiesbaden, 2023), Frankfurt with ca. 750 000 inhabitants (Stadt Frankfurt am Main, 2023) and Darmstadt, having a population of approximately 160 000 inhabitants (Wissenschaftsstadt Darmstadt, 2021), are located around 9 km to the north, 35 km to the north-east and 33 km to the south-east, respectively. Frankfurt Airport

which is the largest airport of Germany is located ca. 25 km east of Mainz and had about half a million commercial flight movements in the last pre-COVID year 2019 (Arbeitsgemeinschaft Deutscher Verkehrsflughäfen, 2023). A map of Mainz and its major surrounding cities in the south-western part of Germany is provided in Fig. 4.2.

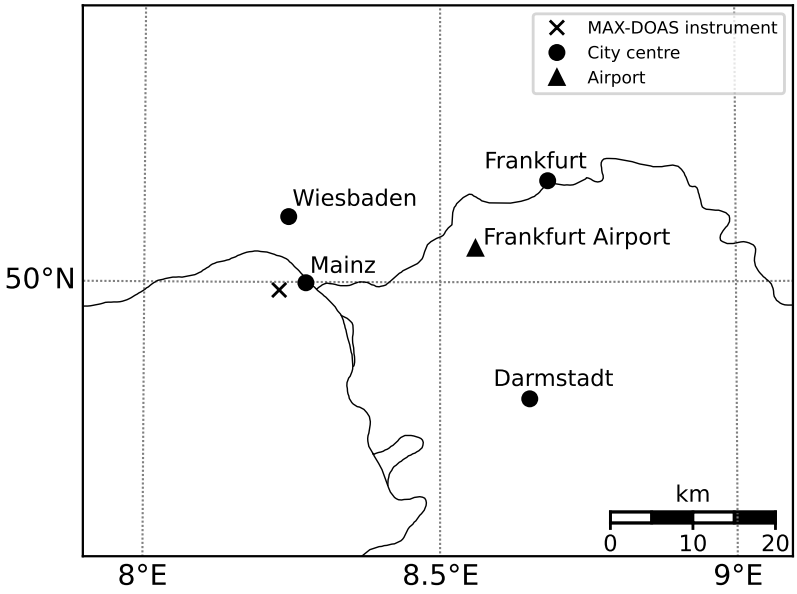


Figure 4.2: Overview of the locations of Mainz and its surroundings in the south-western part of Germany. The location of the instrument is marked by a cross, while city centres are indicated by circles. Further, the location of Frankfurt Airport is represented by a triangle. The map was created using Basemap (MIT License) with GMT data.

Associated with the rather high population density in the Rhine-Main region, heavy traffic occurs not only during rush hours. Further, several industrial plants are located in and to the north and east of Mainz. In the direct western and southern surroundings, mainly agricultural land is situated, while the Taunus, a low mountain range mainly covered with forests, is located north and north-west of the city. Finally, more far to the south (ca. 55 km) along the Rhine river, Ludwigshafen with heavy chemical industries can be found.

The prevailing climate at Mainz is moderate with four distinct meteorological seasons: winter (December to February), spring (March to May), summer (June to August) and autumn (September to November). Figure 4.3 provides an overview of temperature (a) and wind (b) statistics for the measurement period presented in this work (June 2013 to March 2022).

A pronounced annual cycle of the measured air temperatures can be observed with maximum and minimum monthly mean temperatures of around 21 °C and 3 °C in July and January, respectively. The observed span of temperatures is typical for a moderate mid-latitude semi-continental climate, in which, besides overall moderate temperatures, also rather extreme warm and cold temperatures can occur. The relative frequency distribution of the measured wind directions

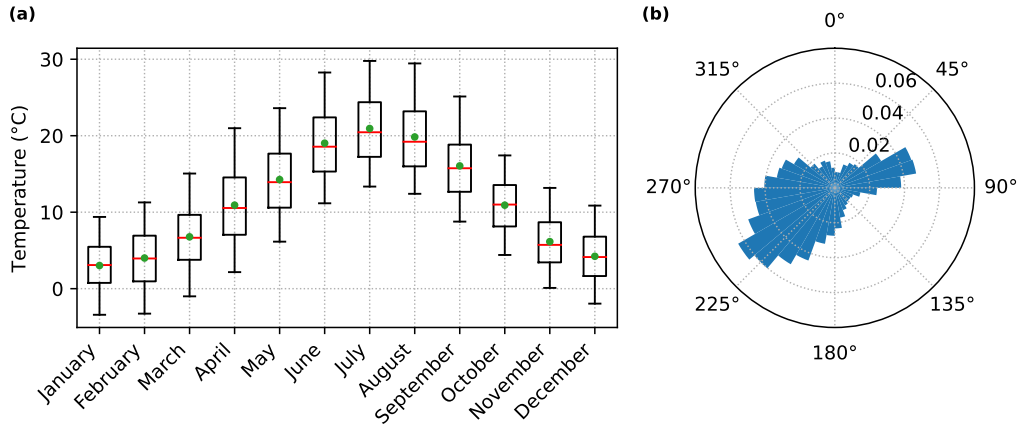


Figure 4.3: Weather statistics for the period from June 2013 to March 2022 obtained by the weather station of the Institute for Atmospheric Physics of the University of Mainz. Panel (a) shows box-whisker-plots for the annual cycle of measured air temperatures. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. The 5th and 95th percentiles are shown by the whiskers of the boxes. Panel (b) shows the relative frequency distribution of the measured wind directions in bins of 10°.

in Mainz is dominated by two maxima. The stronger one indicates frequent south-westerly and westerly winds, caused by low pressure systems originating from the North Atlantic and their corresponding frontal zones. The second weaker pronounced maximum is associated with easterly winds, which occur when, especially in winter and summer, high pressure systems dominate central Europe. It should be mentioned that the mentioned frontal passages are typically accompanied by clouds and rain. Such conditions might lead to a rather low rate of valid MAPA retrievals which are rather sensitive to such complex sky conditions (compare Sect. 3.3). However, given the high frequency of such conditions in Mainz compared to other wind directions, many measurements under these meteorological conditions exist.

Given the observed wind distribution, the measurements at Mainz can be influenced by anthropogenic pollution originating from the cities, the traffic and industries of the Rhine-Main region. Contrary to that, rather clean air masses can be predominant under westerly winds which makes Mainz an interesting place to study the impact of anthropogenic emissions on the formaldehyde abundances.

The MAX-DOAS measurements at Mainz were conducted at the Max Planck Institute for Chemistry (MPIC; 49.991 °N, 8.228 °E; 125 m a.m.s.l.) located at the western edge of the city. A self-built MAX-DOAS instrument, the so-called 4-azimuth MAX-DOAS instrument, which can perform measurements in four different azimuth directions at the same time, was installed on the roof of the institute building ca. 25 m above ground level in June 2013. Only one of the four possible azimuth directions was used for the analyses presented in this work, namely the one looking towards the city at an azimuth angle of 51° from north. Further details of the instrumental setup and its characteristics are discussed in Sect. 4.1.1. Data are available since June 2013, however, several data gaps and

changes in the optical properties of the instrument occurred within the almost nine years of measurements. Details of data availability and quality are discussed in Chapt. 11.

Useful ancillary data are measured in the vicinity of the MAX-DOAS instrument. An AERONET station (details in Sect. 5.2) measuring optical properties of aerosol is also located on the roof of the MPIC, while the weather station of the Institute for Atmospheric Physics of the University of Mainz is located in around 500 m distance from the MAX-DOAS instrument.

4.1.1 *The 4-azimuth MAX-DOAS instrument*

At Mainz data were collected with the so-called 4-azimuth MAX-DOAS instrument which is a self-built MAX-DOAS instrument developed and built in collaboration with the electronics workshop of the Max Planck Institute for Chemistry (MPIC) in Mainz, Germany. Its setup and main characteristics are outlined in this section. For details beyond the overview provided in this work, refer to Krautwurst (2010) and Remmers (2011). Studies using data collected with this instrument are for example summarised in Wang et al. (2017a) and Wagner et al. (2019).

4.1.1.1 *Instrumental setup*

The 4-azimuth MAX-DOAS instrument is composed of two major parts, namely the telescope and the spectrometer units.

The telescope unit is located on the roof of the MPIC building and primarily consists of a metal cross. The telescopes (FOV of ca. 0.6°) with the corresponding stepper motors are attached to its four ends in a way that their viewing directions are perpendicular to each other (compare Fig. 4.4). In the current configuration, the telescopes have viewing directions of 321°, 51°, 141° and 231° from north. The scattered sunlight collected by the four telescopes is transmitted to the spectrometer unit via quartz fibre bundles each consisting of six individual fibres which finally unite to one fibre bundle. Thus, the simultaneously collected light is coupled into the same spectrograph. The described setup allows the collection of spectra in four different azimuth directions at the same time using the same spectrometer which in principle allows to study spatial gradients of trace gases and aerosol with only one instrument. However, as mentioned before, in this work data from only one viewing direction, namely the one looking towards the city of Mainz at an azimuth angle of 51°, are used. This restriction was made to minimise the effort and is justified as horizontal gradients are not the focus of this work.

The spectrometer unit, which is located inside a laboratory roughly 5 m below the telescope unit, consists of a metal box containing the spectrograph which is thermally stabilised at a temperature of 35 °C to guarantee stable optical properties. The 2D-CCD-camera which records the collected spectra is attached to the metal box. Further characteristics of spectrograph and camera are outlined in the next section. Besides these components, several controller units and a PC are part of the setup and placed in the laboratory next to the spectrometer unit.



Figure 4.4: Telescope unit of the 4-azimuth MAX-DOAS instrument located on the roof of the MPIC building. Courtesy: Julia Remmers.

The software controlling and synchronising the whole system including camera and motor was developed in the Satellite Remote Sensing Group of the MPIC.

4.1.1.2 *Characteristics*

Core of the whole setup is the spectrograph, where the collected scattered sunlight is diffracted into its spectral components. Here, a Shamrock SR-303i spectrograph built by the Andor company is used. It consists of a deflecting mirror, two concave mirrors and a grating tower and is built as so-called Czerny-Turner system. An illustration of the principal light path in such a system is illustrated for the spectrometers used in the TubeMAX-DOAS instruments in Fig. 4.11. The grating tower can contain up to three different reflection gratings which in principle offers the possibility to probe different spectral regions. In the current configuration, however, only one grating is used, leading to a covered wavelength range of ca. 320 to 460 nm. Given the optical setup which is described below, the spectral resolution is in the order of 0.6 to 0.9 nm as described by the the full width at half maximum (FWHM). In practice the wavelength range and the spectral resolution can vary, depending on the exact orientation of the grating tower and the location of the image of the respective fibre bundle on the CCD (see below). Details of the actual wavelength range and the spectral resolution during the used measurement period are discussed in Chapt. 11.

The diffracted light is focussed onto the 2D-CCD-camera, which is a camera of the type DV420A-BU by the Andor company. The two-dimensional back-illuminated CCD array consists of 1024 columns and 255 rows and records for each measurement the images of the four fibre bundles of the different telescopes as depicted in Fig. 4.5.

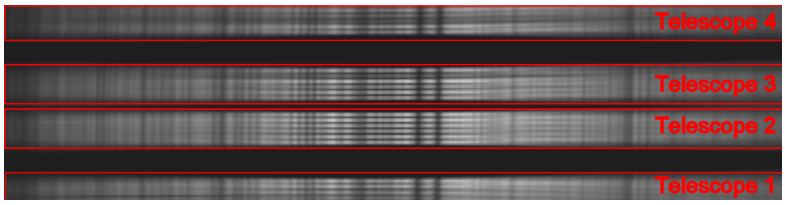


Figure 4.5: Image of the four fibre bundles on the two-dimensional CCD. Courtesy: Julia Remmers.

During the read-out process (details in Krautwurst, 2010), a 16-bit analogue-to-digital (A/D) converter converts the signal induced by the photons into a digital signal which can be further processed by software and results in an one dimensional representation of the measured spectrum for each telescope. The usage of a CCD, an A/D converter and its electronics leads to three major implications that have to be considered before the spectra can be analysed using the DOAS method, namely the offset, the dark current (DC) and the detector linearity. These effects are discussed in more detail for the TubeMAX-DOAS instrument in Sect. 4.2.1.2. Detailed descriptions of the background correction for the 4-azimuth MAX-DOAS instrument can be found in Krautwurst (2010) and Remmers (2011). In order to minimise the influence of thermal fluctuations on the recorded spectra, the CCD-camera is stabilised at a temperature of $-30\,°C$. Further, Krautwurst (2010) showed that the detector exhibits a rather linear response (compare Sect. 4.2.1.2) on intensity changes at typical signal levels. The major advantage of the used camera is its high signal-to-noise ratio which allows comparably short exposure times and thus a high temporal resolution of the measurements.

4.2 BAYFORDBURY, UNITED KINGDOM

The Bayfordbury Observatory (51.775 °N, 0.095 °W; 67 m a.m.s.l.) is operated by the University of Hertfordshire. It is located approximately 30 km north of the centre of London, United Kingdom, and is apart from that only surrounded by some smaller towns and villages. The largest and closest of them are Hatfield (ca. 42 000 inhabitants), Welwyn Garden City (ca. 52 000 inhabitants) and Hertford (ca. 29 000 inhabitants) located around 10 km to the west, 7 km to the north-west and 3 km to the north-east, respectively (City Population, 2023). Further, two international airports, namely Luton (ca. 20 km to the north-west) and Stansted (25 km to the north-east) are located in its vicinity, while the largest airport of London, Heathrow, is located roughly 42 km to the south-west. In its nearest surroundings mainly agricultural fields and smaller forests can be found. The location of the observatory and its surroundings in the south-eastern part of Great Britain are depicted in Fig. 4.6.

The predominant climate at Bayfordbury is moderate with four standard seasons: winter (December to February), spring (March to May), summer (June to August) and autumn (September to November). Located on the British Isles a strong maritime influence leading to frequent south-westerly to westerly winds

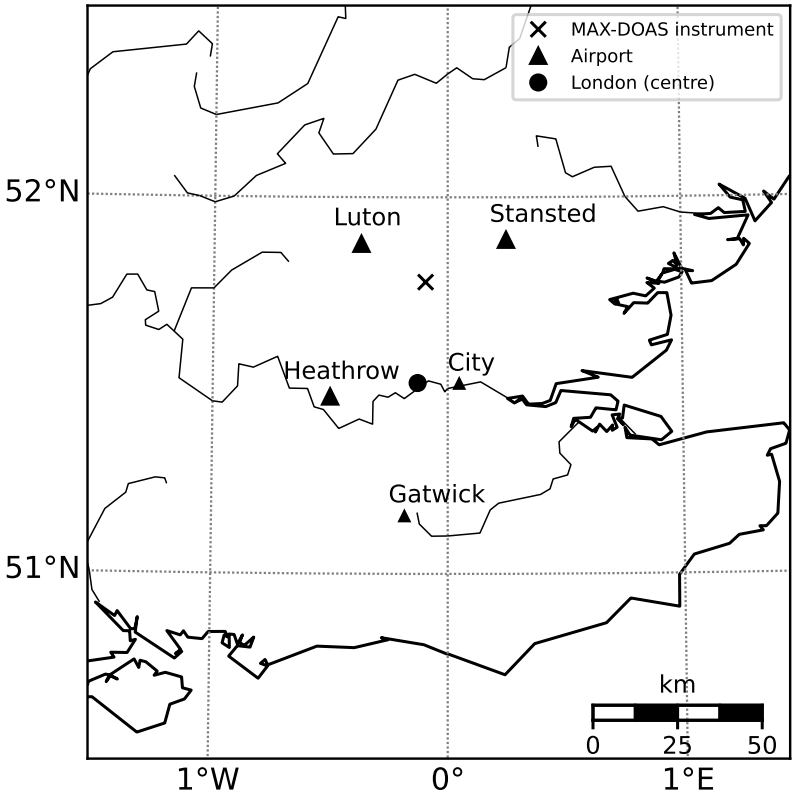


Figure 4.6: Overview of the locations of the Bayfordbury Observatory and its surroundings in the south-eastern part of Great Britain. The location of the observatory is marked by a cross, while the city centre of London is marked by a circle. Large triangles mark international airports whose pollution might influence the measurements at Bayfordbury. Finally, smaller triangles mark airports whose pollution might be masked by the pollution of London. The map was created using Basemap (MIT License) with GMT and GSHHG data.

and relatively balanced temperatures with mild winters and cool summers is expected. Figure 4.7 provides an overview of temperature (a) and wind (b) statistics for the time period of the measurements presented in this work (November 2018 to July 2022). A clear seasonal cycle of the measured air temperatures can be observed with maximum and minimum monthly mean temperatures of around 18 °C and 5 °C in July and January, respectively. The relatively small span of temperatures is typical for a maritime influenced moderate mid-latitude climate. A further feature of such climates on the northern hemisphere are frequent westerly winds which becomes evident also in the wind data for Bayfordbury. Maxima of the relative frequency distribution of the measured wind directions are found under west-north-westerly and west-south-westerly directions. Such westerly winds in north hemispheric mid-latitudes are often associated with frontal passages which typically lead to cloudy and rainy episodes. Further, foggy scenes are rather frequent in those climates, especially during autumn and winter. Like for Mainz (Sect. 4.1), it should be mentioned that such conditions might lead to a rather poor rate of valid profiles in the MAPA retrievals which are rather sensitive to these

sky conditions (compare Sect. 3.3). However, many measurements are performed under such conditions also at Bayfordbury in view of the higher frequency of those conditions compared to other wind directions.

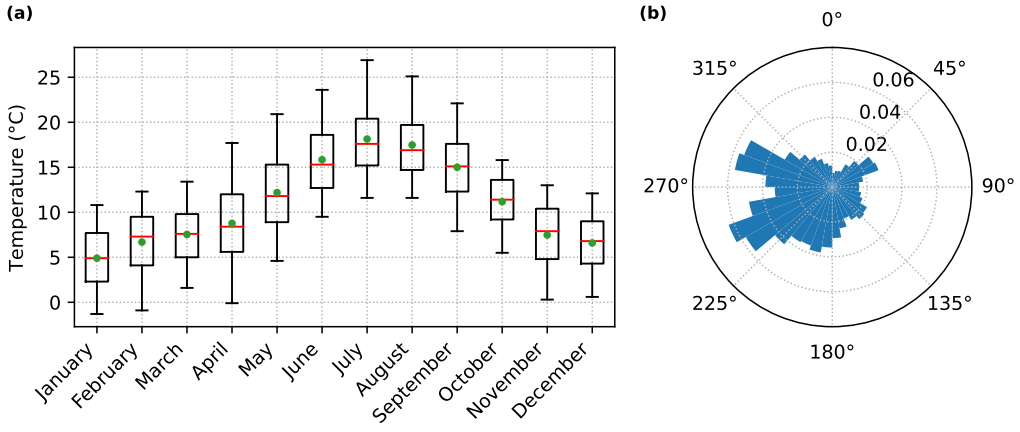


Figure 4.7: Weather statistics for the period from November 2018 to July 2022 obtained by the weather station located at the Bayfordbury Observatory. Panel (a) shows box-whisker-plots for the annual cycle of measured air temperatures. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. The 5th and 95th percentiles are shown by the whiskers of the boxes. Panel (b) shows the relative frequency distribution of the measured wind directions in bins of 10°.

Depending on wind direction, the probed air masses at the site are expected to be influenced by pollution originating from the city of London under southerly winds or relatively clean air masses under westerly, northerly and easterly winds. The influence of the neighbouring towns and airports and emission sources located further away is investigated in detail in the respective analysis section of this work (Chapt. 10). Wind statistics show that in principle all wind directions except straight north winds occur regularly. This makes Bayfordbury an interesting measurement site, where different pollution regimes can be present.

Originally designed as teaching observatory for astronomy and astrophysics, explaining its rather remote location far from heavy light pollution, the observatory is also used as remote sensing site since 2010 and thus hosts several remote sensing instruments including a LIDAR and an AERONET station (University of Hertfordshire, 2023). Data of the latter are used for an assessment of the MAPA retrieval results at various points in this work. Lastly, also all-sky cameras (day and night modes) and a weather station measuring basic atmospheric quantities (compare Fig. 4.7) are deployed on-site.

A RaspberryPi-based version of the TubeMAX-DOAS instrument which is introduced in the following section (Sect. 4.2.1) was installed at Bayfordbury on 21 November 2018. The instrument was placed on a gantry ca. 3 m above ground level with an azimuth viewing direction of 37° from north. A picture of the telescope unit of the Bayfordbury instrument is shown in Fig. 4.8. Data are in principle available without major gaps since then, however, the fan of the cooling unit broke in April 2020 and the spectrometer temperature had to be increased

as consequence. Given the stable properties of the used spectrometer, the data quality was not affected by this failure. Further details of data availability and quality will be discussed in Chapt. 10.

4.2.1 *The TubeMAX-DOAS instrument*

Most of the measurement data used in this work were collected with different versions of the so-called TubeMAX-DOAS instrument, which is a self-built type of MAX-DOAS instrument developed and constructed in collaboration with the electronics workshop of the Max Planck Institute for Chemistry (MPIC) in Mainz, Germany. Its principle setup (Sect. 4.2.1.1) and main characteristics (Sect. 4.2.1.2) are outlined in this section. For further technical details and some successful applications of this instrument refer to Donner (2016), Lukošiūnaitė (2021), Lauster et al. (2021) and Wagner et al. (2021).

4.2.1.1 *Instrumental setup*

The TubeMAX-DOAS instrument consists of two main parts, namely the spectrometer and the telescope units. Besides these two main components, a controller unit and a computer are part of the setup.

The telescope unit houses a stepper motor together with its controller unit inside an orange plastic tube, giving the instrument its name (see picture in Fig. 4.8). In agreement with the manufacturer information, laboratory measurements by Lauster et al. (2019) showed that 16 000 motor steps correspond to an angular movement of 90°, leading to a single motor step size of 0.0056°. Further, these investigations confirmed that the motor accuracy and precision are much smaller than the field of view (FOV) of the instrument's telescope which is approximately 0.7°. The latter is attached to the stepper motor which enables measurements at different elevation angles. Using a quartz lens the telescope focuses the incident scattered sunlight onto a glass fibre bundle at its end which transmits the light to the spectrometer. The bundle consists of four individual glass fibres having a circular layout on the telescope side, while they form a linear array on the spectrometer side, representing its entrance slit. For measurements on mobile platforms, e.g. on ships (compare Sect. 4.5) or cars, the telescope unit can be equipped with a gyroscope to actively stabilise the elevation angles by continuously adjusting the motor position with an accuracy of ca. 0.1°. As shown by Lauster et al. (2019), temperature effects on the properties of the gyroscope can be mostly neglected for measurements on a moving platform.

Besides the spectrometer whose main characteristics are outlined in the next section, the spectrometer unit contains a Peltier element which is mounted on the spectrometer housing and allows to stabilise the spectrometer temperature with only small fluctuations of around 0.1 °C. Thereby, the used hardware allows both, cooling and heating, so that a constant spectrometer temperature can be used throughout the year. In order to control Peltier element and motor and to synchronise them with the measurement routine, a controlling system developed by the electronics workshop of the MPIC is also located inside the spectrometer unit. This system basically sets and reads the parameters of the different parts



Figure 4.8: Telescope unit of the Bayfordbury instrument. The orange tube houses the motor which moves the telescope together with its controlling unit. Courtesy: David Campbell.

and sends them to a computer. Here, two different versions of TubeMAX-DOAS instruments exist.

The first one is called "laptop-based" in the following and uses a quite compact spectrometer unit (see Fig. 4.9) which is connected via USB to a laptop that runs the controlling and measurement software. Besides the self-developed controlling software, this instrument version uses the *DOASIS* software (Kraus, 2006) to run the main measurement script and to record the spectra of scattered sunlight. Further, this spectrometer unit has to be placed inside, as it is not waterproof.

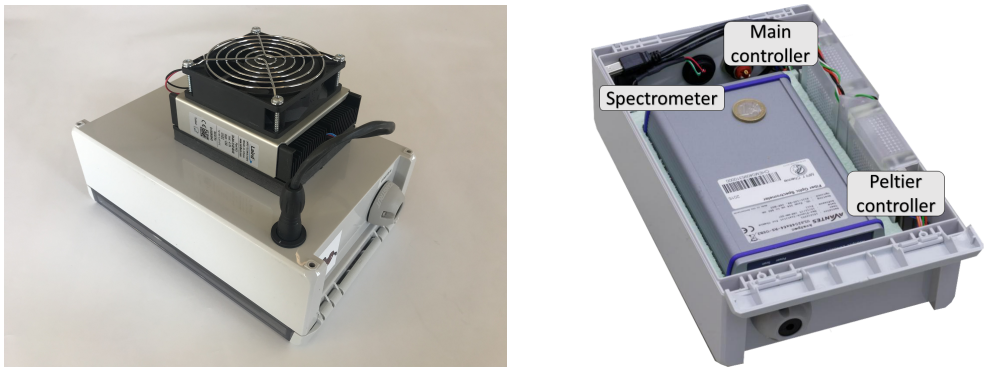


Figure 4.9: Spectrometer unit of the laptop-based TubeMAX-DOAS instrument. On the right, the interior consisting of spectrometer, main and Peltier controller is shown. A 1 €-coin serves as size scale. Courtesy: Steffen Ziegler.

The second version is called "RaspberryPi-based" in the following and utilises a larger and waterproof spectrometer unit (see Fig. 4.10) which additionally contains a RaspberryPi mini-computer that runs all necessary software. Despite having the same measurement characteristics which mainly depend on the used spectrometer, the RaspberryPi-based version has proven to be more stable in terms of instrumental failures. This is because less USB-connections are needed which can be rather unstable and because the spectrometer is directly connected to the RaspberryPi via a serial connector which in turn leads to a slightly slower

read-out of the spectrometer[1]. A further advantage of the RaspberryPi-based version is the fact that after power failures, which regularly occur at Bayfordbury and ATTO, measurements are automatically restarted which is not the case for the laptop-based version. This instrument version only uses software developed in-house at the MPIC which is another reason for the more stable measurements.

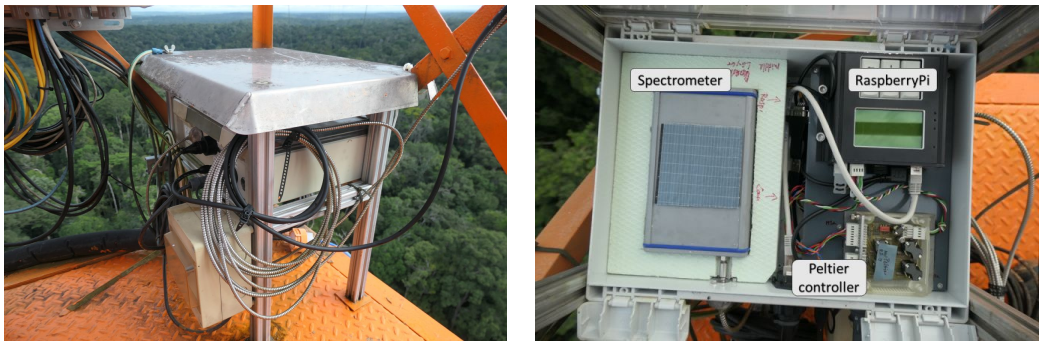


Figure 4.10: Spectrometer unit of the RaspberryPi-based TubeMAX-DOAS instrument installed at an altitude of 80 m at ATTO. On the right, the interior consisting of spectrometer, RaspberryPi and Peltier controller is shown.

An overview of which instrument type is operated at the different stations is provided in Table A.1 in the appendix. This table further lists the employed elevation sequences together with the integration times used to acquire an accumulated spectrum.

4.2.1.2 *Characteristics*

All versions of the TubeMAX-DOAS instrument use the same spectrometer type, namely an AvaSpec-ULS2048x64 spectrometer with built-in CCD-chip manufactured and distributed by the AVANTES company.

The spectrometer is built as symmetrical Czerny-Turner system, where light is coupled in through a slit and collimated by a spherical mirror. Afterwards, the light is split up into its spectral components by a plain grating. Finally, another spherical mirror focuses the diffracted light on the detector array which is a Hamamatsu back-thinned CCD-chip with 2048 columns and 64 rows, where the photons are registered. An illustration of the described light path is presented in Fig. 4.11.

To reduce the stray-light inside the spectrometer, additionally a BG3 filter which removes light of wavelengths longer than ca. 450 nm (SCHOTT AG, 2023) is placed in the light path[2]. For older instruments this filter is situated directly after the slit in the spectrometer housing, while in newer versions the filter is placed directly in front of the fibre bundle inside the instrument's telescope. Given the described optical setup, the used spectrometers cover a wavelength range of around 300 to 465 nm at a spectral resolution of about 0.55 to 0.75 nm as described by the

1 In a further development step this was changed. However, only few months of data collected with this newer instrument version at ATTO (298 m) are included in this work.

2 The filtered wavelength interval is ca. 450 to 750 nm. Also wavelengths above 950 nm and below 300 nm are filtered. The latter, however, is not important for passive DOAS applications.

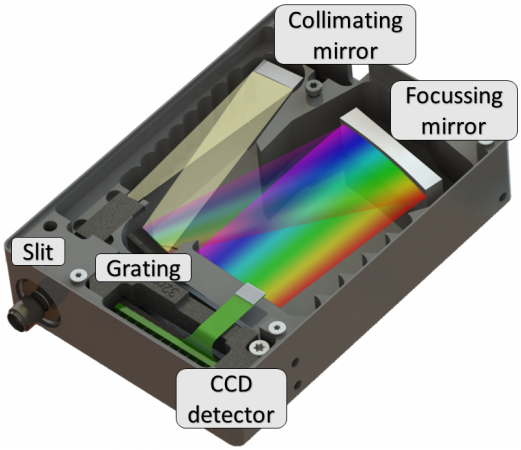


Figure 4.11: Illustration of the light path in a symmetrical Czerny-Turner system as described in the text (adapted with permission from AVANTES BV, 2023).

FWHM of a Super Gaussian parametrisation of the instrumental slit function (SLF; compare Sect. 8.1). The exact values slightly vary between different spectrometers. Further, the used telescope, the adjustment of the fibre bundle in front of the slit and the spectrometer temperature influence the exact values. The spectral resolution additionally slightly varies over the spectral range of the respective spectrometer. Here, the spectral resolutions are provided at 341.5 nm which is the centre wavelength of the applied HCHO fitting interval (compare Table 9.1). Table 4.1 lists the optical characteristics of the spectrometers used at the different measurement stations. For stations where different spectrometers were used or different spectrometer temperatures had to be set, the characteristics of the respective periods are listed.

Table 4.1: Overview of the spectral characteristics of the spectrometers used at the different stations. For stations where different spectrometers were used or different spectrometer temperatures had to be set, the characteristics of the respective periods are listed.

STATION		PARAMETER		
		Wavelength range	Spectral resolution	Spectrometer temperature
Bayfordbury	11/2018 – 04/2020	301 – 464 nm	0.70 nm	15 °C
	04/2020 – 06/2020	301 – 464 nm	0.71 nm	20 – 30 °C
	07/2020 – 07/2022	301 – 464 nm	0.70 nm	28 °C
Mohali	04/2019 – 07/2022	302 – 465 nm	0.71 nm	15 °C
ATTO 80 m	10/2017 – 07/2022	303 – 466 nm	0.67 nm	25 °C
ATTO 298 m	03/2019 – 09/2019	302 – 465 nm	0.62 nm	15 °C
	09/2019 – 04/2020	301 – 465 nm	0.72 nm	15 °C
	04/2022 – 09/2022	306 – 466 nm	0.66 nm	25 °C

The CCD signal is converted into a digital signal by a 16-bit A/D converter. The usage of a CCD, an A/D converter and its electronics leads to three major effects that have to be considered before the spectra can be analysed using the DOAS method:

Firstly, an artificial signal is added to the measured values by the electronics since the A/D converter can only handle positive values. This offset is removed by subtracting an offset spectrum from the spectra. To do so, the offset spectrum (I_{offset}) has to be scaled with the number of scans (N_{meas}) of the measured spectrum (I_{meas}):

$$I_{\text{corr}} = I_{\text{meas}} - \frac{N_{\text{meas}}}{N_{\text{offset}}} \cdot I_{\text{offset}} \tag{4.1}$$

Where I_{corr} denotes the corrected spectrum and N_{offset} is the number of scans of the offset spectrum. In practice, the offset spectra used in this work are obtained by recording spectra during night with an integration time of 3 ms and a scan number of 10 000, to minimise the effect of statistical fluctuations. An offset spectrum typical for the here-used AVANTES spectrometers is shown in Fig. 4.12a.

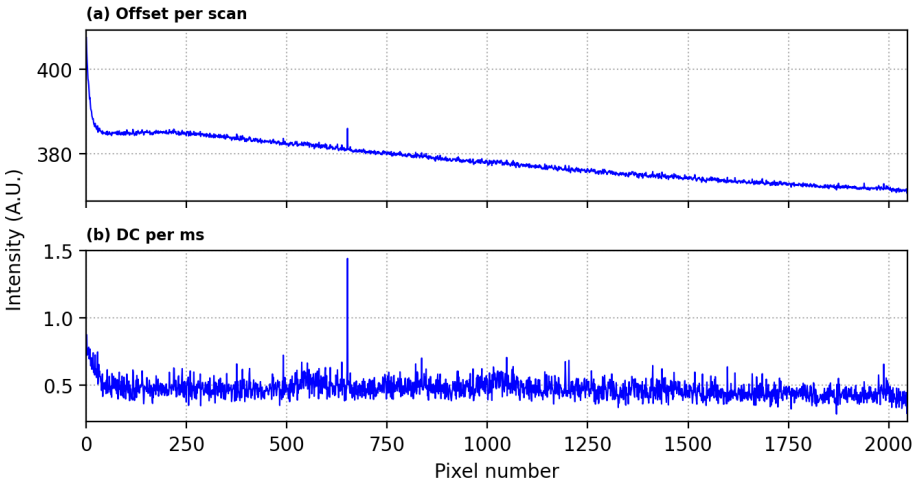


Figure 4.12: Example offset and DC spectra of the AVANTES spectrometers used in this work measured by the Bayfordbury instrument on 20 December 2018. The spectra are normalised to one scan and 1 ms, respectively.

Since CCD-chips are made of semiconductors, also thermal photons induce free electrons and thus cause another artificial signal, namely the dark current (DC). This effect is highly temperature dependent and therefore, spectrometers and detectors are typically temperature stabilised as explained above which is a prerequisite to correct for this effect. The temperature dependence of the DC of the utilised CCDs was investigated in Donner (2014). For the correction of this effect, DC spectra are also collected during night time, however, here a single spectrum is recorded at an integration time of 20 000 ms to increase the amount of registered thermally induced electrons. For the subtraction from the measured

spectra, the dark current spectrum (I_{DC}) has to be scaled by the total integration time of an accumulated measured spectrum (t_{meas}):

$$I_{corr} = I_{meas} - \frac{t_{meas}}{t_{DC}} \cdot I_{DC} \tag{4.2}$$

Here, t_{DC} denotes the integration time of the DC spectrum. Note that also dark current spectra contain the offset and thus, have to be corrected for this effect. An example DC spectrum for the here-used AVANTES spectrometers is depicted in Fig. 4.12b. The most prominent feature is a so-called hot pixel at pixel 651 (corresponding to ca. 358 nm) which are pixels that produce an overproportionate amount of photoelectrons. Such pixels are typically accounted for by introducing (small) spectral gaps in the DOAS retrieval (compare e.g. Sect. 10.1).

Lastly, investigations by Horbanski (2015) showed that the response of the used detector chips is not perfectly linear, most likely caused by the detector electronics. As a consequence, the measured signal does not linearly increase with an increase of integration time or intensity. While this can be expected for nearly every instrument at saturation levels close to 100 %, the used AVANTES spectrometers show this non-linearity also for lower saturation levels. Donner (2016) estimated the effect of non-linearity to be approximately 5 % for the spectrometers used in this study which is quite substantial compared to the small optical depths of the trace gas absorptions (compare Sect. 3.1). Nevertheless, as the non-linearity is specific for each individual spectrometer, different values are expected for different instruments. The observed detector non-linearity can in principle be corrected for as shown in Horbanski (2015). However, the studies in Donner (2016) concluded that most of the non-linearity effect on the spectral DOAS retrieval is accounted for by the intensity offset that is included in the DOAS analysis (compare Sect. 3.1). For that reason, no non-linearity correction is applied during the analyses presented in this work.

4.3 MOHALI, INDIA

Mohali, India, officially named Sahibzada Ajit Singh Nagar, is located in the north-west Indo-Gangetic Plain and part of the so-called *"Tricity"* of Chandigarh which consists of the cities of Chandigarh, Mohali and Panchkula. According to estimates of the 2011 census the three cities individually had 1.2 million, 200 000 and 280 000 inhabitants in 2022, respectively, leading to approximately 1.7 million people living in the entire metropolitan area (Indian Census, 2011). Around 35 km north of Mohali the Himalayan mountain range starts. The measurements at Mohali were performed at the Indian Institute of Science, Education and Research (IISER), situated around 5 km south of the city centre. Thus the *"Tricity"*, including its airport, is located within a sector spanning from north to east. South-east of the measurement site a mix of agricultural lands, smaller towns and cities and industries can be found, while the southern and western sectors are mainly enclosed by agricultural fields and smaller towns. Further, two major power plants are situated within roughly 50 km of the measurement site. The first one is the Guru Gobind Singh Super Thermal Power Plant (PP1) with an installed capacity of 840 MW, located around 45 km to the north-west (342°), however, only

approximately 25 % of its capacity were used in pre-COVID years (Punjab State Power Corporation Ltd., 2023). The second power plant is the Rajpura Thermal Power Plant (PP2) having an installed capacity of 1400 MW, located roughly 18 km south-west of the measurement site (230°). Figure 4.13 summarises the location of the measurement site and its most important surroundings.

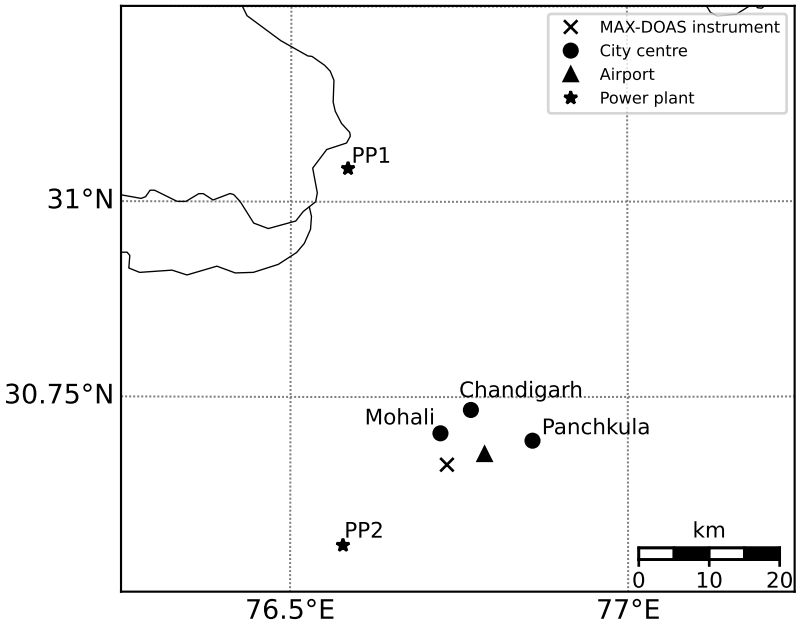


Figure 4.13: Overview of the locations of the Mohali instrument and its most important surroundings. The location of the MAX-DOAS instrument is marked by a cross, while city centres are marked by circles. Further, the positions of the Chandigarh Airport (triangle) and the two power plants (stars) are indicated. The map was created using Basemap (MIT License) with GMT data.

Mohali has a humid subtropical climate which is strongly influenced by a summer monsoon season. Slightly deviating from the definition of the seasons by the Indian Meteorological Department, in this work an emission-based definition of the four seasons in Mohali according to Kumar et al. (2020) is used: winter (November to February), summer/pre-monsoon (March to June), monsoon (July and August) and post-monsoon (September and October). Kumar et al. (2016) suggest a further classification of summer and post-monsoon into cleaner and polluted periods, nevertheless, this is not applied here. An overview of annual temperature (a) and wind (b) statistics for the measurement period presented in this work (April 2019 to July 2022) is depicted in Fig. 4.14.

Winter, especially December and January, is associated with frequent occurrences of fog and haze and rather mild winter temperatures, e.g. the monthly mean temperature in January is 12 °C. Followed by a rapid increase of the temperatures during summer, reaching the maximum monthly mean temperature in June (32 °C). During monsoon only a slight decrease of the temperatures (monthly means of 29 °C) can be observed, however, the rather low spread between the 5th and 95th percentiles indicates frequent cloudy scenes. Average monthly rain rates during monsoon are in the order of 300 mm. Finally, the post-monsoon season is

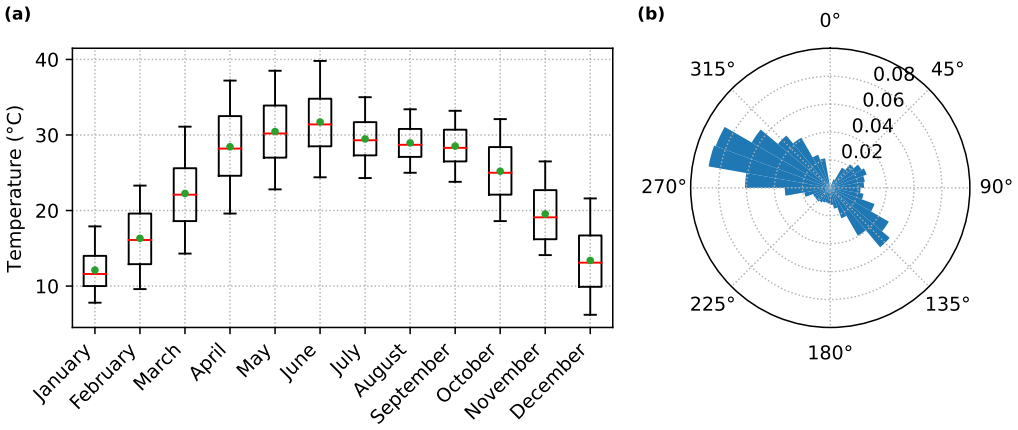


Figure 4.14: Weather statistics for the period from April 2019 to July 2022 obtained by the co-located weather station at Mohali. Panel (a) shows box-whisker-plots for the annual cycle of measured air temperatures. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. The 5th and 95th percentiles are shown by the whiskers of the boxes. Panel (b) shows the relative frequency distribution of the measured wind directions in bins of 10°.

associated with decreasing temperatures. The relative frequency distribution of the wind directions (Fig. 4.14b) shows two prominent maxima. During the main part of the year, the primary wind direction is west to north-west and air masses typically originate from the Asian continent. A sudden change of the main wind direction can be observed in the monsoon season, where the main fetch region is south-east.

Despite overall high anthropogenic pollution due to the large population in the *"Tricity"*, pollution levels at the Mohali site are expected to also depend on the season. Typically large-scale crop residue burning occurs in late summer and late post-monsoon, especially in the rural westerly sector. During winter, private fires for domestic heating and cooking are also known to have a strong influence on the local air quality, leading to high NO_x and aerosol abundances (Sinha et al., 2014; Kumar et al., 2020). Slightly lower pollution levels are expected during monsoon.

First MAX-DOAS measurements at Mohali were performed from January 2013 to June 2017 whose results are discussed in Kumar et al. (2020). Following these measurements, a laptop-based version of the TubeMAX-DOAS instrument (compare Sect. 4.2.1) was installed in Mohali at the IISER (30.664 °N, 76.732 °E; 310 m a.m.s.l.) in April 2019 at around 16 m above ground level with an azimuth viewing direction of 350° from north. The measurements were finally started on 24 April 2019 and data are available since then, only having some smaller gaps caused by instrumental issues. Further details of data availability and quality will be discussed in Chapt. 12. Besides the MAX-DOAS measurements, basic meteorological quantities are measured by a co-located weather station (compare Fig. 4.14). Additionally, in-situ measurements of ozone, nitrogen oxides (NO and NO_2) and different species of VOCs were performed by the local partners. A

more detailed description of the ancillary measurements can be found in Sinha et al. (2014).

4.4 ATTO, BRAZIL

The Amazon Tall Tower Observatory (ATTO; 2.146 °S, 59.006 °W; 130 m a.m.s.l.) is located about 150 km north-east of Manaus, Brazil, in the central Amazon Basin. Situated on a plateau, approximately 130 m a.m.s.l., the dense rainforest surrounding the observatory is considered to be mature, untouched and typical for a non-flooded *terra firme* forest with an average canopy height of 35 m (Andreae et al., 2015; Gomes Alves et al., 2023). Especially in the sector from north to east, for several hundreds of kilometres only smaller settlements are present in otherwise undisturbed forests (Andreae et al., 2015). In this direction, the closest areas with significant human activity are around 1000 km away located on the north-eastern coast of South America. The main stem of the Amazon river dominates the easterly sector, where smaller towns are located along the river (Andreae et al., 2015). The cities of Santarém and Belém are at distances of about 500 km and 1000 km, respectively. Also in the south-eastern direction mainly rainforests cover the land, while areas with higher population lie at distances greater than 1000 km. Around 12 km south of ATTO, the Uatumã river with several smaller communities along it can be found, while ca. 60 km to the north-west Balbina, the closest town, is located at the dam of the Balbina reservoir. However, as outlined below, winds from these directions occur rarely. A map showing the locations of ATTO and Manaus within the Amazon rainforest on the South American continent is provided in Fig. 4.15.

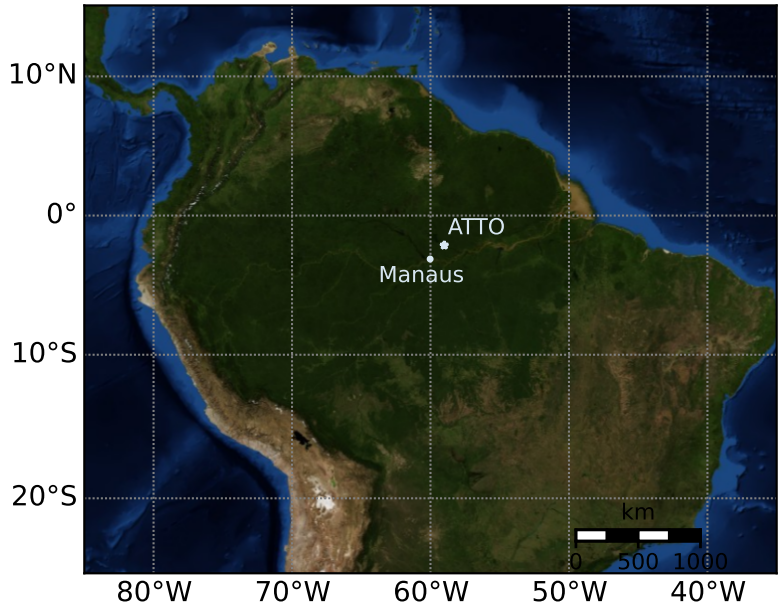


Figure 4.15: Overview of the locations of ATTO and Manaus within the Amazon rainforest on the South American continent. The map was created using Basemap (MIT License) with NASA Blue Marble data (http://visibleearth.nasa.gov).

The climate at ATTO is tropical humid with a mean annual precipitation of 2380 mm (Gomes Alves et al., 2023) and mainly dominated by pronounced wet and dry seasons. Slightly deviating from the meteorological definition based on precipitation rates, in this work a definition of the seasons based on meteorology and pollution levels, also applied in other studies (Pöhlker et al., 2016; Pfannerstill et al., 2021), is used: wet season (February to May), wet to dry transition (June and July), dry season (August to November) and dry to wet transition (December and January). An overview of the statistics of temperatures and wind directions measured[3] at ATTO in the period from March 2018 to April 2022 is presented in Fig. 4.16.

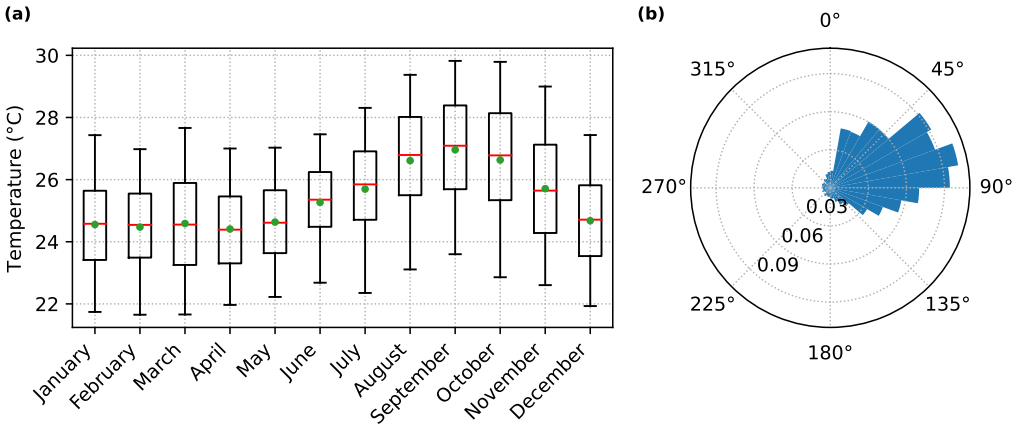


Figure 4.16: Weather statistics for the period from March 2018 to April 2022 obtained by the weather station placed on top of the tall tower (325 m) at ATTO. Panel (a) shows box-whisker-plots for the annual cycle of measured air temperatures. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. The 5th and 95th percentiles are shown by the whiskers of the boxes. Panel (b) shows the relative frequency distribution of the measured wind directions in bins of 10°.

Typical for a tropical climate, the measured temperatures show only a weak annual cycle, having its minimum during the wet season in April and its maximum during the dry season in September, where monthly mean temperatures of 24.5 °C and 27 °C are obtained, respectively. Rapid changes in cloud coverage caused by the alternation between strong convective precipitation and undisturbed sunshine are characteristic in such tropical climates. The influence of this behaviour on the DOAS retrieval will be discussed in Sect. 9.3. The dominant wind directions at ATTO are north-east to east as can be seen in Fig. 4.16b. Depending on the actual position of the ITCZ during the course of the year, north-easterly winds prevail during the wet season, while more easterly winds occur during the dry season. This characteristic change of the main wind direction was also confirmed by backward trajectory analyses performed by Pöhlker et al. (2019).

Given its remote and pristine location outlined in detail above, the ATTO site can be considered as clean in the context of anthropogenic pollution. During

3 Note that the temperature and wind measurements were performed on top of the tall tower at an altitude of 325 m which is not typical for the measurements of these quantities.

the wet season air masses originating from regions, which are mainly covered with pristine rainforest, are transported to the site by north-easterly winds. Also episodes of long-range transport of Saharan dust across the Atlantic towards ATTO can occur in this period (Andreae et al., 2015). In the dry season, however, the site is influenced by air masses from easterly and south-easterly fetch regions, which can contain significant contributions of pollution from biomass burning and other anthropogenic sources in north-eastern Brazil. As shown by the trajectory analyses by Pöhlker et al. (2019), the influence of the pollution plume of Manaus at ATTO under southerly winds can be neglected since such conditions occur in only 1.4 % of the cases investigated in their study.

The ATTO site was originally established in 2010 as a long-term measurement station for aerosol, trace gas, meteorological, and ecological studies in the Amazon forest (Pöhlker et al., 2019). The core of the ATTO site is the 325 m high so-called "tall tower" which got operational in 2017. Already in 2010 and 2011 a 81 m high mast and a 80 m walk-up tower, respectively, were constructed at 1 km distance from today's tall tower. The three towers are equipped with various meteorological and environmental measuring systems performing, amongst others, measurements of reactive gases (e.g. O_3, NO_x and VOCs), greenhouse gases, aerosol properties, micrometeorology and phenology. Most importantly for the studies presented in this work, a weather station is installed at the top of the tall tower and an AERONET station is positioned on the 80 m walk-up tower. For further details on the ATTO site, refer to Andreae et al. (2015).

After completion of the tall tower, a first RaspberryPi-based TubeMAX-DOAS instrument (Sect. 4.2.1) was installed on the railing of the 80 m platform in October 2017 looking in an azimuth direction of 50° from north. Data are available since 12 October 2017, however, especially in the initial phase of the measurements several instrumental issues, leading to larger data gaps, occurred which can be attributed to the hot and humid climate on-site. End of March 2019 a second TubeMAX-DOAS instrument was installed at the 298 m platform having the same viewing direction as the 80 m instrument. Starting its measurements on 29 March 2019, also this instrument suffered from instrumental problems. Further details of data availability and quality will be discussed in Chapt. 13. In the following, these two instruments are referred to as the 80 m and 298 m ATTO instruments, respectively.

4.5 RESEARCH CRUISE MSM 68/2, NORTH ATLANTIC

Starting on 3 November 2017 in Emden, Germany, research cruise MSM 68/2 aboard the research vessel RV MARIA S. MERIAN crossed the English Channel, went through parts of the North Atlantic and ended in Mindelo, Cape Verde, on 14 November 2017. An overview of the exact route taken by the ship is depicted in Fig. 4.17.

Besides MAX-DOAS measurements which were performed using a laptop-based TubeMAX-DOAS instrument with in-built angle stabilisation (compare Sect. 4.2.1), several other atmospheric and oceanographic observations were

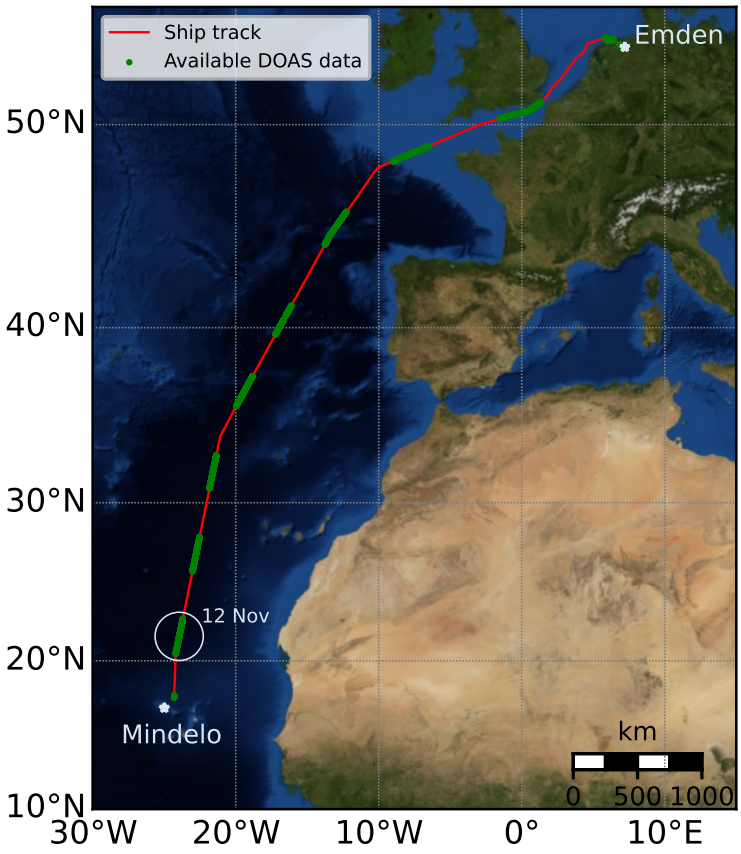


Figure 4.17: Overview of the route taken by RV MARIA S. MERIAN during research cruise MSM 68/2 (red line). Available DOAS data (only during day) are displayed in green, while a grey circle marks the vessels position on 12 November 2017. The map was created using Basemap (MIT License) with NASA Blue Marble data (http://visibleearth.nasa.gov).

conducted during the cruise. A detailed description of the research cruise itself, its main objectives and the results can be found in Fiedler (2018).

As can be seen from the map presented in Fig. 4.17 and since no significant biogenic sources are expected in this region of the North Atlantic, most of the MAX-DOAS data have been collected far from any anthropogenic influence and hence provide an interesting data set under pristine background conditions. The MSM 68/2 data set is not used for detailed atmospheric investigations in this work. However, given the rare pristine measurement conditions, spectra collected on one selected day, namely 12 November 2017, are used in the sensitivity study for the DOAS fit settings of HCHO (Chapt. 6) as an example for real measured spectra with very low NO_2 absorptions and medium HCHO levels.

Many spectra fulfilling several quality criteria were collected on that day indicating a rather nice day with only a few high clouds, which is also confirmed by on-board measurements of global radiation. Air temperatures were around 25 °C on that day, while the mean NO_2 and HCHO VCDs taking only valid data into account (compare Sect. 3.3) were 1.3×10^{14} molec cm^{-2} and 2.8×10^{15} molec cm^{-2},

respectively. Despite this in general favourable measurement conditions with low NO_2 and medium HCHO levels, it has to be pointed out that in the afternoon of this day the vessel headed into regions that were influenced by a dust storm event originating from the African continent. As a consequence, the AOD at 360 nm increased from values below 0.2 during the day to 0.5 shortly before sunset and AODs of 0.6 were measured in the morning of the following day. Nevertheless, the data of 12 November 2017 still seem to be best suited for the intended purpose to serve as an example for real measured spectra with almost no NO_2 absorptions for two reasons. First, the measurements during the beginning of the cruise were influenced by anthropogenic pollution close to the coasts of the European continent and Great Britain with NO_2 VCDs up to 1.5×10^{16} molec cm^{-2}. Second, sky scenes with many broken clouds were present on most of the other available days in pristine conditions.

Details of how the data were used and how the different analyses were performed will be described when needed in Chapt. 6.

5 ADDITIONAL DATA SETS

Besides the MAX-DOAS data collected at the four measurement sites described in the previous chapter, external data sets were used for the studies presented in this work. The following sections provide brief overviews of these additional data sets and their relevance for this work.

5.1 ERA5 DATA

The fifth generation ECMWF ReAnalysis (ERA5) is a global reanalysis data set of atmospheric (and oceanographic) variables developed and provided by the European Centre for Medium-Range Weather Forecasts (ECMWF; Hersbach et al., 2023). Amongst others, it contains a comprehensive sample of key meteorological quantities gridded at $0.25° \times 0.25°$ (corresponding to ca. $30\,\text{km} \times 30\,\text{km}$ at the equator) spatial resolution. The vertical dimension is resolved by 137 atmospheric layers reaching from the surface (lowest model layer represents the lowest 10 m) to 80 km (or 0.01 hPa) and data are available on hourly basis. It should be mentioned that the native model time step is 12 min which is necessary to resolve the involved physical processes properly and that the model calculations are performed with a spatial resolution of 31 km corresponding to 0.281 25° (ECMWF, 2024). For further details on uncertainty estimates as well as model setup and properties, refer to Hersbach et al. (2023).

In this work, a preprocessed version of the ERA5 data which is typically used in the Satellite Remote Sensing Group of the Max Planck Institute for Chemistry (MPIC) is used. This particular version of ERA5 is downloaded on a downsampled $1° \times 1°$-grid with a six hour time step and the 137 original model layers. These coarser resolutions have been shown to be sufficient for many applications. Moreover, only selected atmospheric variables relevant for these specific applications are included in this modified data set. In particular, three-dimensional fields of the wind components, pressure, temperature and water vapour concentration are of relevance here.

The ERA5 data set is utilised for several aspects in this work. In particular, vertical profiles of pressure, temperature, horizontal wind components and water vapour concentrations at the respective measurement site are extracted during the generation of the MAPA input from the MAX-DOAS data. During this preprocessing, the meteorological fields are temporally interpolated to the timestamps of the MAX-DOAS measurements (uneven grid and only during day) and brought to a finer vertical grid. In the preprocessing step of the input data during the MAPA retrieval itself, the vertical profiles of pressure, temperature and water vapour concentrations are used to calculate the a-priori O_4 VCDs needed for the aerosol retrieval (compare Sect. 3.3 and Beirle et al., 2022). Besides that, the investigations on the influence of elevation biases on the MAPA results presented

in Sect. 9.2.2 include a comparison of the water vapour concentration profiles obtained with MAPA to the ones extracted from the ERA5 data. More importantly, in the main part of this work the ERA5 data serve as common consistent data set of meteorological variables at the different measurement sites to study the dependencies of trace gas and aerosol abundances on temperature as well as on wind speed and wind direction (compare corresponding sections in Chapt. 10 to 13). An investigation on how far the ERA5 data set is representative for the actual temperatures, wind speeds and wind directions measured at the different stations given its rather coarse spatial and temporal resolutions is provided in Sect. 9.4. Descriptions of the data handling for the different applications are given in the respective sections.

5.2 AERONET DATA

The aerosol robotic network (AERONET) is a union of several ground-based remote sensing aerosol networks led by NASA and PHOTONS (Photométrie pour le traitement opérationnel de normalisation satellitaire; University of Lille) supported by various institutions and is already operational for more than 25 years (NASA-GSFC, 2024b). AERONET provides globally distributed and highly standardised observations of aerosol optical depth (AOD), optical aerosol properties and precipitable water content of the atmospheric column using photometer measurements of direct and scattered sunlight. The resulting data products have a temporal resolution of at least 15 min and include AODs at wavelengths[1] of 340, 380, 440, 500, 675, 870 and 1020 nm, the precipitable water (PW) content of the atmospheric column[2] in centimetres retrieved at 940 nm and the Ångström exponent (Holben et al., 2001). The latter describes the wavelength dependence of the AOD (Ångström, 1929) and is calculated for different spectral ranges as the slope of the linear regression of $\ln \mathrm{AOD}$ versus $\ln \lambda$ in the respective spectral ranges (Holben et al., 2001). Further, aerosol properties such as aerosol size distributions are inferred from the AERONET measurements but are not used in this work. The resulting data products are available at three quality levels. Level 1.0 data use the pre-field sun calibrations and are unscreened, while level 1.5 data are obtained applying a cloud-screening and automatic quality control checks. The final level 2.0 data are generated by additionally applying the post-field sun calibrations (NASA-GSFC, 2024a). For further details on data acquisition, quality assurance and properties retrieved from the radiance measurements, refer to Holben et al. (1998, 2001).

In this work, mainly the AERONET AODs and the PW content of the atmospheric column are used to compare the corresponding MAX-DOAS results to an independent data set. In particular, the AODs and/or water vapour VCDs retrieved from the MAX-DOAS measurement using MAPA are compared to the corresponding AERONET results to asses the quality of the MAPA retrievals.

1 Some stations provide AODs at additional wavelengths. However, the here-mentioned ones are the standard configuration.

2 This quantity is defined as the height of water, if all the water vapour contained in an atmospheric column was precipitated as rain and is thus easily transferable into a water vapour VCD (compare Eq. 5.2).

These comparisons are for instance used to decide on the different MAPA settings (compare Sect. 9.2.1), the best suited O_4 scaling factor (compare Sect. 9.2.1 and 9.3) and the appropriate elevation bias (compare Sect. 9.2.2). Further, in the main part of this work (Chapt. 10 to 13) the annual variations of AODs and water vapour VCDs obtained by AERONET and MAX-DOAS are compared to each other.

As described in the previous chapter (Chapt. 4), all MAX-DOAS stations considered in this work have a collocated AERONET station except for Mohali where no AERONET data in the vicinity of the MAX-DOAS instrument are collected. Hence, for Mohali, AERONET data were taken from the station at the Amity University Gurugram in Haryana located approximately 45 km south-west of New Delhi and around 250 km south of Mohali itself. Consequently, no good agreement between the MAX-DOAS and the AERONET data is observed for the AODs, while reasonable agreement is found for water vapour (compare Fig. 12.7 below and Fig. B.11 in the appendix). Since not for all stations level 2.0 data were available for the entire measurement periods at the time of the analyses and a consistent data set is desired, level 1.5 data (compare above) are exploited here. Note that no AERONET AODs are available under conditions where the sun is obstructed by clouds as they are retrieved from direct sun measurements. Contrary to that, MAX-DOAS measurements might still yield valid results under such conditions.

To compare AERONET and MAX-DOAS AODs, first the Ångström exponent (γ) for the spectral range from 340 to 380 nm is calculated[3] using the AODs at both wavelengths and solving

$$\frac{\mathrm{AOD}_\lambda}{\mathrm{AOD}_{\lambda_0}} = \left(\frac{\lambda}{\lambda_0}\right)^{-\gamma} \tag{5.1}$$

which describes the relation between AODs and wavelengths (λ) for γ. Once the Ångström exponent is known, the AERONET AOD at 380 nm can be scaled to 367 nm (or 357 nm) which is the wavelength used for the MAPA aerosol inversion (compare Sect. 9.2.1), again using Eq. 5.1.

Since AERONET provides the PW content of the atmospheric column in centimetres, this quantity has to be transferred into a water vapour VCD. This is straightforward and basically just a scaling with the density of water, its molar mass and the Avogadro constant yielding

$$VCD_{\mathrm{H_2O}} = PW \cdot 3.34 \times 10^{22}\,\mathrm{molec\,cm^{-3}}. \tag{5.2}$$

Ultimately, both data sets are sampled on a common time grid which differs for the different specific applications, i.e. data are binned in 15-minute bins, hourly, daily, or monthly.

3 Note that also AERONET provides Ångström exponents for various spectral ranges. However, the one closest to the wavelengths considered by the MAPA retrievals (357 and 367 nm) is calculated for the spectral range between 340 and 440 nm. To have more accurate results, the Ångström exponents were calculated as described in the text.

5.3 TROPOMI FORMALDEHYDE DATA

The TROPOspheric Monitoring Instrument (TROPOMI) was launched on board the Sentinel-5 Precursor (S5P) satellite on 13 October 2017 and flies in a low-Earth afternoon polar orbit with a local equator crossing time of 13:30 LT (Veefkind et al., 2012). Like its predecessor instrument OMI (Levelt et al., 2006), TROPOMI has a 2600 km wide swath providing an almost daily global coverage, but at a much finer spatial resolution of originally $3.5 \times 7\,km^2$ at nadir which has been further refined to $3.5 \times 5.5\,km^2$ on 6 August 2019 (ESA, 2023b). Equipped with an imaging spectrometer measuring in the UV, visible, near-infrared and short-wave-infrared spectral regions the operational level 2 (L2) products include VCDs of O_3, SO_2, NO_2, HCHO, CO and CH_4 as well as cloud and aerosol properties (Veefkind et al., 2012).

In this work, a modified version of the operational TROPOMI L2 HCHO product (reprocessed (RPRO) + Offline (OFFL), version 2 + 2.4.1, ESA, 2020) similar to De Smedt et al. (2021) is used which is only available from May 2018 onwards. The HCHO retrieval is based on the algorithm described in detail by De Smedt et al. (2018) which was adapted several times as reported in the official S5P product description. The DOAS fit yielding HCHO dSCDs is performed in the fitting interval from 328.5 to 359 nm using daily averaged Earth radiances measured in the equatorial Pacific region (on the previous day) as reference spectra. Further, cross-sections of HCHO, NO_2, O_3 (at two temperatures and first-order Puḳīte terms as described in Sect. 3.1.5.2), BrO and O_4 are included in the fit as well as a pseudo absorber accounting for the Ring effect (compare Sect. 3.1.5.1). The conversion of the dSCDs to VCDs is performed using AMFs simulated at 340 nm and accounting for a possible HCHO background in the reference sector over the equatorial Pacific (De Smedt et al., 2021). To circumvent issues related to the cloud correction in the simulation of the AMFs, so-called clear-sky AMFs are used for which no cloud correction is applied. Consequently, the TROPOMI HCHO VCDs used in this work are so-called clear-sky VCDs. For these, measurements with cloud fractions above 0.4 are excluded. In addition to this cloud filtering, the TROPOMI data are filtered based on the quality assurance (QA) values (ESA, 2023a). Here, only data with QA-values above 0.5 are considered which for instance removes most measurements with SZAs above 70°, AMFs below 0.1 or surface albedos above 0.2. For further details on the TROPOMI HCHO retrieval as well as data selection and handling, refer to De Smedt et al. (2018, 2021).

To compare the HCHO VCDs retrieved from the MAX-DOAS measurements at the four stations considered in this work with the TROPOMI VCDs, daily overpass files were created by De Smedt (2023) starting from daily gridded data at $0.05° \times 0.05°$ resolution and applying the described filters. From these gridded data, daily averaged TROPOMI HCHO VCDs in a radius of 20 km around the different measurement sites were calculated like described in De Smedt et al. (2021). Note that in addition to the filters described above, at least 50 % valid grid cells in the respective domains are required.

These daily averaged HCHO VCDs are compared to the corresponding MAX-DOAS VCDs in the respective sections in Chapt. 10 to 13. Descriptions of how the MAX-DOAS data are handled for these comparisons and how the latter are performed are provided in Sect. 10.6. However, no further filtering or modifications are applied to the satellite data.

Part III

IMPROVEMENT OF MAX-DOAS RESULTS

6

SENSITIVITY STUDY FOR THE DOAS FIT SETTINGS OF FORMALDEHYDE

The MAX-DOAS instruments whose measurements are discussed in this work are distributed in different parts of the globe and the measurements are taken with a large variety of underlying meteorological and environmental conditions, e.g. ambient temperatures or levels of anthropogenic pollution (compare Chapt. 4). As a consequence, specific DOAS fit settings adjusted to the local atmospheric conditions have been used for some of the different measurement sites in the past. Also in existing literature dealing with MAX-DOAS measurements of tropospheric HCHO a variety of different DOAS fit settings for HCHO can be found (e.g. Pinardi et al., 2013; Vlemmix et al., 2015; Behrens et al., 2019; Wang et al., 2019; Kumar et al., 2020; Hong et al., 2022). Since HCHO sometimes occurs only in small abundances in the atmosphere and its absorption structures are comparably weak (e.g. the absorption features of NO_2 are 10 times larger than the ones of HCHO in their typical fit intervals), the retrieval results can strongly vary depending on the chosen DOAS fit settings. Thus, a systematic sensitivity study was performed to find the most suitable DOAS fit settings which can be used for the retrieval of HCHO dSCDs for all or at least most atmospheric conditions. This study also allows the quantification of uncertainties which are introduced by deciding on specific DOAS fit settings.

The design of this sensitivity study, its implementation and results are presented in this chapter.

6.1 USED DATA SETS AND STUDY APPROACH

In order to study the sensitivity of the HCHO fit results to different fit parameters, real measured and simulated synthetic spectra were used. The latter allow a direct comparison between the retrieval results obtained with varying DOAS fit settings and the reference values which have been used to simulate these spectra. Contrary to that, spectra measured with a typical MAX-DOAS instrument allow an identification of advantages and disadvantages of specific analysis settings under real measurement conditions, since also instrumental effects and other external factors possibly influence the fit results.

It should be mentioned that also Pinardi et al. (2013) performed a sensitivity study for the DOAS fit settings of HCHO. However, their study was only based on spectra measured at one selected day during the CINDI field campaign (Piters et al., 2012). Further, their study focused on the harmonisation and consistency amongst the participating instruments and therefore was for instance limited to a certain set of spectral ranges.

The measured and synthetic spectra used for the study presented here are described in Sect. 6.1.1, while Sect. 6.1.2 gives an introduction to the approach of this study.

6.1.1 *Synthetic and measured spectra*

The synthetic radiance spectra which are used in this study were simulated at a high spectral resolution of 0.01 nm for a large set of scenarios using the radiative transfer model SCIATRAN (Rozanov et al., 2014). One advantage of this RTM is the possibility to include rotational Raman scattering in the simulations which causes the so-called Ring effect (compare Sect. 3.1.5.1). Thus, in this study it is also possible to study the influence of the Ring correction on the DOAS analysis. The different measurement geometries as well as the trace gas scenarios which were used for the simulations of the synthetic spectra are summarised in Table 6.1.

Table 6.1: Overview of the measurement geometries and atmospheric scenarios for the simulations of the synthetic spectra.

PARAMETER	SCENARIOS
Wavelength range	303 – 470 nm
SZA	10°, 20°, 45°, 60°, 75°, 85°
RAA	0°, 5°, 10°, 30°, 90°, 150°, 170°, 175°, 180°
Elevation angles	1°, 3°, 10°, 30°, 90°
HCHO VCD (0 - 1000 m)	0, 2×10^{15} and 6×10^{15} molec cm^{-2}
NO_2 VCD (0 - 500 m)	1×10^{15} and 2×10^{16} molec cm^{-2}
NO_2 VCD (stratosphere)	5.22×10^{15} molec cm^{-2}, profile with maximum at 24 km
BrO VCD	3.03×10^{13} molec cm^{-2}, stratospheric profile with maximum at 20 km
O_3 VCD	9.29×10^{18} molec cm^{-2}, US standard atmosphere with maximum at 22 km
H_2O VCD	0.5×10^{23} molec cm^{-2} and 2.5×10^{23} molec cm^{-2}, exponential profile with scale height H = 2 km
Aerosol (0 - 1000 m)	AOD at 350 nm: 0.1 and 0.5, phase-function and single scattering albedo for S4 biomass burning aerosol
Surface albedo	0.07
Surface elevation	0 m

Table 6.2 lists the reference trace gas absorption cross-sections used for the simulations of the synthetic spectra. Also the solar reference spectrum is listed.

Table 6.2: List of reference trace gas absorption cross-sections and solar reference used for the simulations of the synthetic spectra.

TRACE GAS	CROSS-SECTIONS
HCHO	Meller and Moortgat (2000), 298 K
BrO	Wilmouth et al. (1999), 228 K
NO_2	Vandaele et al. (1998), 220 K and 294 K
O_3	Serdyuchenko et al. (2014), 193 to 293 K in steps of 10 K
O_4	Thalman and Volkamer (2013), 203 K and 293 K
H_2O	Rothman et al. (2010), 293 K, scaled as suggested by Lampel et al. (2015)
Solar reference	Chance and Kurucz (2010)

Note that for trace gases which are not only present in one atmospheric layer, but in the troposphere and in the stratosphere (e.g. O_3 which shows prominent vertical structures), the temperature dependence of the absorption was accounted for in the simulations. In a DOAS analysis, however, this temperature dependence is typically only approximated by using absorption cross-sections at two or three different temperatures or not considered at all.

After the simulations, the synthetic spectra were convolved using a Gaussian SLF with a FWHM of 0.65 nm which is a typical value for the MAX-DOAS instruments used in this work (compare Table 4.1) and sampled on a 0.1 nm grid. Since measured spectra also contain noise which is caused by photon noise and the electronics of the instrument (see Sect. 4.2.1), also this effect is considered here. Therefore, to study the influence of noise on the fit results, noisy spectra were created for each simulated scenario by adding 100 different random noise spectra with an RMS of 3.4×10^{-4} to the radiance spectra. This RMS is a typical value for the here-used MAX-DOAS instruments. In that way, 100 different spectra with and one spectrum without noise are available for each scenario.

The measured spectra were taken from three measurement sites, namely Bayfordbury, the MSM 68/2 ship cruise and ATTO (compare Sect. 4.2, 4.5 and 4.4). Since Bayfordbury is frequently influenced by polluted air masses with high NO_2 abundances, its measurements serve as examples for spectra containing high NO_2 absorptions. Contrary to that, the spectra collected during the MSM 68/2 ship cruise and at ATTO contain only little to almost no NO_2 absorptions since these measurements have been taken under pristine background conditions on the North Atlantic and in the Amazon rainforest, respectively. For Bayfordbury, data from 29 March 2019 was used which was a clear and sunny day with only very few high cirrus clouds. NO_2 VCDs on that day ranged from ca. 1.2×10^{16} to 3.5×10^{16} molec cm^{-2} and are therefore representative for the high NO_2 scenario (2×10^{16} molec cm^{-2}). The formaldehyde abundances increased from approximately 3.5×10^{15} to 6×10^{15} molec cm^{-2} and, thus, lie between the low and high HCHO scenarios (2×10^{15} and 6×10^{15} molec cm^{-2}). Spectra recor-

ded on 12 November 2017 were used for the MSM 68/2 ship cruise which also was a rather clear day with only few clouds. Measured NO_2 VCDs on that day stayed below ca. 1×10^{14} molec cm^{-2} and are therefore more than one order of magnitude lower than the low NO_2 scenario (1×10^{15} molec cm^{-2}) and can be considered as effectively a no NO_2 scenario. The HCHO abundances, however, are representative for the low HCHO scenario (2×10^{15} molec cm^{-2}) as the VCDs varied between 1.7×10^{15} and 4.2×10^{15} molec cm^{-2} on that day. For ATTO, data of one complete year, namely 2020, measured by the 80 m instrument (compare Sect. 4.4) were used. As presented in Chapt. 13, monthly mean NO_2 VCDs are below roughly 3×10^{14} molec cm^{-2} throughout the year and, thus, are almost one order of magnitude lower than the low NO_2 scenario (1×10^{15} molec cm^{-2}). The HCHO VCDs, on the other hand, are high with monthly means above ca. 6×10^{15} molec cm^{-2} (equals the high HCHO scenario) during the entire year.

6.1.2 *Study approach*

Quantitative measures are needed to investigate the sensitivity of the DOAS fit results to changes in different fit parameters. For synthetic spectra, the HCHO dSCDs retrieved with varying fit settings can be compared to the dSCDs that should be retrieved based on the VCDs that were used to simulate the spectra. To convert these input VCDs to reference dSCDs (compare Sect. 3.2.3), AMFs at wavelengths of 315, 320, 325, 330, 335, 340, 345, 350, 355 and 360 nm were simulated for each scenario listed in Table 6.1. To apply Eq. 3.29, the AMFs are converted to dAMFs using Eq. 3.30 ($dAMF = AMF_{\alpha} - AMF_{90°}$). These are then interpolated to the centre of the relevant fitting interval. Already here it should be noted that during studies regarding the wavelength dependencies of the AMFs, which are presented in Sect. 6.4.4, it was found that AMFs (or dAMFs) at wavelengths outside the absorption bands of formaldehyde should be excluded from the interpolation. In particular, these wavelengths are 335, 350 and 360 nm. Finally, the true reference dSCDs ($dSCD_{\text{true}}$) can be calculated from the input VCDs for each scenario and geometry using Eq. 3.29 and the interpolated dAMFs.

For the comparison of the retrieved dSCDs with the true reference values, absolute and relative deviations are calculated using

$$\Delta dSCD_{\text{abs}} = dSCD_{\text{retrieved}} - dSCD_{\text{true}} \tag{6.1}$$

and

$$\Delta dSCD_{\text{rel}} = \frac{dSCD_{\text{retrieved}} - dSCD_{\text{true}}}{dSCD_{\text{true}}} \cdot 100\,\%, \tag{6.2}$$

respectively. Here, $dSCD_{\text{retrieved}}$ denotes the dSCDs retrieved with varying fit settings and $\Delta dSCD_{\text{abs}}$ and $\Delta dSCD_{\text{rel}}$ indicate the corresponding absolute and relative deviations.

These deviations can be calculated for each elevation angle within a sequence of one specific trace gas and sun geometry scenario. In order to compare the results obtained with different fit settings for such a sequence, it is desirable to have only one representative quantity for each set of fit settings which describes

the deviations for the complete sequence. One possible quantity is the sum of the absolute deviations of the different elevation angles within this sequence which is defined as

$$\Sigma_{\text{abs}} = \sum_{\alpha} \Delta dSCD_{\text{abs}}(\alpha), \tag{6.3}$$

where α are the (off-zenith) elevation angles (1°, 3°, 10° and 30°) of one simulated sequence. Since over- and underestimations of the dSCDs can occur within one sequence there can be compensating effects and the above defined sum can become small, even if the individual deviations were large. Therefore, a second quantity, namely the root mean square (RMS) of the absolute deviations (RMS_{abs}) within one elevation sequence is used in the following investigations which is defined as

$$RMS_{\text{abs}} = \sqrt{\frac{1}{N} \sum_{\alpha} \Delta dSCD_{\text{abs}}(\alpha)^2} \tag{6.4}$$

with N being the number of considered elevations angles (here: $N = 4$). Combining these two quantities allows a classification of the obtained deviations. For instance, a small Σ_{abs} in combination with a large RMS_{abs} indicates large individual deviations for the different elevation angles together with the occurrence of over- and underestimations within one sequence (but not at which elevation these occur), while small values for both quantities indicate a good agreement of retrieved and true reference dSCDs at all elevations.

Besides these measures of deviations, other quantities such as the fit error, the fit RMS and, for noisy spectra, the scatter of the retrieved dSCDs for one scenario are used to assess the quality of the different fit settings.

For real measurement data, no reference values are known and thus no deviations can be calculated. Nevertheless, for these data still the other mentioned quantities are calculated and used to examine the quality of different sets of fit settings.

For large parts of the following investigations, often so-called standard DOAS fit settings were used. Then, individual parameters were successively changed in order to analyse the corresponding effects on the fit results. Table 6.3 summarises these standard settings which are mainly based on the HCHO fit settings used during the CINDI-2 campaign (Kreher et al., 2020). However, a longer fit interval was applied in the here-presented work and for some trace gases different absorption cross-sections were used to be consistent with the simulations. It should be stressed that the standard fit settings already include two O_3 cross-sections at different temperatures. As outlined in Sect. 6.2.3, this is justified even if a sequential FRS is used. Further, the absorption cross-sections of NO_2 and O_3 are convolved using the so-called intensity weighted convolution (Puķīte et al., 2021) which is described in Sect. 3.1.5.3. Lastly, the two pseudo absorbers accounting for the Ring effect for the synthetic spectra were calculated from the 90° spectrum for the case of low HCHO and NO_2 at an SZA of 45° and an RAA of 90°.

Table 6.3: Overview of the standard fit settings used in the sensitivity study.

PARAMETER	SETTINGS
Fitting interval	324 – 359 nm
Trace gas	
HCHO	Meller and Moortgat (2000), 298 K
BrO	Wilmouth et al. (1999), 228 K
NO_2	Vandaele et al. (1998), 294 K, intensity weighted convolution*
O_3	Serdyuchenko et al. (2014), 223 K and 243 K, intensity weighted convolution**
O_4	Thalman and Volkamer (2013), 293 K
Ring effect	Two pseudo cross-sections derived from universal FRS
FRS	90° spectrum of corresponding sequence/scenario
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

*After the optimisation of the NO_2 cross-sections in Sect. 6.2.2, additionally the cross-section at 220 K as well as the linear and square Puķīte terms are included.

**After the optimisation of the O_3 cross-sections in Sect. 6.2.3, additionally the linear and square Puķīte terms are included.

6.2 PUĶĪTE TERMS AND OTHER EFFECTS FOR NO_2 AND O_3 ABSORPTIONS

In this section, mainly the influence of Puķīte terms (compare Sect. 3.1.5.2) for NO_2 and O_3 absorptions on the HCHO fit results is investigated. In addition, the usage of Ring spectra as well as the impact of including additional absorption cross-sections at different temperatures for both species are examined.

The latter might be justified since the absorption features of the trace gases are temperature (and pressure) dependent. Since not always all fractions of a given species are located at the same altitude (and therefore temperature) in the atmosphere, the absorptions potentially occur at different temperatures. In the standard HCHO fit settings (compare Table 6.3), cross-sections at two temperatures (223 K and 243 K) for O_3 and at one temperature (294 K) for NO_2 are included. However, in many MAX-DOAS applications a sequential FRS is used (compare Sect. 3.2.2) and therefore, the stratospheric absorptions (at low temperature) cancel out and any remaining temperature effect is expected to be small, at least for NO_2.

As explained in Sect. 3.1.5.1, inelastic rotational Raman scattering on N_2 and O_2 molecules changes the wavelength of photons and causes a "filling-in" of the solar Fraunhofer lines, the so-called Ring effect. This effect is on the one hand accounted for by including Ring spectra in the DOAS fit. In addition to the "filling-in" of Fraunhofer lines, this effect also occurs for trace gas absorptions if

parts of the absorptions occur before the Raman scattering event where the light is passing through the atmosphere at a different wavelength after the scattering event (Puḳīte et al., 2021). For strong NO_2 or O_3 absorptions and the retrieval of relatively weak and/or less abundant absorbers in the same spectral range, this effect might have a significant influence. Hence, also NO_2 and O_3 Raman cross-sections (also referred to as trace gas Ring cross-sections) are calculated following the approach from Lampel et al. (2015) and Puḳīte et al. (2021) and tested in the DOAS fit to correct for this effect.

Besides these two effects, Puḳīte terms (compare Sect. 3.1.5.2) account for non-linear effects which play a significant role for medium and strong trace gas absorptions. These non-linearities are caused by the wavelength dependencies of the trace gas SCDs and the effective atmospheric light path even within a given fitting interval and are negligible for the retrieval of very abundant and/or strong absorbers. Therefore, they are often not accounted for in standard DOAS applications. However, in the UV spectral range, where HCHO is retrieved, the mentioned wavelength dependencies are particularly strong. Further, the O_3 cross-section shows strong absorption features in that spectral region, while HCHO is usually less abundant and its absorption structures are comparably weak. Therefore, such non-linear effects likely affect the spectral retrieval of HCHO, especially for low SCDs. As described in Sect. 3.1.5.2, Puḳīte et al. (2010) introduced a method to account for these effects during the DOAS fit itself by including pseudo absorption cross-sections, namely the linear and square Puḳīte (or higher order) terms, in the spectral retrieval. Besides O_3, also the higher order terms of NO_2 will be investigated in the following, since NO_2 shows strong absorptions in polluted air masses and might affect the HCHO results as well.

6.2.1 *Treatment of Puḳīte terms*

In the early stage of this sensitivity study, the Puḳīte terms were treated the way it has generally been done up to now, i.e. the "normal" cross-section of an absorber for which Puḳīte terms are used was convolved using an I_0-correction (compare Sect. 3.1.5.3) whereas the Puḳīte terms themselves were convolved without I_0-correction so far. However, it was then recognised that the inconsistent treatment of the "normal" cross-section and the Puḳīte terms is associated with several implications which are examined more closely in the following. In addition, this section also gives recommendations for the correct and consistent treatment of the Puḳīte terms in the DOAS analysis. Nevertheless, first the results of the treatment used so far are briefly presented as they have motivated several of the further investigations.

The first-order Puḳīte terms for an absorber a can be calculated using equations 3.16 and 3.17 (compare Sect. 3.1.5.2). While this is straightforward for the square term, the linear term is sometimes normalised to an arbitrary wavelength λ_0 to maintain the order of magnitude of the original cross-section for a stable DOAS fit. This λ_0 is typically chosen close to the centre of the considered fitting interval. Following this, the normalised linear Puḳīte term is defined as

$$\sigma_{\mathrm{lin,norm},a}(\lambda) = (\lambda - \lambda_0) \cdot \sigma_a(\lambda). \tag{6.5}$$

From a purely mathematical point of view, the fit results should be exactly the same no matter which way of calculation is used for the linear term. If all other fit parameters are kept the same, the fit coefficients of the "normal" cross-section and the corresponding linear term should compensate each other in a way that the sum of both stays exactly the same independent of the way of calculating the linear term. Moreover, the linear higher order term can in principle be calculated from the HR cross-section or the already convolved ones since the multiplication with λ or $\lambda - \lambda_0$ is just a linear operation and can therefore be permuted with the convolution integral.

Nevertheless, consistency checks during the early stage of the studies presented in Sect. 6.2.2 and 6.2.3 showed that the first expectation is not always fulfilled. The permutation of calculation and convolution of the Puķīte terms is in principle possible, however, also here, small, mostly negligible, deviations are introduced since during the convolution process of the HR cross-sections using the *QDOAS* software (Danckaert et al., 2017) an interpolation is performed.

Figure 6.1 shows HCHO dSCDs (upper panel) and the corresponding fit RMS (lower panel) as function of elevation for a sequence of synthetic spectra in the case of high NO_2, an SZA of 45° and an RAA of 180°. The analysed spectra did not contain noise. The results were obtained using the standard fit settings from Table 6.3 and adding the linear Puķīte term for NO_2 which was computed using each of the four previously described approaches to the DOAS fit. The reddish pair of lines shows the results using the non-normalised linear terms, whereas the greenish pair depicts the same for the normalised ones. A black line shows the true reference dSCDs based on the VCDs used for the simulations.

When comparing the red and magenta or the dark and light green lines, hardly any difference is visible for the permutation of calculation and convolution for both ways of calculation. A very detailed look shows that the difference is slightly larger for the normalised linear term (greenish lines), this is, however, barely visible at the y-axes scales in Fig. 6.1. Contrary to these marginal differences, especially at the two lowest elevation angles clear differences are found between the results obtained using the normalised and non-normalised linear Puķīte terms (compare pairs of reddish and greenish lines). This finding is valid for both dSCDs and RMS.

Figure 6.2 shows fit examples for the 1° spectrum of the elevation sequence used above for the different described scenarios. The columns represent the different ways of calculating the linear term, while the rows show the fit RMS (first row) and the fit results for HCHO (second row), NO_2 (third row) and the NO_2 linear term (fourth row).

Also here, hardly any differences are visible for the permutation of calculation and convolution with the results for the normalised linear terms showing marginally larger differences (compare third and fourth columns). The marginal observed differences can be explained by numerical noise which is introduced by the interpolation during the convolution using *QDOAS*. In contrast to that, clear differences are found when the fit results using the normalised and non-normalised terms are compared. For instance, a distinct residual structure at 333 nm collocated with an NO_2 absorption feature is visible in the two left

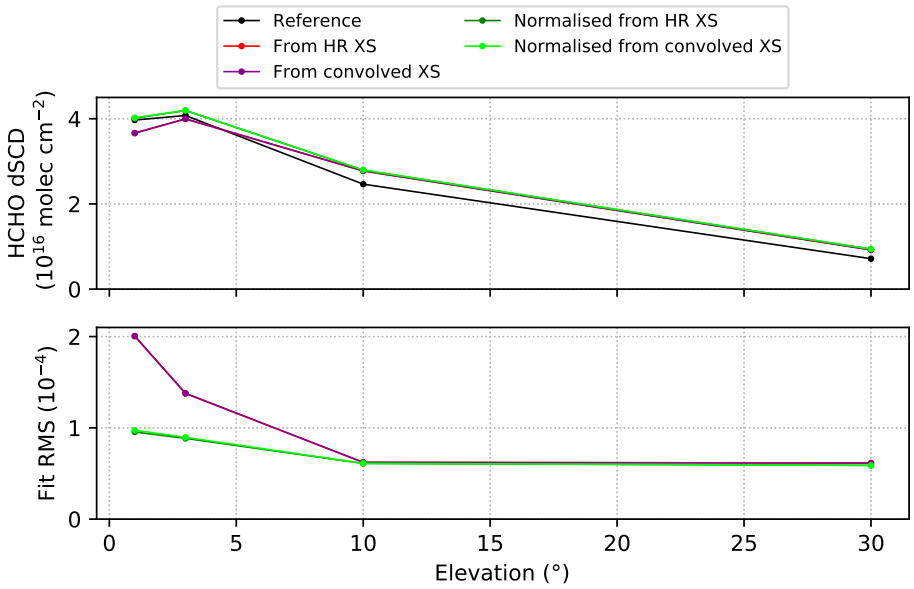


Figure 6.1: Retrieved HCHO dSCDs (upper panel) and fit RMS (lower panel) for an elevation sequence of synthetic spectra in the case of high NO_2, an SZA of 45° and an RAA of 180° (without noise) using the same settings but different ways of calculating the linear Puķīte term for NO_2. Different ways of calculating the linear term are indicated by different colours. The reference dSCDs are marked in black.

columns which is not visible for the normalised linear terms in the two right columns.

For the order of calculation and convolution, the clear recommendation is to calculate all Puķīte terms before the convolution from the HR cross-sections since the considerations in Puķīte et al. (2010) and Puķīte and Wagner (2016) are performed in HR space (compare Sect. 3.1.5.2).

The observed deviations between normalised and non-normalised linear terms can be explained by an inconsistent treatment of the "normal" NO_2 cross-section and the corresponding linear term. While in the above example the "normal" NO_2 cross-section was convolved using an I_0-correction (compare Sect. 3.1.5.3), the Puķīte terms so far have been convolved without applying an I_0-correction. It is thus desirable to treat all cross-sections consistently and in principle three options are available:

1. Convolve "normal" cross-sections and corresponding Puķīte terms using the standard convolution mode.

2. Perform I_0-corrected convolutions for "normal" cross-sections and the corresponding Puķīte terms.

3. Use the so-called intensity weighted convolution proposed by Puķīte et al. (2021) as described in Sect. 3.1.5.3 for "normal" cross-sections and the corresponding Puķīte terms.

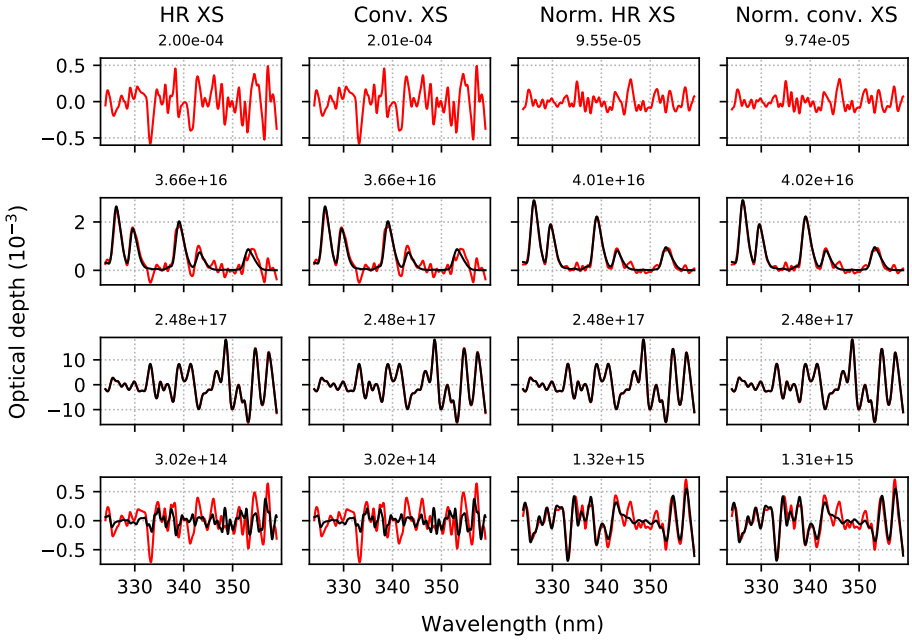


Figure 6.2: Fit examples for the 1° spectrum of the used elevation sequence (compare Fig. 6.1) for different approaches to calculate the linear Puķīte term (columns). The different rows show the fit residuals and the fit results for HCHO, NO_2 and the NO_2 linear Puķīte term, respectively. The numbers above the subplots provide the corresponding dSCDs or the RMS. Black lines indicate the fitted cross-section, while red lines show the fitted cross-section plus the residual or the residual itself (first row).

The first option is the easiest one, but especially for the retrieval of weak and/or less abundant absorbers like HCHO the usage of an I_0-correction for the convolutions of strong absorbers included in the fit, like NO_2 and O_3, is important as outlined in Sect. 3.1.5.3.

Since the highly resolved synthetic spectra are convolved using a Gaussian function with a FWHM of 0.65 nm, also for these the solar I_0-effect occurs and the second option appears to be reasonable.

For the "normal" cross-sections, the I_0-corrected convolution is straightforward and typically values of 1×10^{17} and 1×10^{20} molec cm^{-2} are used as typical SCDs (compare Sect. 3.1.5.3) for NO_2 and O_3, respectively. Already Wagner (1999) stated that the actual choice of the typical SCD for the "normal" cross-section is not critical. Nevertheless, also here a short investigation was performed and this statement can be confirmed. For the higher order terms, however, the I_0-corrected convolution is more subtle and the expression "slant column density" is not fully correct anymore. As a consequence of Eq. 3.21, the product $-SCD \cdot \sigma_{HR}$, also referred to as optical depth (compare Sect. 3.1.3), has to have the right order of magnitude. For the linear Puķīte term, the factor λ is increasing σ_{HR}. Therefore, the typical SCD was divided by the mean wavelength of the standard fitting interval (341.5 nm) and a value of 2.93×10^{14} molec cm^{-2} is obtained for NO_2. For the square term, the typical SCD has to be divided by the mean absorption strength of the trace gas within the fitting interval (3.97×10^{-19} cm^2 molec^{-1})

yielding 2.52×10^{35} molec cm^{-2}. Due to limitations in *QDOAS*, the HR square term was multiplied with 1×10^{10} and then convolved with I_0-correction using a SCD of 2.52×10^{25} molec cm^{-2}.

In addition, there is a second implication with the I_0-correction for absorbers for which Puḷīte terms are used since the "normal" cross-section of an absorber and the corresponding higher order terms can in principle not be convolved separately (compare Sect. 3.1.5.3). Therefore, Puḷīte et al. (2021) propose the so-called intensity weighted convolution for such absorbers (compare Sect. 3.1.5.3) which equals option three in the list above.

Figure 6.3 compares the HCHO dSCDs (upper panel) and the corresponding fit RMS values (lower panel) obtained with different treatments of the Puḷīte terms of NO_2. The reddish, greenish and brownish groups of lines indicate results obtained with the standard settings and additionally including the linear, normalised linear or square Puḷīte terms in the fit, respectively. Different colours within one group mark the different convolution types, namely the standard treatment with I_0-correction for the normal NO_2 cross-section and standard convolution for the Puḷīte terms as well as options two and three from the list above.

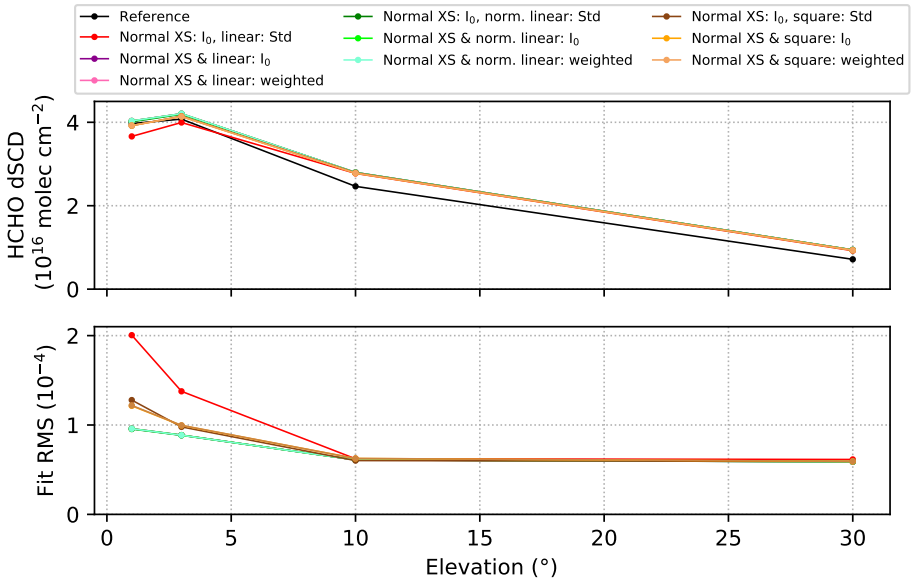


Figure 6.3: Retrieved HCHO dSCDs (upper panel) and fit RMS (lower panel) for an elevation sequence of synthetic spectra in the case of high NO_2, an SZA of 45° and an RAA of 180° (without noise) for different ways of convolving the NO_2 cross-section and its Puḷīte terms. The convolution types are indicated by the different colours within one group. The reddish, greenish and brownish groups of lines indicate the results obtained with the linear, normalised linear or square Puḷīte terms added to the standard fit settings, respectively. The reference dSCDs are marked in black.

A detailed comparison shows consistent results within the three groups of lines indicating that the different convolution types do not have a large impact. This is also valid for the fit RMS (lower panel). Nevertheless, there is one clear exception,

namely when the non-normalised linear term is convolved with the standard convolution, while the "normal" cross-section is convolved using an I_0-corrected convolution (red line). For this case, at the lowest two elevations, large differences to the other scenarios are found for dSCDs and RMS. At higher elevations, this effect becomes smaller but remains visible. As mentioned above, this treatment of the terms is not consistent and should be avoided. This is actually also valid for the normalised linear and the square terms, however, for these two terms the differences are smaller. All other reddish and greenish lines lie very close to each other indicating consistent results. In particular, the differences between I_0-corrected convolution and intensity weighted convolution for all NO_2 cross-sections are very small. The brownish lines mark cases where square Puḳīte terms are included in the fit instead of linear terms and show slightly different results. Also the RMS-values are larger for these scenarios. Nevertheless, the brownish lines show consistent results amongst each other. Comprehensive investigations on the topic which higher order terms should be included in the DOAS fit are performed below. Further, it should be stressed that despite generally consistent results for option two (I_0-correction for "normal" cross-section and Puḳīte terms), this way of treating the convolution is mathematically not correct.

Similar investigations were also conducted for the O_3 cross-section and its higher order terms and mainly the same findings as for NO_2 are valid. Also for O_3, the scenario where the non-normalised linear term is convolved in the standard mode, while the "normal" cross-section is convolved using an I_0-correction, shows largest discrepancies. The typical SCD which is needed for the I_0-corrected convolution is 1×10^{20} molec cm^{-2} for O_3 and values of 2.93×10^{17} molec cm^{-2} and 4.79×10^{40} molec cm^{-2} (mean absorption: 2.09×10^{-21} cm^2 molec^{-1}) are obtained for the linear and square Puḳīte terms, respectively. However, for the convolution with I_0-correction of the Puḳīte terms of O_3, difficulties can occur. Since the O_3 absorptions become rather large at low elevations, the product $-SCD \cdot \sigma_{HR}$ in the exponent of Eq. 3.21 (compare Sect. 3.1.5.3) gets large, especially for the square term. This leads to non-linearities during the I_0-corrected convolution and thus to problems with *QDOAS* which makes an appropriate scaling of the Puḳīte terms necessary. For instance, the square term is multiplied with 1×10^{36} and the typical SCD in the above equations is adjusted to 1×10^4 molec cm^{-2}. For the normalised and non-normalised linear terms, the typical SCD has to be scaled by the centre wavelength of the fitting window as explained above for NO_2. All in all, for O_3 the inconsistent application of Eq. 3.21 and separate convolution of "normal" cross-section and corresponding Puḳīte terms can become important. If the I_0-corrected convolution of the Puḳīte terms for O_3 is preferred, an appropriate scaling as outlined is needed but this approach is not recommended.

As a summary, the following recommendations for the treatment of the Puḳīte terms are given:

1. The Puḳīte terms should be calculated from the HR trace gas absorption cross-sections since also their derivation is performed in HR space (see also Puḳīte et al., 2010).

2. The intensity weighted convolution should be used for absorbers for which the "normal" cross-section and the corresponding Puḳīte terms are used. For specific applications, an I_0-corrected convolution can be used for all cross-sections of an absorber which is mathematically not correct.

3. Non-normalised linear terms according to Eq. 3.16 should be used since the normalisation to λ_0 is somehow arbitrary and consistent results are found for the normalised and non-normalised linear terms, when both are treated consistently and an intensity weighted convolution is used.

For the study presented in the following, all of these three recommendations have already been applied, i.e. the Puḳīte terms were calculated from the HR absorption cross-sections of the respective trace gases and all cross-sections of absorbers for which Puḳīte terms are included were convolved with an intensity weighted convolution. Also the respective non-normalised versions according to Eq. 3.16 were used for the linear Puḳīte terms.

6.2.2 *NO_2 absorptions*

The standard fit settings summarised in Table 6.3 include one NO_2 cross-section at 294 K. In this section, the impact of additional NO_2 cross-sections on the HCHO fit results is investigated. For air masses containing high NO_2 abundances, the NO_2 absorptions can become rather strong even in the UV spectral range. Thus, non-linearities as described in Sect. 3.1.5.2 can become important and the usage of NO_2 Puḳīte terms might be useful. Further, parts of the total NO_2 content are located at higher altitudes in the free troposphere or the stratosphere and therefore absorbing at lower temperatures. This effect is in principle accounted for by using a sequential FRS (compare Sect. 3.2). Nevertheless, besides the Puḳīte terms, in the following also the impact of a second NO_2 cross-section at a lower temperature, namely 220 K, on the HCHO fit results is investigated. Lastly, also the influence of an NO_2 Ring spectrum is tested.

6.2.2.1 *Synthetic spectra*

To obtain robust results, in total 36 elevation sequences of synthetic spectra with noise for different measurement geometries, namely the combinations of SZAs of 20°, 45° and 60° and RAAs of 30°, 90° and 180°, for the cases of low and high NO_2 and HCHO abundances (compare Table 6.1) were analysed using the standard fit settings summarised in Table 6.3. However, in addition different combinations of Puḳīte terms and additional NO_2 cross-sections (all 16 possible combinations of them) were included in the fit. First tests showed that the choice of the "noisy" FRS matters since systematic patterns are introduced by the selection of a specific spectrum. Therefore, the analyses were performed using the 90° spectrum without noise of the corresponding scenario as FRS.

As described above, each elevation sequence of a scenario is available 100 times with different noise spectra having the same RMS added. Hence, 100 fit results for each combination of NO_2 cross-sections and for each elevation of

a given scenario were obtained. From these results first the mean dSCDs and mean fit RMS were calculated for each elevation. Next, absolute and relative deviations from the true reference values were calculated (compare Sect. 6.1.2) using Eq. 6.1 and 6.2, respectively. Then, the mean retrieved dSCDs, deviations and mean fit RMS are plotted as functions of elevation for each sequence. Figure 6.4 shows a representative example for the case of low NO_2 abundance, high HCHO abundance, an SZA of 45° and an RAA of 180°. The different panels depict the previously explained quantities, whereby the results for different combinations of Puķīte terms and other NO_2 cross-sections are colour-coded as indicated in the legend.

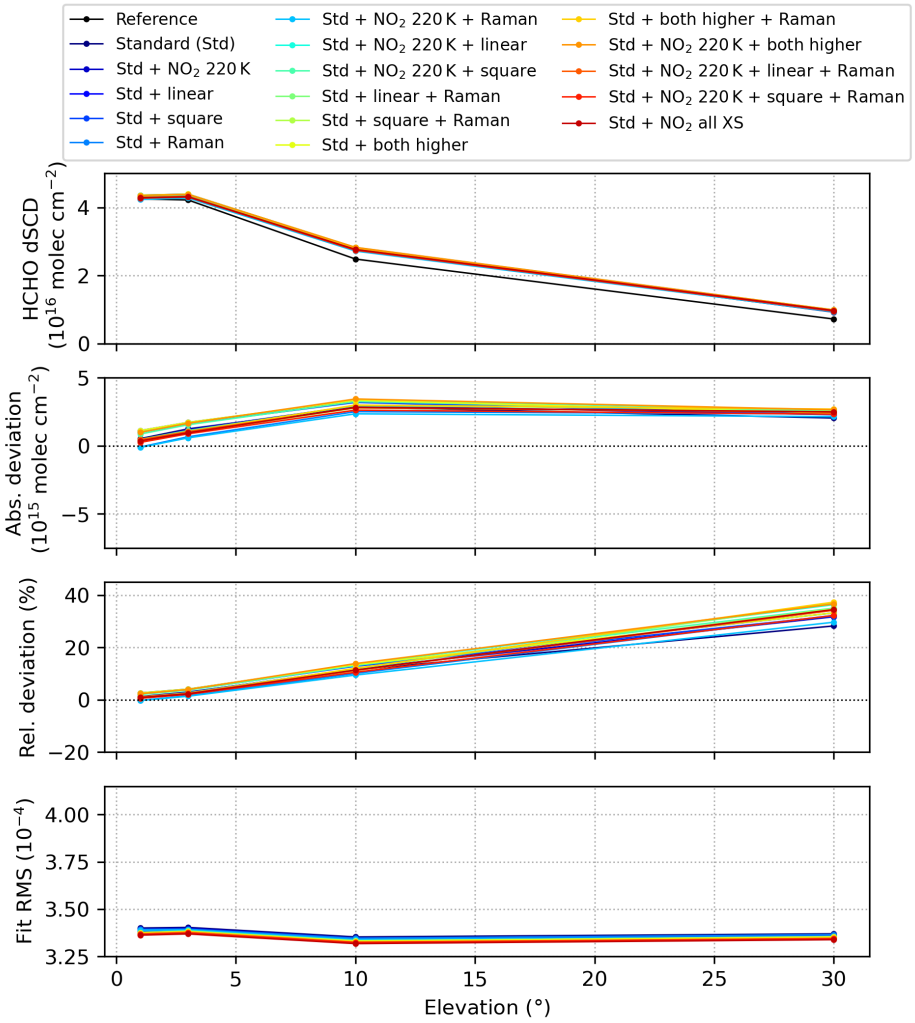


Figure 6.4: Mean retrieved HCHO dSCDs, absolute and relative deviations from the true reference values and fit RMS as functions of elevation for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. Results for different combinations of Puķīte terms and other NO_2 cross-sections are indicated by different colours.

All combinations of NO_2 cross-sections lead to an overestimation of the HCHO dSCDs. This is also valid for all other geometries which are not shown here and

the corresponding low HCHO scenarios. While the overestimations are small at low elevations, larger overestimations are found for elevation angles of 10° and 30° for this and almost all other geometries. Further, the differences between the results obtained with different combinations of NO_2 cross-sections are small which is understandable since the here-presented scenario contains only small NO_2 absorptions. Also the fit RMS show no significant differences amongst the different retrieval scenarios (approximately 1 %). Also this finding is valid for the other (not presented) geometries. Moreover, the differences between the results obtained with different combinations of NO_2 cross-sections are smaller than the deviations of the results from the true reference values, especially for elevations of 10° and 30°. Consequently, hardly any conclusion on which combination is the best can be drawn from the low NO_2 scenario.

Figure 6.5 shows the previously explained quantities for the same scenario but for the case of high NO_2 abundance.

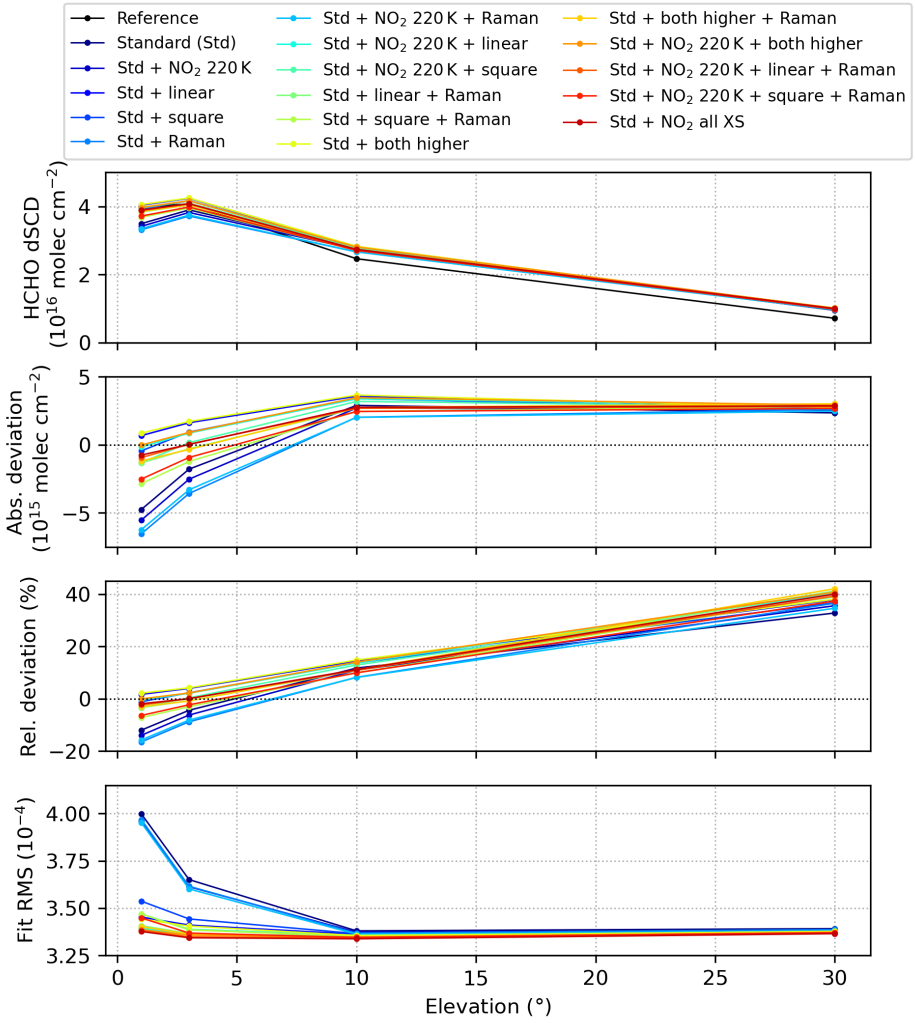


Figure 6.5: Same as Fig. 6.4 but for the case of high NO_2 abundance.

Here, for the lowest elevations (1°, 3° and 10°) the spread between the results obtained with different combinations is large indicating strong impacts of addi-

tional NO_2 cross-sections, while the influences of the different combinations at 30° are small. Both findings also hold for the not presented geometries and the low HCHO scenario. The standard settings lead to strong underestimations of the HCHO dSCDs for the two lowest elevations (compare darkest blue line in the second panel of Fig. 6.5). This is in principle valid for all geometries with one exception, namely for an SZA of 60° and an RAA of 30° (not shown), for which the deviations are rather small for the standard settings at these elevations. Many combinations of NO_2 cross-sections lead to an improved agreement between retrievals and simulations compared to the standard settings, i.e. lower absolute deviations at low elevations. Nevertheless, a few combinations, e.g. only adding the NO_2 Raman cross-section, result in worse agreements at low elevations. For all fit scenarios, overestimations occur at 10° and 30° elevation. Similar general patterns, i.e. elevation dependence of absolute deviations, are found for the other not presented geometries and the low HCHO scenario. It should be noted that the observed systematic elevation dependence of the absolute deviations, in particular the change of the sign of the deviations which is found for some combinations, directly affects the retrieved profile shape of HCHO. For elevations below 10°, significant improvements of the fit RMS of up to 15 % can be achieved by many combinations of additional NO_2 cross-sections. This is again also valid for the not presented geometries and scenarios.

From this kind of plots, it is difficult to assess which of the 16 combinations of NO_2 cross-sections performs overall best. As outlined in Sect. 6.1.2, it is desirable to have one quantity describing the deviations for a complete sequence for each tested combination of cross-sections. Therefore, the sums (Σ_{abs}) and the RMS (RMS_{abs}) of the absolute deviations within each sequence for each combination of NO_2 cross-sections were calculated using Eq. 6.3 and 6.4. As an example, Fig. B.1 in the appendix presents the resulting RMS_{abs} and Σ_{abs} for all 16 combinations of NO_2 cross-sections for the sequence shown in Fig. 6.5. The four combinations not including any Puķīte term for NO_2 clearly yield the largest RMS_{abs} reflecting the large deviations from the true reference dSCDs obtained with these combinations for this sequence. In order to find the combination which yields the overall smallest deviations from the true reference values, mean Σ_{abs} and mean RMS_{abs} as well as the corresponding standard deviations were calculated for each combination of NO_2 cross-sections over all sequences, i.e. geometries. The combination with lowest mean RMS_{abs} together with the lowest mean Σ_{abs} should indicate the best scenario. As all findings so far are independent of the HCHO scenario, both are treated together in the following while the means were calculated separately for the cases of low and high NO_2 abundances.

Figure 6.6 presents the mean RMS_{abs} (upper panel) and Σ_{abs} (lower panel) for all 16 combinations of NO_2 cross-sections for all scenarios with low NO_2 abundance. The error bars indicate the standard deviations of both quantities over all these scenarios.

As described above, the differences between the results obtained with different combinations of NO_2 cross-sections are rather small for the low NO_2 scenario. In particular, the difference between the minimum and maximum mean RMS_{abs} is smaller than the means themselves. Especially, within their respective standard

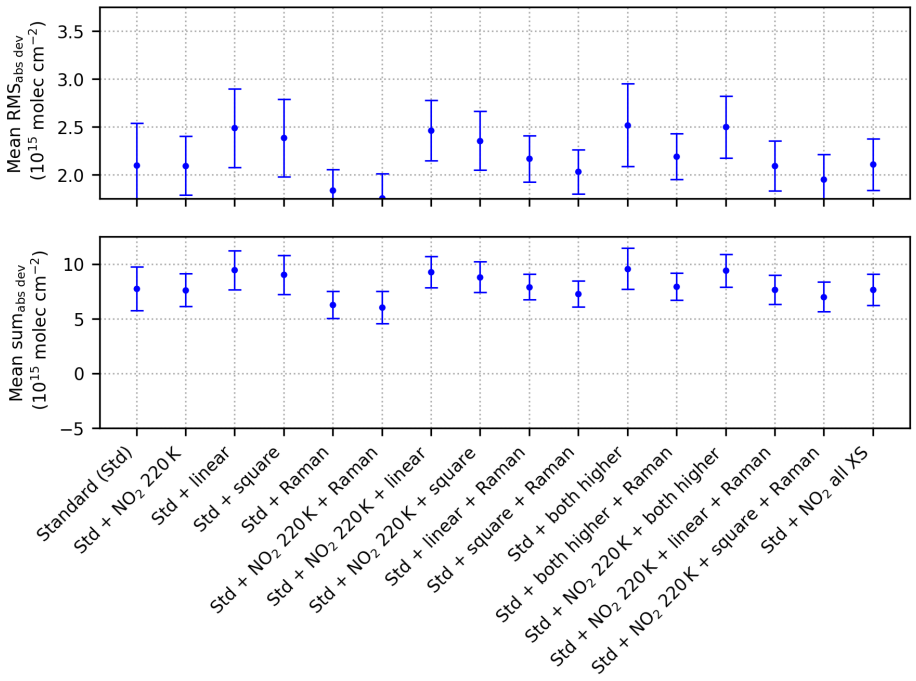


Figure 6.6: Mean RMS (upper panel) and mean sums (lower panel) of the absolute deviations from the true reference dSCDs for all 16 combinations of NO_2 cross-sections for all scenarios with low NO_2 abundance. Error bars indicate the corresponding standard deviations over all these scenarios.

deviations all scenarios yield similar overall deviations from the reference dSCDs. This also holds for the mean sums of the absolute deviations. Both findings stress that hardly any conclusion can be drawn from the low NO_2 scenario. Hence, the focus is put on the high NO_2 scenario in the following.

Similar to Fig. 6.6, Fig. 6.7 depicts the mean RMS_{abs} (upper panel) and Σ_{abs} (lower panel) for the different combinations of cross-sections but for the high NO_2 scenarios.

If the focus is solely put on the lowest mean RMS_{abs}, a group of roughly 10 combinations of NO_2 cross-sections can be considered as "best combinations" since they show similar deviations from the reference dSCD within their respective uncertainties which are also comparable amongst each other. The absolutely lowest mean RMS_{abs} is found for the fit scenario which includes additionally the NO_2 cross-section at 220 K, the linear Puḳīte term and the NO_2 Raman cross-section. However, all these fit scenarios are cases where NO_2 cross-sections are combined which individually lead to notable under- and overestimations at elevations of 1° and 3°, while at 10° and 30° elevation the individual cross-sections yield similar results (compare Fig. 6.5). In combination, these two opposite deviations compensate each other leading to comparably small overall deviations.

For instance, for the combination yielding the absolutely lowest mean RMS_{abs}, the standard settings (Std) and the additional inclusion of the NO_2 cross-section at 220 K only as well as the additional usage of the Raman cross-section only individually lead to strong underestimations of the HCHO dSCDs at 1° and 3°,

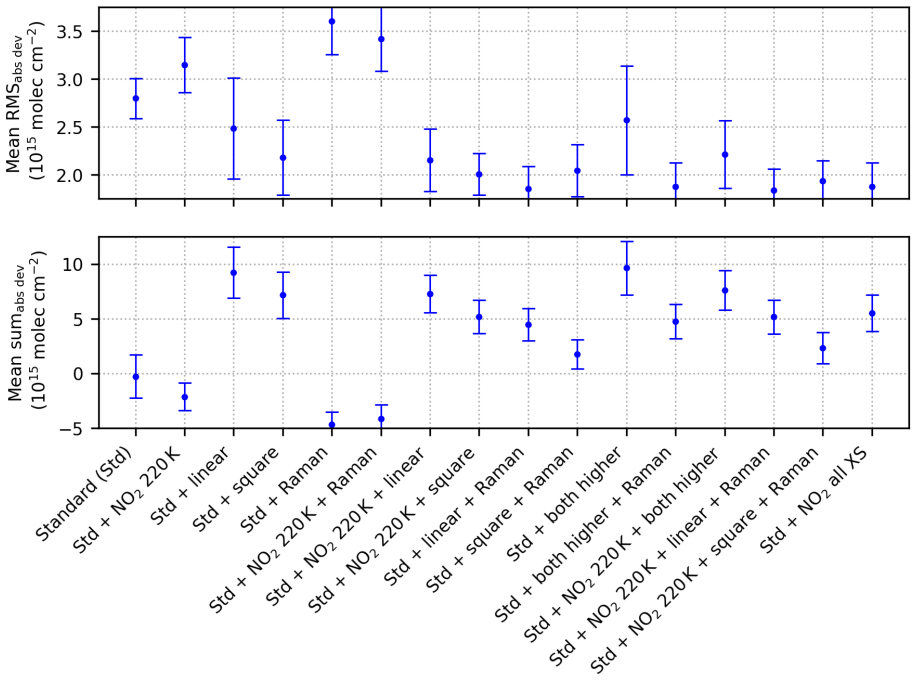


Figure 6.7: Same as Fig. 6.6 but for the case of high NO_2 abundance.

while the scenario with only adding the linear Puķīte term yields overestimations at these elevations (compare Fig. 6.5). If all of these cross-sections are included together, smaller absolute deviations are found at the lowest two elevations leading to the absolutely lowest mean RMS_{abs}. Similar considerations are valid for the other combinations of cross-sections showing small deviations. The bad overall performances of the fit scenarios including only one additional cross-section become also evident in Fig. 6.7 which shows large mean RMS_{abs} for these scenarios. For the linear Puķīte term, also a high mean sum of the absolute deviations is found reflecting the overestimations at the lowest two elevations, while this quantity is negative for the other cross-sections indicating strong underestimations at these two elevations. Note that a small mean sum of the absolute deviations together with a large mean RMS_{abs} only indicates strong underestimations within a sequence but not at which elevations these occur. The previous conclusions can only be drawn when also information of plots like Fig. 6.5 are considered.

Altogether, no clear decision on the best combination of NO_2 cross-sections can be made in view of the compensations of under- and overestimations obtained with the individual cross-section scenarios. Further, some of these combinations perform better for the high NO_2 scenarios than for the low NO_2 cases which is not really reasonable. For the low NO_2 scenarios, the following conclusions hold:

1. The different combinations of additional NO_2 cross-sections do almost not improve the fit results and their agreement with the true reference dSCDs compared to the standard settings.

2. The differences between the results obtained with different combinations of NO_2 cross-sections are small and similar for almost all scenarios.

3. Even for the low NO_2 scenarios for which the NO_2 absorptions should not largely affect the DOAS fit, the RMS_{abs}-values are not 0 but a systematic bias exists (compare studies in Sect. 6.4).

As a consequence of these three findings, apparently perfect agreement between retrievals and simulations cannot be achieved. Therefore, under the given circumstances, it can be assumed that there is a systematic deviation from the reference dSCDs which so far cannot be fully explained (for details refer to Sect. 6.4) and which is almost independent of NO_2 as the variations between the results obtained with different combinations of NO_2 cross-sections are small. Hence, no better agreement can be expected for the high NO_2 scenarios as well and the absolute deviations obtained for the low NO_2 scenarios can be used as reference for the deviations which are expected for the high NO_2 scenarios. These considerations lead to a second approach for finding the best combination of NO_2 cross-sections included in the HCHO fit. For this approach, the deviations obtained for the low NO_2 scenario of a given geometry are used as reference instead of the true reference dSCDs. It should be stressed that this second approach is only used under the given circumstances that even for the low NO_2 scenario deviations from the true reference dSCDs remain which are not fully understood (compare Sect. 6.4).

For this second approach, the absolute deviations obtained using Eq. 6.1 for the high NO_2 scenario of a given sequence (defined as sun geometry and HCHO scenario) including different combinations of NO_2 cross-sections in the DOAS fit are compared to the corresponding absolute deviations obtained for the low NO_2 scenario including all possible NO_2 cross-sections. The latter combination was chosen since it should in principle reproduce the simulations best as it accounts for all possible effects that can be accounted for by the different combinations of cross-sections. For this comparison, the differences between the deviations obtained with a specific set of NO_2 cross-sections for the case of high NO_2 ($\Delta dSCD_{\text{abs, high } NO_2}$) and the deviations for the corresponding low NO_2 scenario including all possible NO_2 cross-sections in the fit ($\Delta dSCD_{\text{abs, low } NO_2 \text{all } NO_2 \text{ XS}}$) are calculated for each elevation angle (α) using

$$CQ(\alpha) = \Delta dSCD_{\text{abs, high } NO_2}(\alpha) - \Delta dSCD_{\text{abs, low } NO_2 \text{all } NO_2 \text{ XS}}(\alpha). \tag{6.6}$$

These differences are referred to as "comparison quantity (CQ)" in the following.

As an example, the absolute deviations from the true reference dSCDs (upper panel) and the corresponding CQ-values (lower panel) are shown as functions of elevation in Fig. 6.8 for the scenario with high NO_2 and HCHO abundances, an SZA of 45° and an RAA of 180°. The results obtained with different combinations of NO_2 cross-sections are indicated by different colours, while the black line in the upper panel represents the absolute deviations of the corresponding low NO_2 scenario obtained with all possible NO_2 cross-sections included in the fit.

Several fit scenarios using different combinations of NO_2 cross-sections reproduce the absolute deviations of the corresponding low NO_2 scenario quite

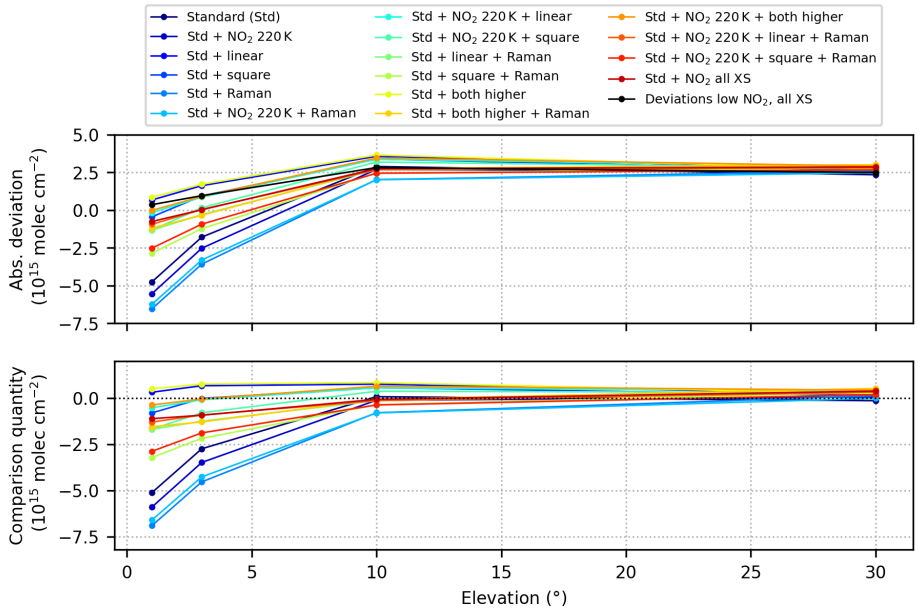


Figure 6.8: Absolute deviations from the true reference dSCDs (upper panel) and corresponding *CQ*-values (lower panel) as functions of elevation for the case of high NO_2 and HCHO, an SZA of 45° and an RAA of 180°. Results obtained with different sets of NO_2 cross-sections are indicated by different colours. The black line in the upper panel represents the absolute deviations of the corresponding low NO_2 scenario obtained including all possible NO_2 cross-sections.

well (compare coloured and black lines in the upper panel) and thus remove systematic influences caused by NO_2 absorptions from the fit. This becomes even more evident when the *CQ*-values obtained using Eq. 6.6 in the lower panel are considered. Here, the large spread of the different coloured lines at the lowest two elevations indicates that some combinations of NO_2 cross-sections can and some cannot reproduce the absolute deviations of the low NO_2 scenario. Similar general patterns are found for the other not presented geometries and only small differences in specific details are observed. In particular, the findings are independent of the HCHO scenario. Consequently, the two HCHO scenarios are considered together in the following.

Like for the absolute deviations, it is desirable to have one quantity describing *CQ* obtained with a specific set of NO_2 cross-sections for a complete sequence. Hence, similar to Eq. 6.3 and 6.4, the sum (Σ_{CQ}) and the root mean square (RMS_{CQ}) of *CQ* within a given sequence for a certain set of NO_2 cross-sections can be calculated. The latter is defined as

$$RMS_{CQ} = \sqrt{\frac{1}{N}\sum_{\alpha} CQ\,(\alpha)^2} \tag{6.7}$$

and basically describes the RMS_{CQ} along a specific coloured line in the lower panel of Fig. 6.8.

This was done for each geometry, both HCHO scenarios and each combination of NO_2 cross-sections. Figure B.2 in the appendix depicts the resulting RMS_{CQ}

and Σ_{CQ} for all 16 combinations of NO_2 cross-sections for the sequence shown in Fig. 6.8. Similar to the considerations regarding the absolute deviations above, the four combinations not including any Puķīte term for NO_2 lead to the largest RMS_{CQ} reflecting the large differences between the absolute deviations obtained for the high and low NO_2 scenarios with these combinations for this sequence (compare lower panel in Fig. 6.8). To find the overall best combination, the mean Σ_{CQ} and the mean RMS_{CQ} were calculated over all scenarios for each set of NO_2 cross-sections in the same way as for the absolute deviations above. The fit scenario yielding the lowest mean RMS_{CQ} (possibly in combination with a small mean Σ_{CQ}) performs best and indicates the combination of NO_2 cross-sections for which the absolute deviations of the low (including all NO_2 cross-sections) and high NO_2 scenarios deviate least. The resulting mean RMS_{CQ} (upper panel) and mean Σ_{CQ} (lower panel) obtained for the 16 different combinations of NO_2 cross-sections are presented in Fig. 6.9, whereby the standard deviations of both quantities over all scenarios are indicated by error bars.

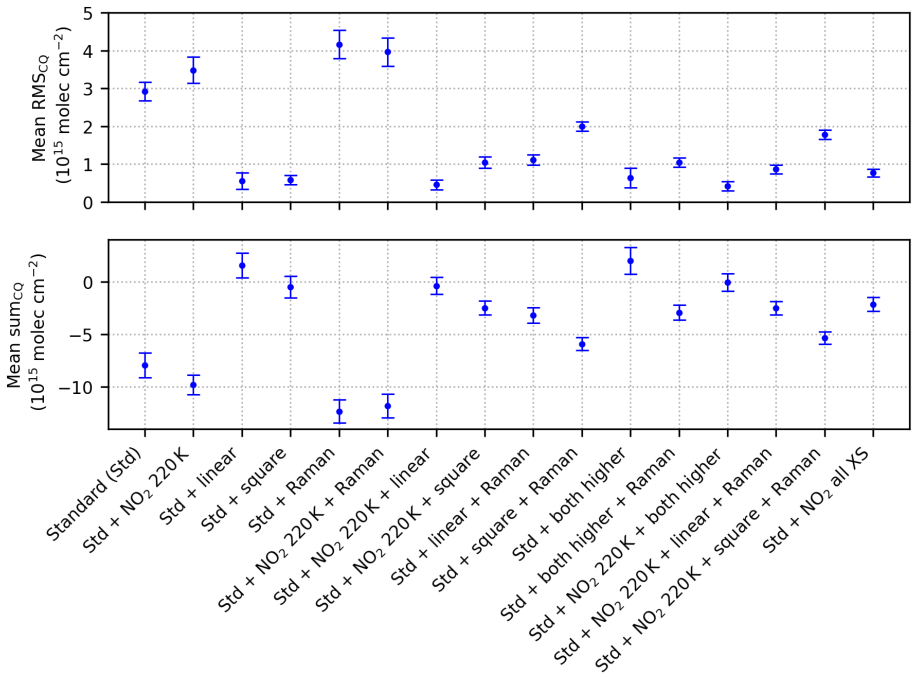


Figure 6.9: Mean RMS_{CQ} (upper panel) and mean Σ_{CQ} (lower panel) obtained for all 16 combinations of NO_2 cross-sections. Error bars indicate the corresponding standard deviations over all scenarios.

Obviously, four combinations, namely the standard fit settings, the fit scenarios solely adding the NO_2 cross-section at 220 K or the NO_2 Raman cross-section as well as their combination, show the largest mean RMS_{CQ}-values indicating large discrepancies to the corresponding low NO_2 scenarios. The negative mean Σ_{CQ}-values suggest that these discrepancies are mainly caused by largely negative CQ at low elevations (compare Fig. 6.8). These four combinations of NO_2 cross-sections performed also bad in terms of absolute deviations from the true reference dSCDs (compare Fig. 6.7). The combinations "Std + square + Raman" and

"Std + NO_2 at 220 K + square + Raman" are amongst the best sets of NO_2 cross-sections regarding the deviations from the reference dSCDs (compare Fig. 6.7) but lead to comparably large RMS_{CQ}-values. The remaining 10 combinations of NO_2 cross-sections yield RMS_{CQ}-values which are rather close to each other. When taking their respective error bars (standard deviations) into account, a detailed look shows that about five sets of NO_2 cross-sections yield very similar and rather small RMS_{CQ}-values. Interestingly, the best combination in terms of deviations from the reference dSCDs (additionally including the NO_2 cross-section at 220 K, the linear Puķīte term and the NO_2 Raman cross-section) is not amongst them.

Despite the fact that the combination of NO_2 cross-sections yielding the lowest mean RMS_{CQ} can in principle be identified, namely the fit scenario which additionally includes the NO_2 cross-section at 220 K and both Puķīte terms, like for the absolute deviations a point is reached at which no clear "winning" combination can be easily selected. Another aspect is the question how many cross-sections should at maximum be included in the DOAS fit to not run into over-fitting. A different and probably more reasonable consideration is the combination of both optimisation approaches, namely the minimisation of the absolute deviations and the consistency between the deviations obtained for the low and high NO_2 scenarios of a given geometry (and HCHO scenario). Note that the initial aim of the study was of course to reproduce the true reference dSCD as good as possible. However, as previously discussed, there are reasons why both concepts alone do not lead to clear conclusions. Thus, a combination of both approaches is chosen and two "decision trees" are used in parallel since both approaches are equally justified:

1. First check for the lowest mean RMS_{abs} (deviations from the true reference dSCDs) plus a certain tolerance interval and perform a preselection of combinations. Then, check for the minimum mean RMS_{CQ} (consistency between low and high NO_2 scenarios) plus a certain tolerance interval for the preselected combinations of NO_2 cross-sections.

2. First optimise the consistency between low and high NO_2 scenarios (RMS_{CQ}) and perform a preselection of combinations. Then, check for the lowest mean RMS_{abs} (deviations from the reference dSCDs) of the preselected sets of NO_2 cross-sections. Again applying certain tolerance intervals for both steps.

For "decision tree 1", first the mean RMS_{abs} describing the absolute deviations from the true reference dSCDs (as explained above) for the different combinations of NO_2 cross-sections are compared amongst each other and the one with the absolutely lowest mean as well as the ones within a certain tolerance interval are preselected. For that, the upper panel of Fig. 6.10 shows the mean RMS_{abs} for all 16 combinations of NO_2 cross-sections calculated over all scenarios with high NO_2 abundance. Again error bars indicate the corresponding standard deviations. In addition, different tolerance intervals for the lowest mean RMS_{abs} (fit scenario additionally including the NO_2 cross-section at 220 K, the linear Puķīte term and the NO_2 Raman cross-section, black dashed line) are indicated by coloured dashed lines.

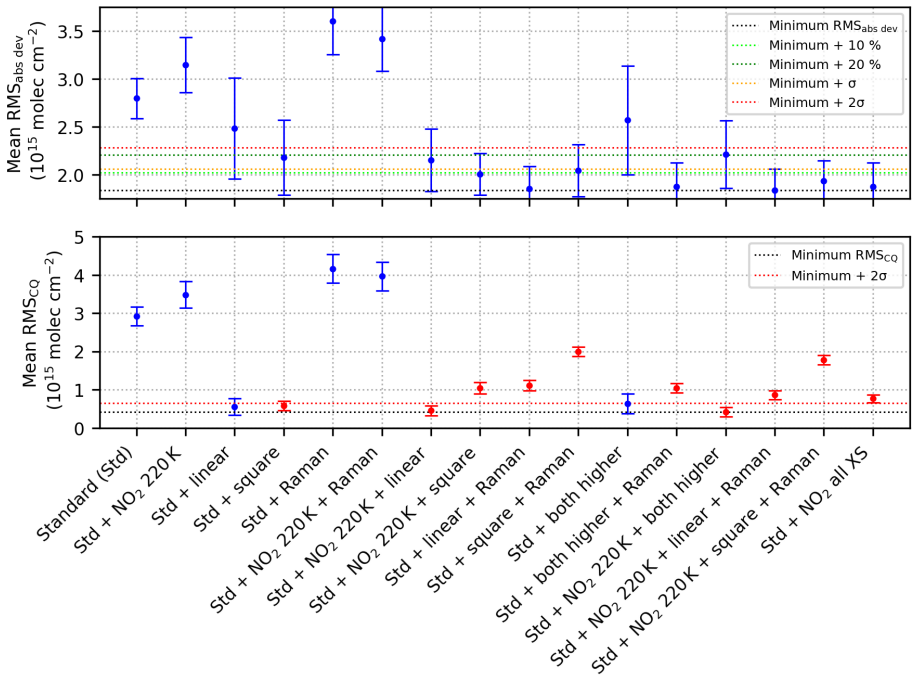


Figure 6.10: Mean RMS_{abs} (upper panel) and mean RMS_{CQ} (lower panel) obtained for all 16 combinations of NO_2 cross-sections calculated over all scenarios with high NO_2 abundance. Error bars indicate the corresponding standard deviations over all these scenarios. Different tolerance intervals for the lowest mean RMS_{abs} (black dashed line) are indicated by coloured dashed lines as described in the legend. In the lower panel, the 10 combinations of NO_2 cross-sections which remain after the preselection as described in the text are marked in red.

In principle, different choices are conceivable for the tolerance interval. For instance, one could simply select the best three or five combinations in each decision step which, however, would be rather arbitrary. Further, also relative measures such as 10 % or 20 % of the minimum mean RMS_{abs} could be reasonable choices (greenish dashed lines in the upper panel of Fig. 6.10). However, even that procedure would still be somehow arbitrary. Hence, the standard deviation over all scenarios (geometries) with high NO_2 abundance of the combination yielding the lowest mean RMS_{abs} (σ in the upper panel indicated by reddish dashed lines) is used in the following. This choice is well justified as it is a statistical property of the data themselves. Selecting only sets of NO_2 cross-sections which yield mean RMS_{abs}-values within 2σ of the minimum mean RMS_{abs} (red dashed line) leads to 10 remaining combinations. The latter are marked in red in the lower panel of Fig. 6.10 which depicts the mean RMS_{CQ} for all 16 combinations of NO_2 cross-sections for all scenarios with high NO_2 abundance. The lowest mean RMS_{CQ} of these preselected combinations is found for the fit setting which additional includes the NO_2 cross-section at 220 K and both Puḷīte terms. Actually, this scenario yields also the overall lowest mean RMS_{CQ}. The red dashed line in the lower panel indicates the 2σ tolerance interval of this set of NO_2 cross-sections. Note that this σ is different from the one in the upper panel since the

combinations yielding the lowest mean RMS of both quantities differ. Again selecting only combinations of NO_2 cross-sections (of the preselection according to the first step) within this 2σ tolerance leads to three remaining combinations. These are the fit scenario which includes additionally only the square Puḳīte term, the combination including the NO_2 cross-section at 220 K and the linear Puḳīte term and the scenario additionally including the NO_2 cross-section at 220 K and both Puḳīte terms. Interestingly, the combination of NO_2 cross-sections yielding the lowest mean RMS_{abs} is not included in this group of "winners" of "decision tree 1" indicating that for this fit scenario no good consistency between the low and high NO_2 scenarios is found. Lastly, it should be mentioned that the choice of the tolerance interval is crucial. In particular, the three "winners" of "decision tree 1" would not be options if the tolerance interval was 10 % or 1σ in the preselection (compare upper panel of Fig. 6.10). Consequently, the three "winner" combinations are not the best in terms of agreement with the true reference dSCDs but also not the worst.

As described above, the order of the optimisations is reversed for "decision tree 2". Here, first the consistency of the absolute deviations from the true reference dSCDs between the high and low NO_2 scenario is assessed. To do so, the mean RMS_{CQ} obtained for all 16 combinations of NO_2 cross-sections are presented in the upper panel of Fig. 6.11, together with their corresponding standard deviations (error bars) and different tolerance intervals of the minimum mean RMS_{CQ} (coloured dashed lines). The latter is obtained for the combination of NO_2 cross-sections which additionally includes the NO_2 cross-section at 220 K and both Puḳīte terms and is marked by a black dashed line. Similar to "decision tree 1", the standard deviation of this fit scenario is used as selection criterion and combinations yielding mean RMS_{CQ}-values within 2σ (red dashed line in the upper panel of Fig. 6.11) of the minimum mean are preselected for the second optimisation step.

These five remaining combinations are marked in red in the lower panel of Fig. 6.11 which shows the mean RMS_{abs} and the corresponding standard deviations (error bars) for all 16 combinations of NO_2 cross-sections for all scenarios with high NO_2 abundance. Again a red dashed line indicates the 2σ tolerance interval of the NO_2 cross-section combination (of the preselected combinations) yielding overall the lowest deviations from the reference dSCDs, namely the one adding the NO_2 cross-section at 220 K and the linear Puḳīte term. Like for "decision tree 1" this is a different combination as in the first optimisation step. Nevertheless, contrary to the first "decision tree", all five preselected sets of NO_2 cross-sections are within in the 2σ (compare lower panel of Fig. 6.11) of the minimum mean RMS_{abs}. Again the choice of the tolerance interval is crucial, especially for the first optimisation step. Since σ is comparably small for the minimum mean RMS_{CQ} (compare reddish dashed lines in the upper panel), a stricter selection criterion for the first optimisation step would not leave many combinations left.

A comparison of the "winner" groups of both "decision trees" shows that all three "winners" of "decision tree 1" are also within the "winner" group of "decision tree 2". Thus, these three combinations can be considered as the best

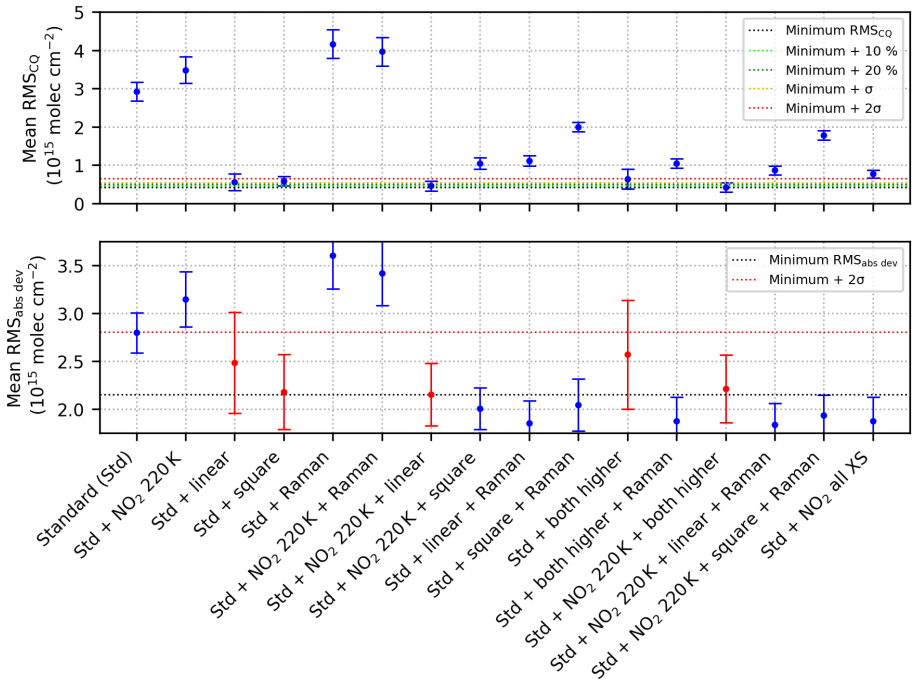


Figure 6.11: Mean RMS_{CQ} (upper panel) and mean RMS_{abs} (lower panel) obtained for all 16 combinations of NO_2 cross-sections calculated over all scenarios with high NO_2 abundance. Error bars indicate the corresponding standard deviations over all these scenarios. Different tolerance intervals for the lowest mean RMS_{CQ} (black dashed line) are indicated by coloured dashed lines as described in the legend. In the lower panel, the five combinations of NO_2 cross-sections which remain after the preselection as described in the text are marked in red.

suited, namely the fit scenario which includes additionally only the square Puḳīte term, the combination including the NO_2 cross-section at 220 K and the linear Puḳīte term and the scenario additionally including the NO_2 cross-section at 220 K and both Puḳīte terms. All of these three combinations perform very similar in terms of absolute deviations from the true reference dSCDs (compare figures above) with the combination additionally including the NO_2 cross-section at 220 K and the linear Puḳīte term being slightly better. However, none of them is amongst the overall best combinations in this regard. Yet they lead to clear improvements of the agreement between retrievals and simulations for the high NO_2 scenarios compared to the standard scenario (compare Fig. 6.7). Also in terms of consistency between high and low NO_2 scenario all three combinations perform equally well, whereby the scenario additionally including the NO_2 cross-section at 220 K and both Puḳīte terms performs overall best. Since the combination that includes additionally only the square Puḳīte term is not the best in either of the optimisation steps it is not considered in the following. Putting slightly more weight on the consistency between the deviations from the true reference dSCDs obtained for the high and low NO_2 scenarios, the final choice is the fit scenario which additionally includes the NO_2 cross-section at 220 K and

both Puķīte terms and thus this combination is used for the further optimisations of the DOAS fit settings for HCHO.

6.2.2.2 *Measured spectra*

Although the final decision on which set of NO_2 cross-sections should be included in the DOAS fit of HCHO is mainly based on the results obtained for the synthetic spectra presented in the previous section, the performances of the different combinations of NO_2 cross-sections under real measurement conditions are investigated as well in this section. For that, measured spectra from three measurement sites as described in Sect. 6.1.1 were used. Naturally, for these spectra no comparison to reference values is possible. Nevertheless, as described above, the effects of different combinations of NO_2 cross-sections on the fit RMS and the retrieved HCHO dSCDs can be examined. In particular, the performance of the finally selected combination based on the results for the synthetic spectra, namely the one that additionally includes the NO_2 cross-section at 220 K and both Puķīte terms, is assessed.

The uppermost panel of Fig. 6.12 shows time series of HCHO dSCDs retrieved including different sets of NO_2 cross-sections (represented by coloured lines) in the DOAS fit for spectra recorded at Bayfordbury on 29 March 2019 at 3° elevation. In addition, the resulting VCDs of NO_2 and HCHO are shown as black and green dots, respectively. On the used y-scale, the results obtained with the different combinations of NO_2 cross-sections are hardly distinguishable and only in the late afternoon around 5 p.m. a clearer separation of the different fit scenarios is visible. Therefore, (absolute) differences of the HCHO dSCDs obtained with different sets of NO_2 cross-sections with respect to the ones obtained using the standard settings (only one NO_2 cross-section) were calculated and presented in the second panel of Fig. 6.12.

The HCHO dSCDs retrieved including different sets of NO_2 cross-sections in the DOAS fit differ by roughly 2×10^{15} to 3×10^{15} molec cm^{-2} (peak-to-peak) during large parts of the day. For the synthetic spectra, the spread between the results obtained with different fit scenarios at 3° elevation was ca. 4×10^{15} to 5×10^{15} molec cm^{-2} for the high NO_2 scenarios depending on the specific sun geometry which is approximately two times larger. This difference between the results for synthetic and real measured spectra despite similar NO_2 VCDs is probably caused by differences between the assumed (simulations) and real profile shapes (i.e. vertical distributions) of NO_2. In particular, the entire tropospheric VCD is concentrated in the lowest 500 m for the synthetic spectra, while the results of the Bayfordbury measurements (compare Sect. 10.5) show that the majority of the NO_2 is distributed up to a height of ca. 1 km. Consequently, the NO_2 absorptions and therefore also the dSCDs at low elevations are larger for the synthetic spectra (ca. 2×10^{17} molec cm^{-2}) than for the real spectra (ca. 1.5×10^{17} molec cm^{-2}). Like for the synthetic data, the differences between the dSCDs retrieved with different sets of NO_2 cross-sections included in the DOAS fit become smaller with increasing elevation as also the NO_2 absorptions get weaker (not shown). For the synthetic spectra in the case of high NO_2, the standard fit settings underestimated the true reference dSCDs of HCHO at 3° elevation

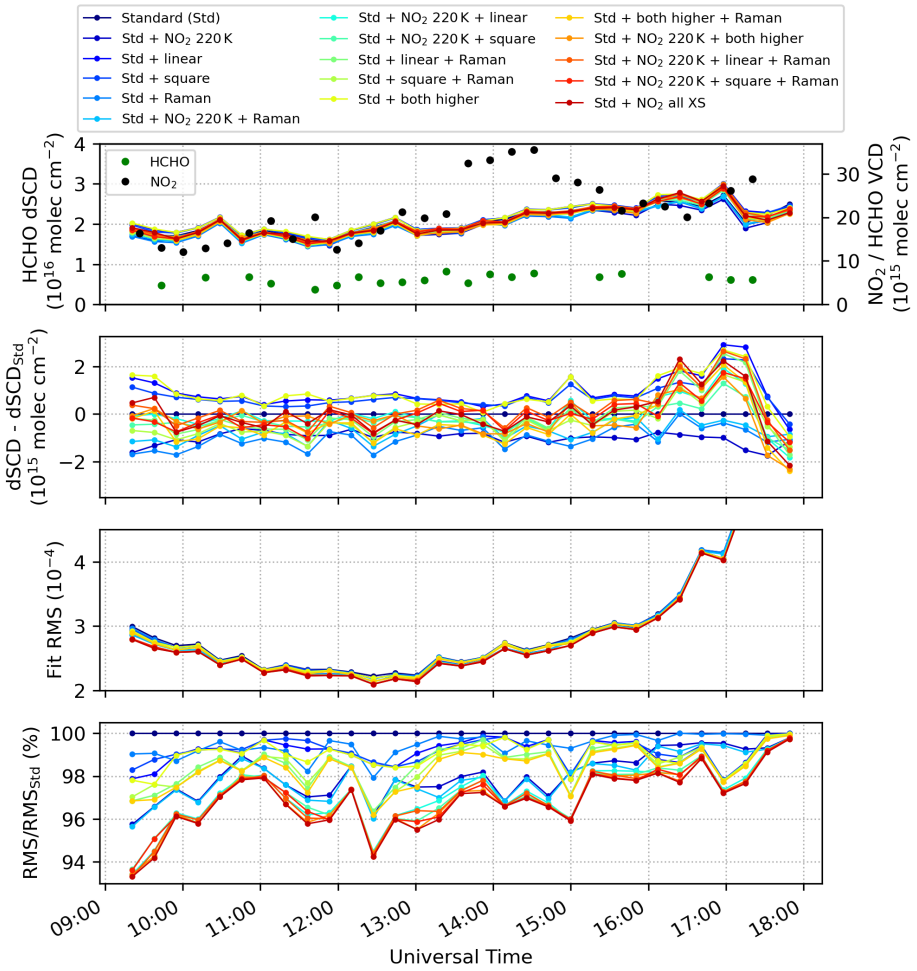


Figure 6.12: Time series of HCHO dSCDs retrieved including different sets of NO_2 cross-sections in the DOAS fit (uppermost panel), corresponding differences with respect to the standard settings (second panel), fit RMS-values (third panel) and ratios of these fit RMS-values with respect to the ones obtained for the standard settings (lowermost panel) for spectra recorded at Bayfordbury on 29 March 2019 at 3° elevation. Results obtained with different combinations of NO_2 cross-sections included in the DOAS fit are indicated by different colours. The uppermost panel additionally shows the resulting NO_2 (black dots) and HCHO (green dots) VCDs.

for almost all sun geometries and both HCHO scenarios. Further, the dSCDs obtained with the standard fit settings were amongst the lowest compared to the other combinations of NO_2 cross-sections (compare Fig. 6.5). Contrary to that, the standard settings lead to dSCDs which are roughly in the middle of the dSCDs retrieved with the other tested cross-section sets (compare darkest blue line in the second panel of Fig. 6.12). Towards evening with increasing SZA the peak-to-peak differences between the dSCDs increase reaching up to 4.5×10^{15} molec cm^{-2} which is comparable to the results for the synthetic spectra. Combinations leading to clearly lower or higher HCHO dSCDs for the synthetic spectra, e.g. the settings solely adding the NO_2 Raman cross-section or adding

both Puķīte terms, respectively, show the same behaviour also for the real spectra. This indicates combinations being extreme cases which should therefore not be chosen (compare Sect. 6.2.2.1). However, the results of most combinations of NO_2 cross-sections are rather similar and scatter around the results obtained with the standard settings, whereby mostly slightly smaller dSCDs are retrieved during large parts of the day with the additional NO_2 cross-sections. This finding differs from the results for the synthetic spectra for which most combinations of NO_2 cross-sections yielded larger dSCDs for most geometries (compare Fig. 6.5). For the measured spectra, only towards evening most combinations lead to higher dSCDs compared to the ones of the standard settings. At 3° elevation, the "final choice" of NO_2 cross-sections ("Std + NO_2 cross-section at 220 K + both Puķīte terms", light orange lines) yielded about 2×10^{15} to 2.5×10^{15} molec cm^{-2} larger dSCDs than the standard settings for most scenarios of synthetic spectra. Contrary to that, the dSCDs obtained with both settings are comparable for the spectra measured at Bayfordbury, whereby the final choice yields slightly lower dSCDs. Only towards evening, when also the NO_2 VCDs increase, the final choice leads to ca. 1×10^{15} molec cm^{-2} larger HCHO dSCDs. However, the increased NO_2 abundances in the early afternoon do not lead to higher HCHO dSCDs for the final choice. This can again probably be explained by different NO_2 profile shapes causing different sensitivities in the early afternoon and towards afternoon. Consistent to the results for the synthetic spectra, these findings also hold for the other elevation angles at which the effects are absolutely smaller. Since no reference values are known for the real data, the effects of different combinations of NO_2 cross-sections on the dSCDs can only be described. Therefore, next the fit RMS obtained with the different combinations is examined as measure for the quality of the spectral fits.

The corresponding time series are shown in the third panel of Fig. 6.12. Like for the dSCDs, the small differences between the different fit scenarios are hardly visible at the used y-scale. Thus, the ratios of the RMS values obtained with the different fit settings with respect to the ones for the standard settings are depicted in the lowermost panel. Here, a first finding is that all additional NO_2 cross-sections lead to an improvement of the fit RMS which is a mathematical consequence since adding more parameters to the DOAS fit must decrease the residual. However, it is also evident that adding for example only the NO_2 Raman cross-section to the fit improves the RMS less than adding the linear or square Puķīte terms (compare different blue lines in the lowermost panel). Further, some combinations with two additional NO_2 cross-sections, e.g. the cross-section at 220 K and the Raman cross-section, improve the RMS equally as just including one additional cross-section (the one at 220 K). The latter indicates an overall small additional effect of the Raman cross-section. Almost all combinations with three additional cross-sections perform as good as the scenario including all NO_2 cross-sections. Amongst them is the combination that includes the cross-section at 220 K and both Puķīte terms which was the final choice for the synthetic spectra. This combination improves the fit RMS by up to 6 % (ca. 2 to 4 % on average during the day). A last finding is that also the effect on the fit RMS becomes smaller with increasing elevation angle like for the synthetic spectra (not shown).

Figure 6.13 shows similar time series as Fig. 6.12 but for spectra measured during the MSM 68/2 ship cruise on 12 November 2017 at 3° elevation.

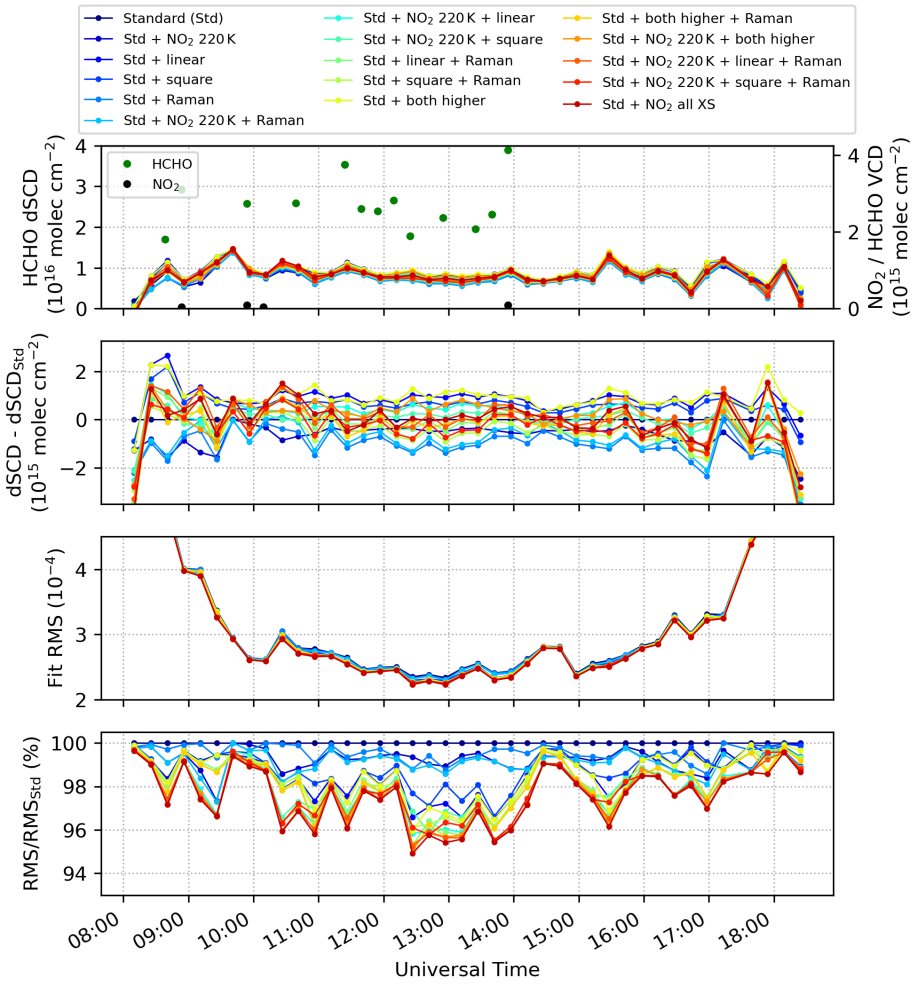


Figure 6.13: Same as Fig. 6.12 but for spectra measured during the MSM 68/2 ship cruise on 12 November 2017 at 3° elevation.

Despite lower HCHO and NO_2 dSCDs, whereby the latter are even close to the detection limit of the used instrument, the findings for the results obtained with different combinations of NO_2 cross-sections are similar to the ones for the Bayfordbury measurements. In principle, most conclusions regarding the differences between the HCHO dSCDs retrieved with different fit settings are also true for the measurements taken during the MSM 68/2 ship cruise. No clear explanation for the observation that the results for the low and high NO_2 scenarios are similar for the real measured spectra can be given so far. For the synthetic spectra in case of low NO_2 abundance, the spread between the HCHO dSCDs obtained with different sets of cross-sections at 3° elevation was approximately 1×10^{15} molec cm^{-2} slightly depending on the specific sun geometry and HCHO scenario which is about two times smaller than for the spectra measured during the MSM 68/2 ship cruise. Similar to Bayfordbury, a possible reason for this finding are differences between the atmospheric profiles assumed in the simu-

lations and the real atmospheric profile shapes of NO_2 and HCHO. As already mentioned, also the other main findings obtained for the Bayfordbury measurements hold for the measurements conducted during the MSM 68/2 ship cruise. However, different to the Bayfordbury measurements for which mostly slightly smaller dSCDs were retrieved during large parts of the day, many combinations yield similar or slightly larger dSCDs here. Also the finally chosen combination of NO_2 cross-sections (orange lines) leads on average to slightly larger HCHO dSCDs compared to the ones retrieved with the standard fit settings.

Looking at the RMS of the fits using different sets of NO_2 cross-sections shows that the general patterns are different than for Bayfordbury. In particular, apparently two groups of lines exist in the lowermost panel of Fig. 6.13 while the different coloured lines were more separated for the Bayfordbury results. One group consists of fit scenarios which do almost not improve the fit RMS and represents combinations which mainly include only one additional NO_2 cross-section (more blueish lines). The second group mainly consists of combinations with two or more additional NO_2 cross-sections leading to significant improvements of the corresponding fit RMS-values particularly around noon. However, for the measurements during the MSM 68/2 ship cruise, the effects on the fit RMS are smaller with only 1 to 2 % improvement on average. The latter is most likely related to the lower NO_2 abundances on the North Atlantic.

The investigations based on the synthetic spectra showed that the inclusion of additional NO_2 cross-sections in the DOAS fit has almost no effect on the retrieved HCHO dSCDs for the low NO_2 scenarios, while clear improvements are achieved under high NO_2 conditions. The latter is also valid for the fit RMS under real measurement conditions at Bayfordbury. However, a further finding for the measurements during the MSM 68/2 ship cruise is that seemingly the variability of the retrieved dSCDs is larger compared to the Bayfordbury measurements (compare Fig. 6.12 and 6.13) for which the dSCDs of both species are larger. Hence, the question remains if additional NO_2 cross-sections should be included in the DOAS retrievals for stations which experience low NO_2 abundances throughout the year, e.g. ATTO. Further, it is important to know to which extent these additional cross-sections possibly increase the observed variability of the HCHO results.

To answer these questions, spectra measured by the ATTO instrument at 80 m for one complete year (2020) were analysed with two different fit scenarios. Here, the ATTO data were chosen since the ATTO site experiences low NO_2 abundances throughout the year while the HCHO abundances are comparably high in all seasons. The first scenario includes the finally chosen combination of NO_2 cross-sections ("Std + NO_2 cross-section at 220 K + both Puḳīte terms") whereas the second scenario includes only one cross-section for NO_2, namely the standard one at 294 K (compare Table 6.3). Note that these fits were performed using the standard settings for HCHO of the final analyses for this work which are summarised in Table 9.1. These settings already include the optimisation of the O_3 cross-sections (compare Sect. 6.2.3). The uppermost panel of Fig. 6.14 shows time series of HCHO dSCDs retrieved with both fit scenarios for spectra at 3° elevation.

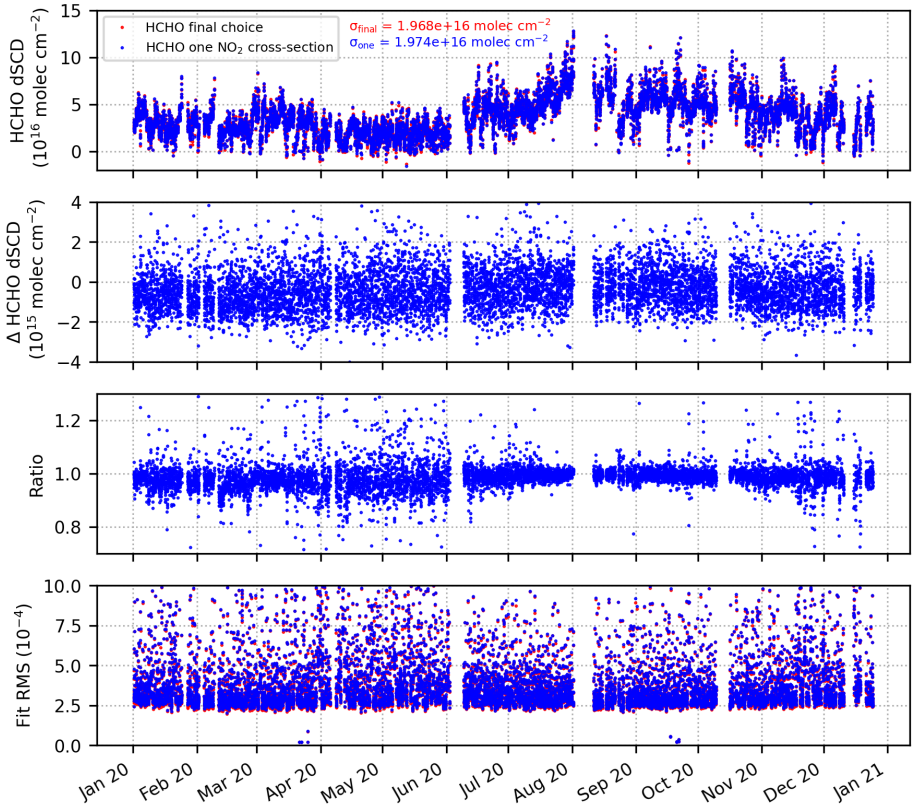


Figure 6.14: Time series of HCHO dSCDs retrieved with the final choice of NO_2 cross-sections (red dots) and with only one NO_2 cross-section (blue dots) included in the DOAS fit (uppermost panel), corresponding differences (second panel), ratios (third panel) and fit RMS-values (lowermost panel) for spectra recorded at ATTO in the year 2020 at 3° elevation. In the uppermost panel additionally the standard deviations of the HCHO dSCDs obtained with both fit scenarios are provided in the respective colours.

Consistent to the findings above, the dSCDs are only marginally affected by the inclusion of the NO_2 cross-section at 220 K and both Puḳīte terms (orange lines in Fig. 6.13) under the low NO_2 conditions at ATTO. The absolute differences depicted in the second panel show that the dSCDs are on average slightly lower for the scenario which includes only one NO_2 cross-section by approximately 0.46×10^{15} molec cm^{-2}. This tendency is also confirmed by the corresponding ratios of the results obtained with both fit scenarios (median ratio: 0.99). Further, a weak annual cycle is found for the differences in the dSCDs. During the parts of the year which show higher HCHO abundances, the differences between the results of both fit scenarios are practically 0 while they are larger for lower HCHO dSCDs. To estimate the effect of the additional NO_2 cross-sections on the variability of the retrieved dSCDs, the standard deviations of both time series were calculated (σ_{final} and σ_{one}). Assuming that systematic variations of the dSCDs are captured identically by both fit scenarios, changes in the standard deviations can be solely attributed to changes in the statistical fluctuations. The respective standard deviations are 1.968×10^{16} molec cm^{-2} and 1.974×10^{16} molec cm^{-2} for

the fits including the final choice of NO_2 cross-sections and only one cross-section, respectively. Both values differ by less than 1 %, whereby the final choice of cross-sections leads to marginally lower variability. Also the fit RMS-values of both fit scenarios differ only slightly with the optimised fit settings yielding smaller values. Lastly, it can be stated that all findings are independent of the elevation angle.

Summarising all findings of this and the previous section, the inclusion of additional NO_2 cross-sections in the DOAS fits for HCHO is well justified and even essential in the case of high NO_2 abundances. For synthetic spectra, the agreement between retrieved and true reference HCHO dSCDs is significantly improved with additional NO_2 cross-sections for the cases of high NO_2 abundances. Further, enhanced consistency between the results for the high and low NO_2 scenarios of a given geometry (and HCHO scenario) is achieved. The combination of NO_2 cross-sections which was finally selected based on the results for the synthetic spectra using a combination of both previously described approaches, namely the one additionally including the NO_2 cross-section at 220 K and both Puķīte terms, also leads to improvements of the fit RMS for measured spectra. Further, a small influence on the retrieved HCHO dSCDs is observed including this combination in the DOAS fit. However, here it should be stressed that small inconsistencies between the findings obtained with synthetic and real spectra exist. In particular, the influences on the retrieved dSCDs are similar for Bayfordbury (high NO_2) and the MSM 68/2 ship cruise (low NO_2) which is different for the synthetic spectra for which the effects are clearly weaker for the low NO_2 scenarios. As described above, this is probably related to differences between the assumed profile shape of NO_2 used in the simulations of the synthetic spectra and the true atmospheric NO_2 profiles. Lastly, the investigations based on the ATTO data showed that the (statistical) variability of the retrieved HCHO dSCDs is almost not affected if additional (probably not needed) NO_2 cross-sections are included in the DOAS fit also under very low NO_2 conditions. Given this observation, the same set of NO_2 cross-sections is used for all NO_2 conditions (all stations) in the following also in sake of consistency. Hence, also the results obtained for the real measured spectra support the usage of additional NO_2 absorption cross-sections and thus the NO_2 cross-section at 220 K as well as the linear and square Puķīte terms are added to the settings summarised in Table 6.3 which are used for the following optimisation steps.

6.2.3 *O_3 absorptions*

Ozone (O_3) shows high abundances in the stratospheric ozone layer, whereby the vertical extent is several kilometres. Thus, like for NO_2, the ozone absorptions occur at different temperatures which was also considered in the simulations of the synthetic spectra. First tests showed that it is essential to include two O_3 cross-sections at two different temperatures, namely 223 K and 243 K, in the DOAS fits for HCHO. Therefore, the standard fit settings summarised in Table 6.3 include these two O_3 cross-sections. In addition, O_3 exhibits strong absorption features in the UV spectral range, where also HCHO is retrieved. Together with

its high abundances in the stratosphere, this leads to large optical depths and non-linearities as explained above which might strongly affect the spectral retrieval of HCHO. Thus, the usage of Puḳīte terms for O_3 could be reasonable. In principle, all stratospheric effects should cancel out when a sequential FRS is used (compare Sect. 3.2). However, the stratospheric O_3 absorptions are very strong and the stratospheric AMFs also show a weak elevation angle dependence. Hence, the stratospheric effects do not completely cancel out. For these reasons, the effects of Puḳīte terms for O_3, the importance of the second O_3 cross-section at 243 K and the influence of a O_3 Raman cross-section on the HCHO fit results are investigated in the following. Lastly, it should be mentioned that following the results of the previous studies on the NO_2 cross-sections, the standard fit settings summarised in Table 6.3 were modified. Consequently, the NO_2 cross-section at 220 K as well as the linear and square Puḳīte terms for NO_2 are already included in the fits performed for the investigations on the O_3 cross-sections.

6.2.3.1 *Synthetic spectra*

To study the effects of different sets of O_3 cross-sections included in the DOAS fit for HCHO, a similar approach as for the investigations on the NO_2 cross-sections was applied. In particular, the same 36 elevation sequences of synthetic spectra with noise, for the combinations of SZAs of 20°, 45° and 60° and RAAs of 30°, 90° and 180° for the cases of low and high NO_2 and HCHO abundances (compare Table 6.1) were analysed using the modified standard settings (compare Table 6.3 and optimised set of NO_2 cross-sections) and including 16 different combinations of O_3 cross-sections. For ozone, the Raman cross-section as well as the Puḳīte terms can in principle be calculated either from the cross-section at 223 K or the one at 243 K. Tests showed that this choice is not critical. Thus, the respective cross-sections were calculated from the O_3 cross-section at 223 K. Lastly, again the 90° spectrum of a given elevation sequence, defined as a specific combination of sun geometry and trace gas scenarios, without noise was used as FRS.

Like for the studies regarding the NO_2 cross-sections (compare Sect. 6.2.2.1), after calculating mean HCHO dSCDs and mean fit RMS-values over the results obtained for the 100 different noises, absolute and relative deviations from the true reference dSCDs (compare Sect. 6.1.2) were calculated for each elevation of a given sequence and each combination of O_3 cross-sections using Eq. 6.1 and 6.2. Then, these quantities were plotted as functions of elevation for each scenario (sun geometry and trace gas scenario). An example of such a plot is presented in Fig. 6.15 for the sequence with low NO_2, high HCHO, at an SZA of 45° and an RAA of 180°. The four previously introduced quantities are shown in the different panels, where results obtained with different combinations of O_3 cross-sections included in the fits are colour-coded. Finally, it should be stressed once more that here the standard combination ("Std") already includes two O_3 absorption cross-sections at two different temperatures and the three additional NO_2 cross-sections as discussed in Sect. 6.2.2.

As already described above, the need of two O_3 cross-sections at different temperatures becomes also evident in Fig. 6.15. In particular, a clear worsening of the agreement between retrieved and true reference HCHO dSCDs is found

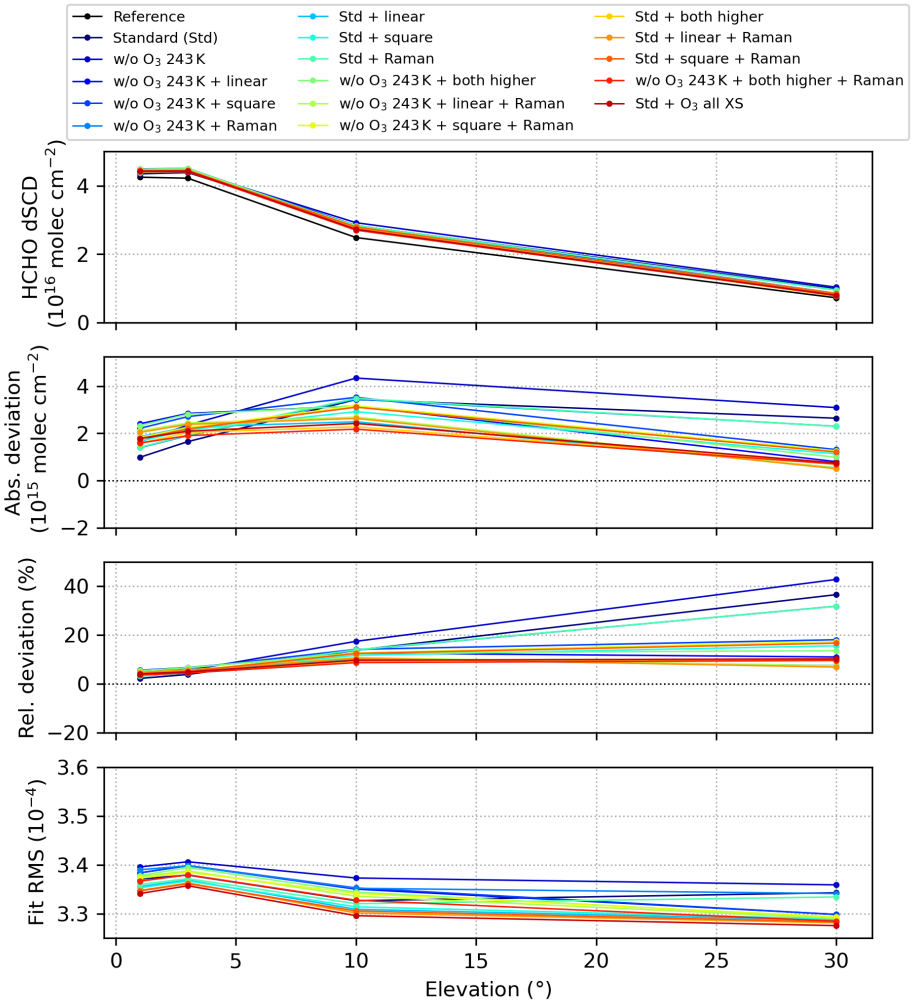


Figure 6.15: Mean retrieved HCHO dSCDs, absolute and relative deviations from the true reference values and fit RMS as a functions of elevation for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. Results for different combinations of Pukīte terms and other O_3 cross-sections are indicated by different colours.

at all elevations for the fit scenario which does not include the O_3 cross-section at 243 K (compare two darkest blue lines). Also for other combinations which do not include both O_3 cross-sections larger deviations can be observed at the lowest two elevations, while for most of these combinations improvements of the agreement between retrieved and true reference dSCDs is achieved at 10° and 30° elevation. This finding is valid for most scenarios.

For the here-presented scenario and most (but not all) of the not presented ones, worse agreements between retrievals and simulations are found for all sets of O_3 cross-sections at 1° and 3° elevation compared to the standard fit settings (compare second panel). Only for a few of the not presented sun geometries small improvements or almost no differences are observed at the lowest elevations for some sets of cross-sections. Contrary to that, many of the tested combinations of O_3 cross-sections lead to significant improvements of the absolute deviations at

10° and 30° elevation, whereby the variability amongst the different combinations is comparably small at these elevations. These improvements at the highest two elevations are found for seven out of nine sun geometries. Both findings for the absolute deviations are independent of the NO_2 and HCHO scenarios. Combining these two effects (worsening of the deviations at low elevations and improvements at high elevations), for the presented and many other scenarios, some combinations including additional O_3 cross-sections lead to a smoother elevation dependence of the deviations from the reference dSCDs, i.e. the absolute deviations within one elevation sequence for a given set of O_3 cross-sections vary less compared to the standard settings.

Also for the fit RMS, the need of the second O_3 cross-section at 243 K becomes evident, especially at low elevations, where all combinations not including this cross-section yield larger values than the standard settings (darkest blue line). At 10° and 30°, improvements of the fit RMS-values are observed for almost all combinations that include additional O_3 cross-sections, while also here the one including only one O_3 cross-section performs worst. The general patterns are similar for all scenarios, whereby the specific degree of improvement depends on the sun geometry. Altogether (considering all scenarios), at low elevations the RMS can be improved by up to 5 %, while the maximum improvement at 30° is 3 %.

All findings from above do almost not depend on the HCHO and NO_2 scenarios, i.e. the (relative) patterns of the results for the different combinations of O_3 cross-sections are similar amongst each other for different trace gas scenarios for a given sun geometry. In contrast to that, the results obtained with different sets of O_3 cross-section included in the DOAS fit depend rather strongly on the sun geometry. This was not the case for the investigations regarding the NO_2 cross-sections for which the strongest dependence was found on the NO_2 scenario. A discussion of this difference is provided at the end of this section. To illustrate the dependence of the results on the sun geometry, Fig. 6.16 shows the same four quantities as functions of elevation as Fig. 6.15 but for an SZA of 60° and an RAA of 180°.

A comparison of these two figures reveals the differences in the relative behaviours of the results obtained with the different sets of O_3 cross-sections amongst each other, in particular, the order of the coloured lines and their respective elevation dependencies. For the investigations on the NO_2 cross-sections, these two sun geometries showed the same relative patterns for the results obtained with the different cross-section combinations, whereby only the absolute values of the deviations differed. A further finding is that especially the sun geometries with SZAs of 45° and 60° in combination with a small RAA, namely 30°, show overall bad agreements between retrieved and true reference dSCDs (not shown).

Like for the studies regarding the additional NO_2 cross-sections, from this kind of plots it is difficult to judge which of the 16 combinations of O_3 cross-sections performs best since information from 36 figures have to be compiled. Therefore, again the sums (Σ_{abs}) and the RMS (RMS_{abs}) of the absolute deviations within each sequence for each combination of O_3 cross-sections were calculated using Eq. 6.3 and 6.4. An example of the resulting RMS_{abs} and Σ_{abs} for all 16

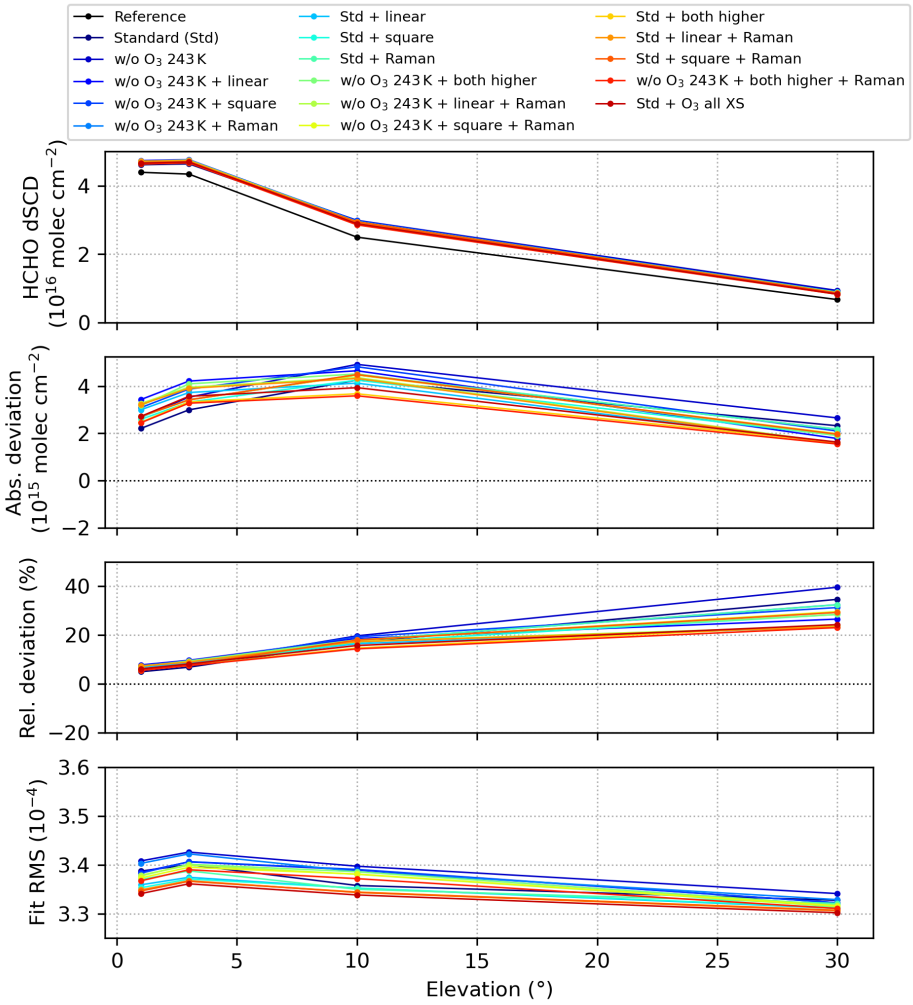


Figure 6.16: Same as Fig. 6.15 but for an SZA of 60° and an RAA of 180°.

combinations of O_3 cross-sections for the sequence shown in Fig. 6.15 is presented in Fig. B.3 in the appendix. The observations from above that two O_3 cross-sections at different temperatures are needed and that several combinations including additional O_3 cross-sections can reduce the absolute deviations from the true reference dSCDs (for high elevations) are also represented by this example. Next, means and standard deviations over all scenarios (i.e. sun geometries and trace gas scenarios) were determined for both quantities for each set of tested O_3 cross-sections. Similar to Sect. 6.2.2.1, the combination yielding the lowest mean RMS_{abs} together with a small mean Σ_{abs} is regarded as the best combination as it leads overall to the lowest deviations from the true reference HCHO dSCDs.

The general findings from above were shown to be almost independent of the NO_2 and HCHO scenarios. This remains valid for the Σ_{abs} and RMS_{abs}-values. More specifically, for a given sun geometry, the relations of both quantities obtained for different combinations of O_3 cross-sections amongst each other are independent of the trace gas scenario. This also holds for their respective means over all sun geometries. Only a quasi constant (absolute) shift between

the trace gas scenarios is found (not shown). Given these findings, all trace gas scenarios are considered together in the following. An overview of the mean Σ_{abs} (lower panel) and RMS_{abs}-values (upper panel) for all 16 sets of O_3 cross-sections calculated over all scenarios is provided in Fig. 6.17, whereby the error bars indicate the corresponding standard deviations.

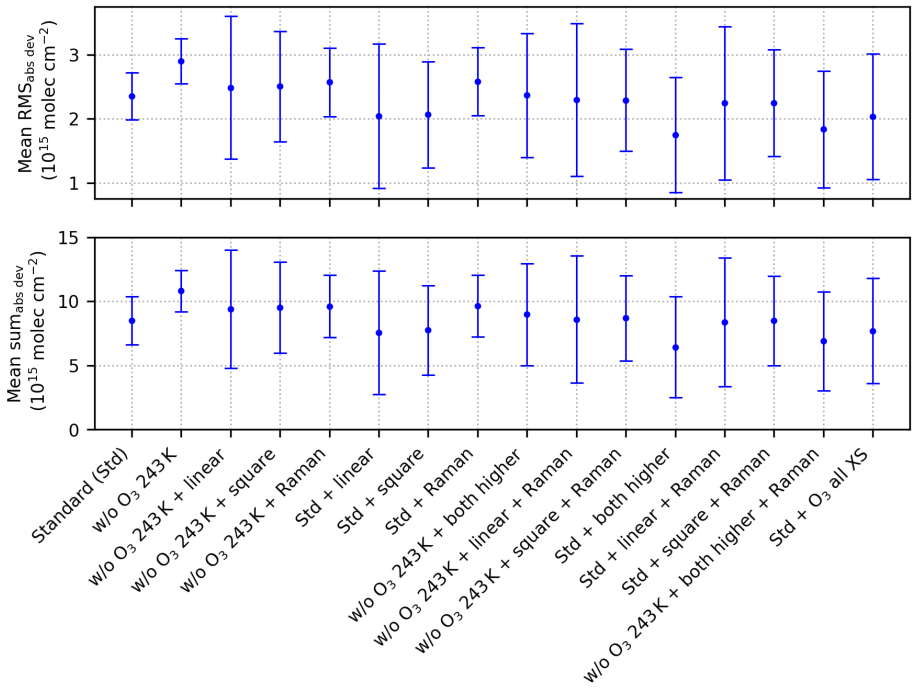


Figure 6.17: Mean RMS (upper panel) and mean sums (lower panel) of the absolute deviations from the true reference dSCDs for all 16 combinations of O_3 cross-sections over all scenarios. Error bars indicate the corresponding standard deviations.

A first observation is that the inclusion of additional O_3 cross-sections in the DOAS fit increases the variability of both quantities between the different sun geometries compared to the standard settings and the combination without the cross-section at 243 K (compare standard deviations). Specifically, this indicates that the different cross-section combinations perform differently well for different sun geometries which was already observed above. Since also the standard deviations are independent of the NO_2 and HCHO scenarios, this applies to all trace gas scenarios. Despite the increased variability, the RMS_{abs} is slightly reduced for many scenarios which include additional O_3 cross-sections. The absolutely lowest mean RMS_{abs} (and mean Σ_{abs}) is found for the combination which includes besides the two standard cross-sections the linear and square Puḳīte terms ("Std + both higher") indicating that this combination leads, at least for some sun geometries, to improved agreements between retrieved and true reference HCHO dSCDs. As outlined above, these improvements mainly occur at elevations of 10° and 30°, while at 1° and 3° for some sun geometries also slightly worse agreements compared to the standard settings can occur. Nevertheless, the combination which additionally includes both Puḳīte terms leads to an overall

improvement for six out of nine investigated sun geometries. Only for two cases an overall worsening of the results is observed, namely for the sun geometries with an RAA of 30° for SZAs of 45° and 60° which were identified as problematic already before. As mentioned above, another finding is that most combinations including additional O_3 cross-sections yield smoother elevation dependencies of the deviations. Consequently, these fit scenarios lead to comparable deviations at the different elevations within a sequence and thus the effects of the imperfect fit results on the profile retrieval, more precisely, the resulting profile shape are minimised. Amongst these combinations is the one which yields the lowest mean RMS_{abs} (compare "Std + both higher" in Fig. 6.15).

In conclusion, the set of O_3 cross-sections which additionally includes both Puķīte terms can be considered as best choice based on the results for the synthetic spectra. Hence, this combination of cross-sections is used for the further optimisations of the DOAS fit settings for HCHO since the advantages outweigh the small drawbacks.

Lastly, it should be stressed once more that the results for the different sets of O_3 cross-sections depend rather strongly on the sun geometry (compare above), while they are almost independent of the trace gas scenario. This was not the case for the investigations regarding the NO_2 cross-sections, where mainly the NO_2 scenario was important. This is somehow a surprising finding. One possible explanation is probably the altitude at which, on average, the last scattering event of the incident photons occurs before they are detected. In particular, the question is whether the scattering occurs below or above the bulk of O_3 and NO_2, respectively. To get a general impression, boxAMFs (compare Sect. 3.2.3), which describe the sensitivity to an absorber within a given altitude range, were simulated for two sun geometries, namely for SZAs of 20° and 60° at an RAA of 30°, for the case of low NO_2 and high HCHO. Besides some detailed patterns, the radiative transfer simulations confirm the hypothesis that usually the atmospheric scattering takes place below (or within the lowest part of) the ozone layer, while the scattering typically occurs above the tropospheric NO_2 layer except for the lowest elevation angles. Hence, the strong dependence of the results for the investigations regarding the O_3 cross-sections on the sun geometry can be understood.

6.2.3.2 *Measured spectra*

Measured spectra from the same selected days as above for the considerations on the NO_2 cross-sections recorded at Bayfordbury and during the MSM 68/2 ship cruise (compare Sect. 6.1.1 and 6.2.2.2) were used to examine the performances of different combinations of O_3 cross-sections included in the DOAS fit under real measurement conditions. Again these investigations are based on improvements of the fit RMS as well as the influences on the retrieved HCHO dSCDs, whereby a special focus is put on the set of O_3 cross-sections which additionally includes the linear and square Puķīte terms which was found to be best based on the findings obtained for the synthetic spectra. For these spectra, largest effects (mainly improvements) of the different cross-section combinations were found at 10° and 30° elevation. Therefore, Fig. 6.18 shows results for spectra measured at

10° on 29 March 2019 at Bayfordbury. Like Fig. 6.12, the uppermost panel shows time series of HCHO dSCDs retrieved including different combinations of O_3 cross-sections (represented by different coloured lines) in the DOAS fit together with the resulting VCDs of NO_2 (black dots) and HCHO (green dots). Time series of the differences between the results obtained with varying combinations of O_3 cross-sections and the ones obtained with the standard settings are shown in the second panel to better reveal the differences amongst the coloured lines in the uppermost panel.

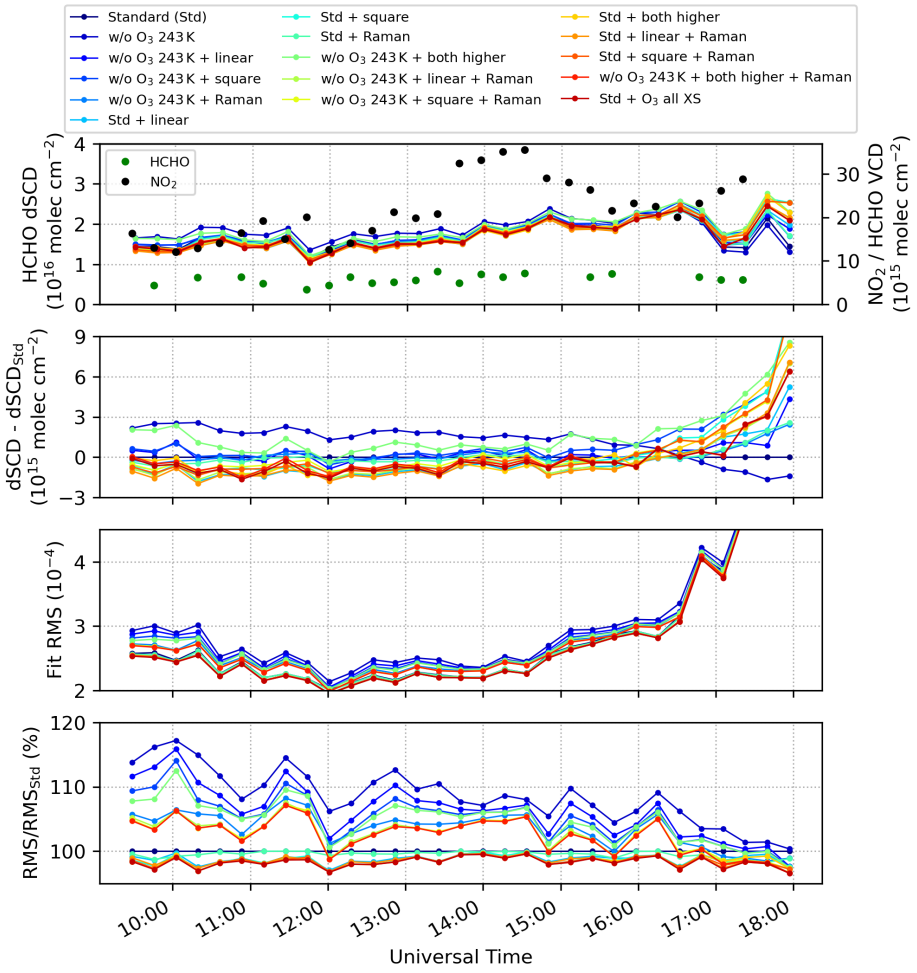


Figure 6.18: Time series of HCHO dSCDs retrieved including different sets of O_3 cross-sections in the DOAS fit (uppermost panel), corresponding differences with respect to the standard settings (second panel), fit RMS-values (third panel) and ratios of these fit RMS-values with respect to the ones obtained for the standard settings (lowermost panel) for spectra recorded at Bayfordbury on 29 March 2019 at 10° elevation. Results obtained with different combinations of O_3 cross-sections included in the DOAS fit are indicated by different colours. The uppermost panel additionally shows the resulting NO_2 (black dots) and HCHO (green dots) VCDs.

The HCHO dSCDs retrieved with different sets of O_3 cross-sections differ by about 2.5×10^{15} to 4×10^{15} molec cm^{-2} during large parts of the day. Towards

evening with increasing SZA, the spread of the dSCDs reaches values of up to 1×10^{16} molec cm^{-2}. The results of two combinations which do not include the second O_3 cross-section at 243 K, namely the one including only the cross-section at 223 K ("w/o O_3 243 K") and the one additionally only including both Puḳīte terms ("w/o O_3 223 K + both higher"), differ most from the results of the remaining combinations during large parts of the day. This is consistent with the findings obtained using the synthetic spectra and stresses the importance of the second O_3 cross-section at 243 K. Ignoring these two fit scenarios, the spread amongst the remaining combinations of O_3 cross-section is in the range of 1.5×10^{15} to 2×10^{15} molec cm^{-2} during large parts of the day still reaching values of up to 9×10^{15} molec cm^{-2} towards evening. These values are comparable to the spread amongst these combinations which was observed for the synthetic spectra at 10° elevation for many scenarios. However, for some sun geometries, the spread was even below 1×10^{15} molec cm^{-2} for the synthetic spectra. Another observation for the synthetic spectra was that for roughly half of the sun geometries, the effects of the different additional O_3 cross-sections were rather small at low elevations and comparably large at 10° and 30° elevation (compare Fig. 6.15). Contrary to that, for the real data, only a weak elevation dependence of the influences of the different fit scenarios on the fit results is observed (not shown). Nevertheless, the effects are generally smaller at low elevations. Therefore, the results of the spectra measured at Bayfordbury are more representative for sequences of synthetic spectra with sun geometries which show weak elevation dependencies of the impacts of the different fit scenarios on the retrieved HCHO dSCDs, for instance, the one presented in Fig. 6.16, where similar influences of the different cross-section combinations are found at all elevation angles. This can probably be explained by the fact that on average in spring at the mid-latitude station Bayfordbury the sun geometries are better represented by scenarios with SZAs above 45°. The dSCDs retrieved with different fit scenarios are comparable amongst each other and are overall smaller than the ones obtained using the standard settings (compare second panel in Fig. 6.18), except for the clearly problematic ones. This is consistent to the results for the synthetic spectra for which the standard combination of O_3 cross-sections overestimated the HCHO dSCDs at 10° elevation for all scenarios and the majority of the other combinations yielded (depending on the sun geometry) approximately 1×10^{15} to 2×10^{15} molec cm^{-2} smaller dSCDs (compare Fig. 6.15 and 6.16). Also the relative patterns at the other elevations are similar to the synthetic spectra. For instance, at low elevations not always lower dSCDs compared to the standard settings are found for the tested combinations. Towards evening, the spread between the different fit scenarios increases and almost all sets of O_3 cross-sections lead to larger dSCDs than the standard combination which can be explained by enhanced O_3 absorptions at these large SZAs leading to larger effects on the HCHO results. Actually, also for the synthetic spectra, many combinations had a tendency towards higher dSCDs already at an SZA of 60° (except for 10° elevation). At 10° elevation, the final choice of O_3 cross-sections ("Std + both Puḳīte terms") yielded on average about 1.5×10^{15} molec cm^{-2} (up to 2.5×10^{15} molec cm^{-2} at maximum) smaller HCHO dSCDs compared to the standard settings for most sequences of synthetic spectra.

For the spectra measured at Bayfordbury, the final choice of O_3 cross-sections leads to on average ca. 1×10^{15} molec cm^{-2} smaller dSCDs than the standard fit scenario during large parts of the day. Only towards evening with increasing SZA larger dSCDs are retrieved. Thus, the findings obtained with synthetic and real measured spectra are overall consistent, whereby the effects are slightly smaller for the real data. This also holds for the other elevations. Since no true reference values are known for the real data, the fit RMS is assessed to examine the influences of the different cross-section combinations on the fit quality.

Time series of the fit RMS-values for the different sets of O_3 cross-sections are shown in the third panel of Fig. 6.18, while the lowermost panel presents the corresponding ratios with respect to the values of the standard settings. The most striking feature is that all combinations not including the O_3 cross-section at 243 K show significantly larger RMS-values. This is also true for the other elevations. While the worsening is up to 17 % at 10°, it is only up to 10 % at the not presented elevations. These findings again stress the importance of the second O_3 cross-section. All other combinations of additional O_3 cross-sections lead to small improvements of the fit RMS compared to the standard combination. The largest improvements are found at 10° elevation with roughly 2 % on average and ca. 4 % at maximum. The lines representing the different fit scenarios leading to improvements are hardly to discriminate, i.e. the results for these scenarios are similar with the final choice being amongst them. For the synthetic spectra, the maximum difference between maximum and minimum fit RMS-values for different scenarios and fit settings (including the ones without the O_3 cross-section at 243 K) was only roughly 5 % and thus significantly smaller than the maximum effect for the Bayfordbury data. This can probably be explained by the fact that for synthetic spectra the SZAs (and the RAAs) of all spectra within in a given sequence are exactly the same. This means that the sun geometries for a given spectrum at low elevation and the corresponding FRS spectrum used for the DOAS fit are identical. Naturally, this is not the case for real measured spectra. Although the FRS spectrum used to perform the DOAS fit for a given spectrum is interpolated between the 90° spectra before and after the sequence to its measurement time, small differences in the sun geometries possibly remain and not all radiative transfer effects cancel out completely. Actually, this might also explain the large spread of the results towards evening (compare second panel), where this effect might be even stronger. A last observation is that for the fit scenarios leading to worse RMS, a relatively strong decrease of the effect on the RMS during the course of the day is found. However, no explanation for this effect can be provided so far.

Figure 6.19 shows similar time series as Fig. 6.18 but for spectra measured during the MSM 68/2 ship cruise on 12 November 2017 at 10° elevation.

For the synthetic spectra, the results were almost independent of the trace gas scenario and mainly depended on the sun geometry. However, comparing Fig. 6.18 and 6.19 shows that the patterns of the results obtained with different combinations of O_3 cross-sections included in the DOAS fit are different in some aspects for the two data sets of measured spectra. For instance, for the MSM 68/2 data all combinations not including the O_3 cross-section at 243 K lead to larger

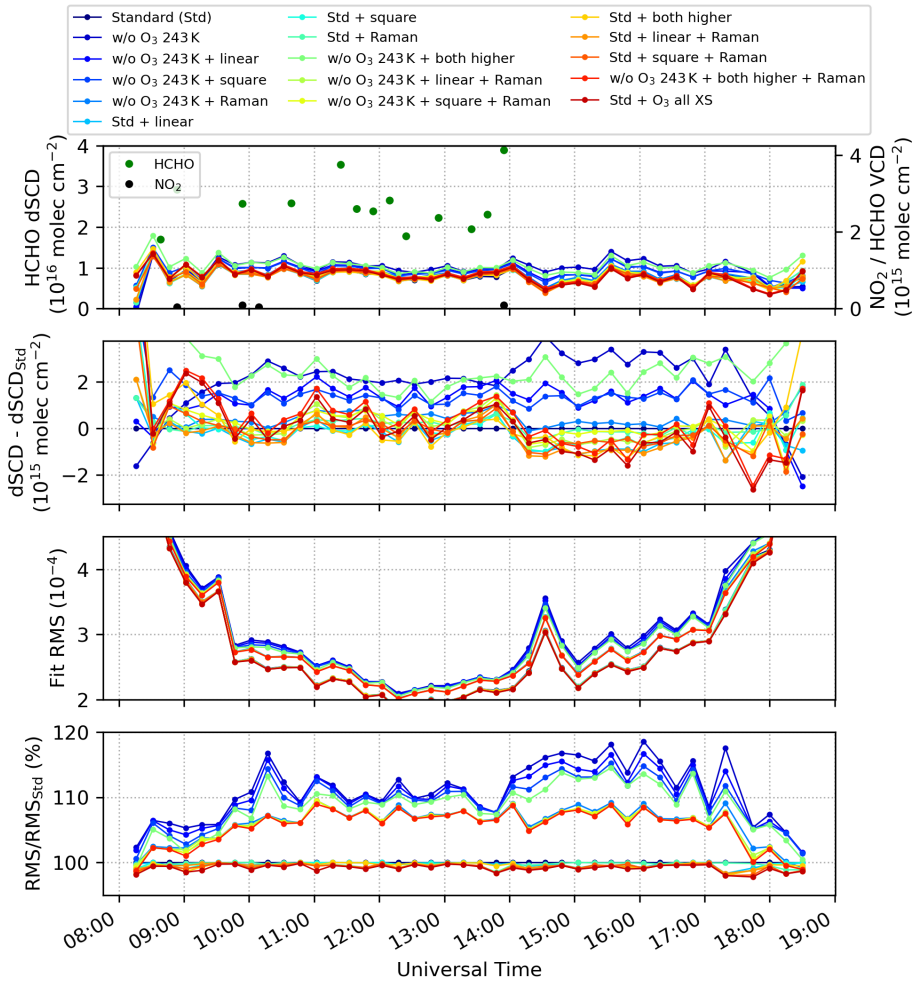


Figure 6.19: Same as Fig. 6.18 but for spectra measured during the MSM 68/2 ship cruise on 12 November 2017 at 10° elevation.

HCHO dSCDs than the standard combination (except for some combinations in the afternoon), while this was only the case for two of such combinations for the Bayfordbury data (compare second panels in both figures). This finding is independent of the elevation angle. When these problematic fit scenarios are ignored, a further observation is that until approximately 2 p.m. no consistent effects of the different (remaining) combinations of O_3 cross-sections on the retrieved HCHO dSCDs are found, whereby the corresponding results scatter around the ones obtained with the standard settings. In the early morning, even slightly larger dSCDs are obtained which is different from the Bayfordbury measurements. Nevertheless, after 2 p.m. systematically, by roughly 1×10^{15} molec cm^{-2}, lower dSCDs are retrieved with all reasonable sets of O_3 cross-sections. These effects are similar for all of these combinations within approximately 1×10^{15} molec cm^{-2} and of the same order as for the Bayfordbury measurements. Towards evening with increasing SZA, large scatter (ca. 2×10^{15} to 2.5×10^{15} molec cm^{-2}) between the results obtained with different fit scenarios is observed which again is likely caused by strong O_3 absorptions. At the other not presented elevations, the effects

of different O_3 cross-section scenarios are smaller. For the low elevations, this is consistent with the previous observations, especially with the results for the synthetic data, while for 30° this is different. The latter might be related to the comparably weak HCHO absorptions on the North Atlantic, however, the HCHO VCDs on the selected day are of the same order of magnitude as the ones of the low formaldehyde scenario of the synthetic spectra (compare Sect. 6.1.1). Thus, other differences between both data sets, such as assumptions on the profile shape, are probably even more important. The final choice based on the results obtained with the synthetic spectra ("Std + both Puḷīte terms") yields around 1.5×10^{15} molec cm^{-2} larger HCHO dSCDs until 10 a.m. followed by a period in which the results scatter around the ones of the standard fit scenario. Finally, starting at 2 p.m. up to ca. 1.5×10^{15} molec cm^{-2} smaller dSCDs are retrieved with the final choice of O_3 cross-sections. In general, the effects of the finally chosen combination of O_3 cross-sections are of the same order as for Bayfordbury, while the detailed patterns differ, e.g. the results at the not presented elevations or during the morning. Further, the overall variability is larger for the MSM 68/2 data which might be, at least partly, explained by lower HCHO dSCDs.

The general findings for the fit RMS (third and lowermost panels of Fig. 6.19) are comparable to the ones for Bayfordbury. In particular, all combinations not including the O_3 cross-section at 243 K lead to clearly larger fit RMS-values than the standard settings which is also valid for the other elevations. While the worsening is up to 19 % at 10° elevation, it reaches up to ca. 15 % for the not presented elevations. Especially the latter is larger as seen for the Bayfordbury data, where particularly at the lowest elevations the worsening was below 10 %. The remaining combinations of additional O_3 cross-sections, including the final choice, lead to only small improvements of the fit RMS and the effects are smaller than for Bayfordbury. More specifically, the improvements are at maximum ca. 2 % for the MSM 68/2 data, whereby the average is below 1 % for all elevations. Like for Bayfordbury, the maximum difference between the maximum and minimum fit RMS-values is larger than observed for the synthetic spectra where it was 5 % at maximum. This finding can be explained as outlined in detail above for Bayfordbury.

One major finding of this and the previous section is that including a second O_3 cross-section at 243 K in the spectral retrieval of HCHO is essential since clearly worse agreements between retrievals and simulations were found for almost all scenarios of synthetic spectra for several fit scenarios which do not include this cross-section. Also the fit RMS is increased in these cases, whereby this effect was found to be particularly important for the real measured spectra. The latter aspect can be explained by the fact that not all radiative transport effects cancel out despite the usage of an interpolated sequential FRS for the measured spectra. Another finding is that the effects of the different combinations of O_3 cross-sections on the fit results depend rather strongly on the sun geometry. Further, the variability between the results obtained for different sun geometries is increased compared to the standard fit settings. Nevertheless, improved agreements between retrieved and true reference dSCDs can be achieved with several combinations of additional O_3 cross-sections for six out of nine sun geometries.

Largest improvements are observed at elevations of 10° and 30°, while at the two lower elevation angles not always an improvement is found. Consequently, several tested combinations lead to a smoother elevation dependence of the absolute deviations from the true reference dSCDs and can therefore reduce the influence on the profile shapes of the HCHO retrieval as explained above. For the real data, in general similar results were obtained. However, the effects were found to be smaller and especially for the MSM 68/2 data in some aspects deviations were observed. Moreover, the described elevation dependence of the influences of the different combinations of O_3 cross-sections on the fit results is less pronounced for the real measured spectra. Based on the deviations of the HCHO dSCDs retrieved with different sets of O_3 cross-sections included in the DOAS fit from the true reference values for the synthetic spectra, the combination which additionally includes the linear and square Puḳīte terms was found to be best suited. Also for real data, this combination showed reasonable results and slightly improved the fit RMS. Thus, this set of O_3 cross-sections was selected and is therefore used in the DOAS fits for the further optimisations steps.

6.3 CHOICE OF THE FITTING INTERVAL

A further important setting of the DOAS fit for a given trace gas is the wavelength range in which the fit is performed. Formaldehyde exhibits pronounced absorption bands in the UV spectral region between 225 and 375 nm (Meller and Moortgat, 2000). However, in this work passive MAX-DOAS measurements are used. Thus, only sunlight with wavelengths that reach the Earth's surface with sufficiently high intensity can be used. Since in the stratospheric ozone layer sunlight with short wavelengths is practically completely absorbed, the usable absorption bands of HCHO are limited[1] to a range from around 310 to 360 nm. In general, best fit results are expected when many of these absorption bands are covered by the fitting interval of the DOAS fit. Moreover, typically longer intervals lead to more stable results in terms of small statistical fluctuations in the resulting trace gas dSCDs. At the same time, absorption cross-sections of absorbers which have considerable absorption features in the respective fitting interval, e.g. BrO in the case of the HCHO retrieval, have to be included in the DOAS fit besides the cross-section of the target species. This can lead to spectral interferences which cause artificial enhancements or reductions of the dSCDs of the trace gas of interest. Hence, shorter fitting intervals excluding such interfering absorption bands can be better suited than longer intervals including more HCHO bands. For that reason, in the following the performances of different fitting intervals for the DOAS fit of HCHO are investigated.

6.3.1 *Study approach*

To investigate the performances of different fitting intervals, synthetic spectra (with noise) and spectra measured during the MSM68/2 ship cruise were analysed using varying fitting intervals. For that, the boundaries of the fitting intervals were

1 The HCHO absorption bands above 360 nm are too weak to be used in a reasonable way.

varied in steps of 0.5 nm ranging from 310 to 350 nm and from 330 to 362 nm for the lower and upper boundaries, respectively. Intervals with lengths below 7 nm were skipped since the widths of the individual HCHO absorptions bands are in the order of approximately 5 nm at the spectral resolution of the spectrometers used in this work. Again, the standard fit settings summarised in Table 6.3 were used, whereby the additional cross-sections for NO_2 and O_3, namely the NO_2 cross-section at 220 K as well as the linear and square Puḳīte terms of both gases, which were found to be useful in the previous sections were already included in the fits.

Synthetic spectra

For the synthetic spectra (with noise), the same 36 elevation sequences as for the previous studies (combinations of SZAs of 20°, 45° and 60° and RAAs of 30°, 90° and 180°) for the cases of low and high NO_2 and HCHO (compare Table 6.1) were analysed using the described modified standard settings (compare Table 6.3 + modifications) and varying fitting intervals.

To assess the quality of the results obtained with the different fitting intervals, several quantities are examined. In particular, these are the fit RMS, the fit error, the variability of the HCHO dSCDs retrieved for the spectra with 100 different noise variants added, the deviations from the true reference dSCDs (absolute and relative) and the RMS of the absolute deviations within a given sequence.

For the fit RMS and the fit error, mean values were calculated for each scenario, defined as a specific set of trace gas scenario, sun geometry and elevation angle, over the results obtained for the 100 noise variants. The variability of the results is defined as the standard deviation of the 100 dSCDs retrieved for one specific scenario. The deviations were calculated from the mean of the 100 retrieved HCHO dSCDs using Eq. 6.1 an 6.2 as explained in Sect. 6.1.2 and 6.2.2.1. Lastly, the RMS of the absolute deviations from the true reference values within a sequence (RMS_{abs}) was calculated using Eq. 6.4.

For each tested fitting interval, these quantities were determined and plotted as functions of the lower and upper boundaries of the different fitting ranges following the approach from Vogel et al. (2013). Thereby, the values of a given quantity, e.g. the mean fit RMS, are displayed as coloured contours and the x- and y-axes indicate the lower and upper boundaries of the tested fitting intervals, respectively. Thus, each point in such a plot represents the result of one specific fitting interval. The resulting "maps" allow an easy interpretation of the dependencies of the results on the tested fitting ranges. Combining the findings of the previously explained quantities, an identification of well- and less-suited fitting intervals is possible.

Measured spectra

For the measured data, again spectra measured during the MSM 68/2 ship cruise on 12 November 2017 were used since the HCHO abundances on that day were comparably low (mainly marine background HCHO) and rather stable (compare uppermost panels in Fig. 6.13 and 6.19) especially around noon.

As no true HCHO dSCD are known for real data, the assessment of the results for the different scenarios is limited to the fit RMS, the fit error and the variability of the retrieved dSCDs. To create the previously described plots, a 4-hour time period from 11 a.m. to 3 p.m. (UT) in which the dSCDs are almost constant was chosen and mean fit RMS-values and mean fit errors for this time period were calculated for each tested fitting interval. The variability of the retrieved HCHO dSCDs for a given fitting interval is defined as the standard deviation of the dSCDs in this time period. Note that the crucial assumptions here are that the true atmospheric HCHO VCD is constant over the considered time period and that the variability of the retrieved HCHO dSCDs is solely caused by statistical fluctuations which are related to the DOAS fit and the choice of the fitting interval but not by real variations of atmospheric radiative transfer and/or the atmospheric HCHO content. As becomes evident in the uppermost panel of Fig. 6.19, the HCHO dSCDs in the mentioned time interval were rather constant with a standard deviation of only 1.06×10^{15} molec cm^{-2} which corresponds to roughly 10 % of the mean dSCD in this period (8.35×10^{15} molec cm^{-2}). Further, the SZA varied by only 11° in the considered time range. Another aspect in this regard is that the measurements during the MSM 68/2 ship cruise were performed under pristine background conditions on the North Atlantic far from any significant emission sources (compare Sect. 4.5). Hence, a homogeneous HCHO distribution can be presumed. In view of these considerations, the described assumptions are most likely reasonably well fulfilled.

Apart from that the real data were treated like the synthetic data and the calculated quantities were plotted in the above described way.

6.3.2 *Analyses and investigations*

In this section, the different quantities which are used to examine the quality of different fitting intervals are considered separately. For each quantity, first, the main findings are illustrated for one representative scenario of synthetic spectra. Here, the results obtained for the spectra for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° at 1° elevation were chosen since this scenario can be regarded to be representative for the main characteristics of the results and was already presented as example for the investigations on the higher order terms of NO_2 and O_3 (compare Sect. 6.2.2.1 and 6.2.3.1). Then, deviating results for the other (not shown) scenarios are discussed accordingly and possible differences amongst the scenarios and dependencies on different parameters are carefully assessed. For each quantity, this is directly followed by an evaluation of the corresponding results for the real measured spectra. Here, always the results obtained for the spectra at 1° elevation are shown as examples. Finally, an overall discussion is performed at the end of this section and the finally chosen fitting interval is justified.

Fit RMS

First, the influence of the choice of the fitting interval on the resulting fit RMS is assessed. For that, the mean fit RMS is shown as function of the lower and

upper boundaries of the tested fitting intervals in Fig. 6.20 for the representative scenario specified above. The lengths of the different intervals are represented by the upper x-axis and the corresponding skewed black lines. Fitting intervals which were used in several previous studies are marked by black circles. In addition, the normalised HCHO absorption cross-section is depicted on the x- and y-axes to provide an overview on the absorption bands that are covered by the different fitting ranges.

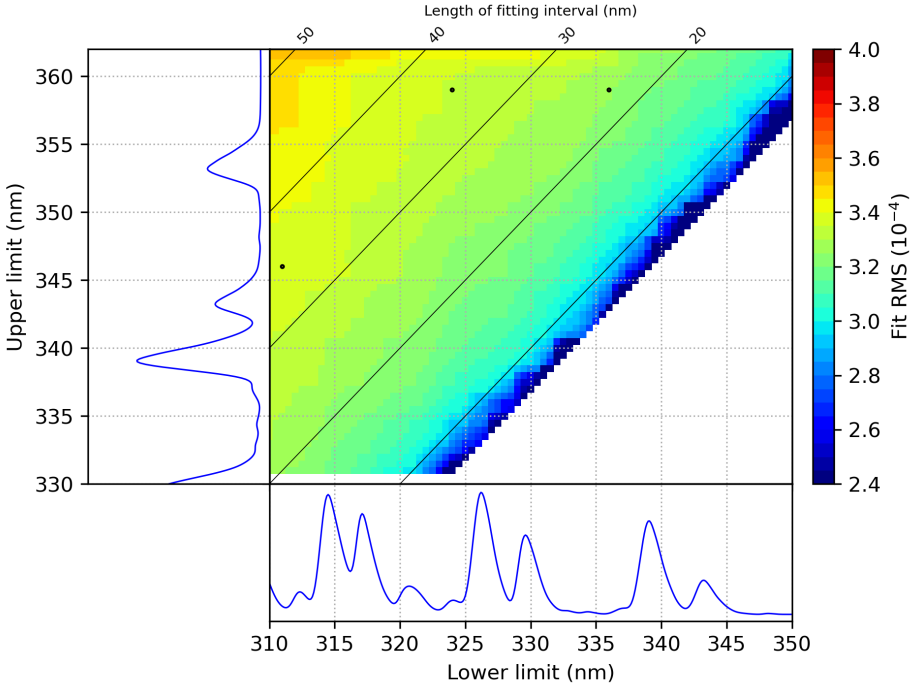


Figure 6.20: Mean fit RMS at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes.

Indicated by the observation that the coloured contours are almost parallel to the skewed black lines, the fit RMS is mainly determined by the length of the applied fitting interval with long intervals generally leading to larger RMS-values than shorter intervals. This overall characteristic is valid for all scenarios and independent of the elevation. Also the main 2D-patterns are almost independent of the scenario, while the absolute values naturally vary between different scenarios. For instance, for almost all sun geometries, the fit RMS decreases with increasing elevation within one sequence as expected. A more detailed look reveals that especially long intervals with lower boundaries at comparably short wavelengths tend to have slight higher RMS-values than intervals of the same lengths but lower boundaries at longer wavelengths. This effect is only weakly pronounced for the presented example scenario and stronger for other scenarios, e.g. the

scenario with an SZA of 60° at an RAA of 180°, and can most likely be explained by strongly increasing O_3 absorptions towards short wavelengths which negatively affect the DOAS fits. Further, long intervals with upper boundaries at long wavelengths show overall larger RMS-values. This effect is strongest at 1° and 3° elevation and is thus probably related to the inclusion of more and stronger NO_2 absorption bands. Hence, this effect is marginally stronger for the high NO_2 scenario (not shown). Both of the latter effects are for most scenarios only weakly pronounced.

The dependence of the mean fit RMS on the fitting interval for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017 at 1° elevation is presented in Fig. 6.21.

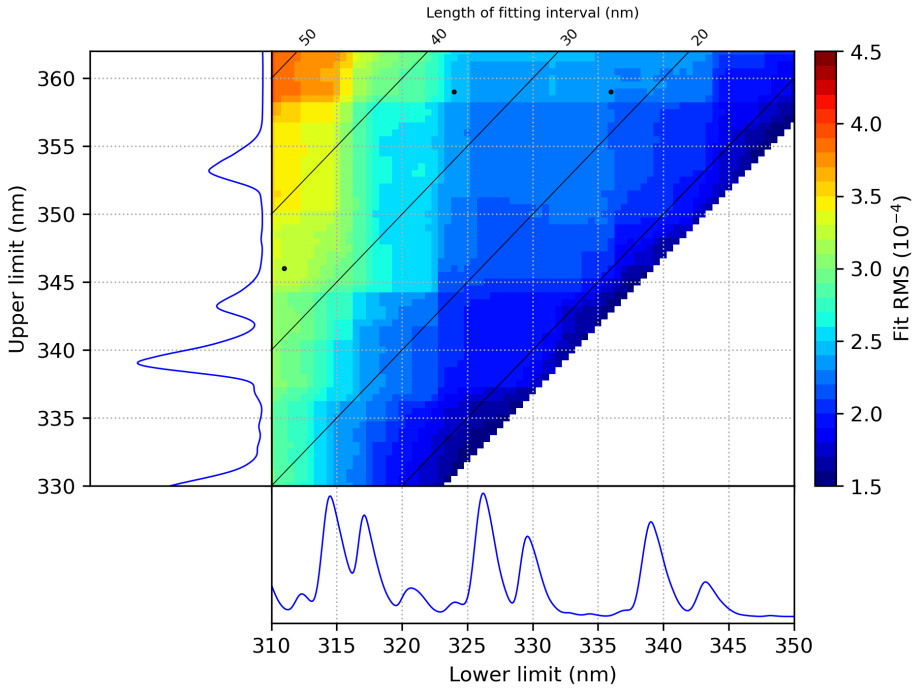


Figure 6.21: Same as Fig. 6.20 but for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017 at 1° elevation. Note the different colour maps in both figures.

In general, similar findings as for the synthetic spectra are true for the real measured data. In particular, long fitting intervals are mostly associated with larger RMS-values compared to shorter ones. Again, this general pattern is independent of the elevation angle, whereas the absolute values decrease strongly with increasing elevation which is stronger pronounced than for the synthetic data. However, for the measured spectra, the coloured contours are less parallel to the skewed black lines. This indicates that the observation, which was already made for the synthetic spectra, that long fitting intervals with lower boundaries at comparably short wavelengths tend to have higher RMS-values than intervals of the same lengths but lower boundaries at longer wavelengths is stronger pronounced for the real data. At 30° elevation, this becomes even the dominant effect (not shown). This difference between both data sets can at least partly be

explained by the fact that even if an interpolated FRS is used in the DOAS fits for the real data not all radiative transfer effects completely cancel out (compare Sect. 6.2.3.2) due to slightly varying sun geometries within one elevation sequence which is not the case for synthetic spectra. Thus, the increasing O_3 absorptions towards short wavelengths affect the results for the measured spectra stronger. Another effect possibly contributing here is that towards the edges of the detector (compare spectral range of used detectors in Sect. 4.2.1.2) its characteristics are worse which potentially amplifies the effects caused by strong O_3 absorptions. These slight differences between synthetic and real measured data illustrate the need to include measured spectra in the here-presented sensitivity study. Since the NO_2 absorptions in the spectra measured during the MSM 68/2 ship cruise are significantly smaller than the ones in the synthetic spectra, even for the low NO_2 scenario (compare Sect. 6.1.1), no effect of NO_2 absorptions at relatively long wavelengths as seen for the synthetic spectra above are found for the measured spectra. A last aspect is that the absolute values of the fit RMS are overall smaller for the real data (note different colour maps in Fig. 6.20 and 6.21). This suggests that, despite a careful selection, the noise level assumed for the simulations was probably slightly higher than in the real measurements.

Fit error

The *QDOAS* software which is used in this work to perform the DOAS fits also determines the statistical errors of the retrieved trace gas dSCDs. It is important to note that this fit error is determined assuming that the fit residual is only caused by statistical uncertainties (of the linear fit parameters) and does not contain systematic errors, e.g. wrong wavelength calibrations, uncertainties of the fitted absorption cross-sections or systematic residual structures (Stutz and Platt, 1996). Since unweighted fits are performed here, the statistical errors on the retrieved dSCDs can be estimated based on the reduced χ^2 (Danckaert et al., 2017; Stutz and Platt, 1996). Further information on the error analysis in DOAS fits can be found in Stutz and Platt (1996) and more specifically for the here used *QDOAS* software in Danckaert et al. (2017).

Figure 6.22 shows the mean fit error as function of the lower and upper boundaries of the varying fitting intervals for the same representative scenario as above.

As described above, the (statistical) fit errors in *QDOAS* are estimated based on the reduced χ^2. Therefore, it is closely linked to the fit RMS and thus also the fit error mainly depends on the length of the applied fitting interval. However, since the reduced χ^2 is calculated considering the degrees of freedom in the fit which increase with increasing length of the fitting interval, long intervals are associated with smaller fit errors. This becomes also evident in Fig. 6.22 by almost parallel coloured contours and skewed black lines. The general 2D-patterns are almost independent of the elevation angle. These two findings are also quasi independent of the sun geometry and the trace gas scenario. Further, intervals of the same lengths tend to have slightly smaller fit errors with decreasing wavelengths of the lower boundaries (e.g. intervals with a length of 25 nm in Fig. 6.22). This can be explained by the facts that the HCHO absorptions are

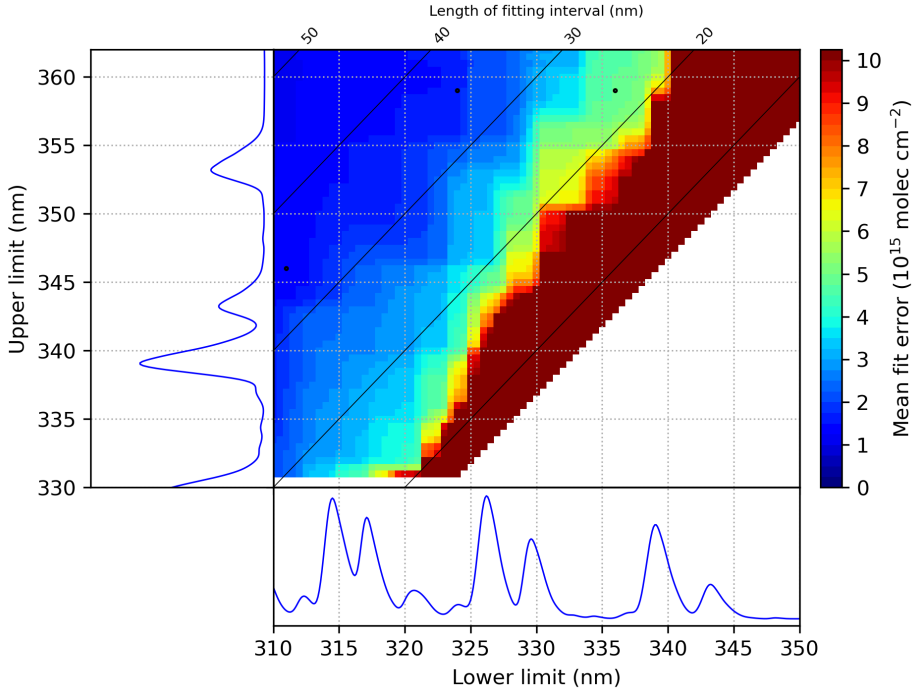


Figure 6.22: Mean fit error at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes.

stronger at shorter wavelengths than at longer wavelengths (compare absorption bands at ca. 326 and 339 nm) and that they are (spectrally) closer to each other for shorter wavelengths. The latter leads to more included absorption bands for these fitting intervals with lower boundaries at shorter wavelengths.

The dependence of the mean fit error on the choice of the fitting interval for the measured spectra is illustrated in Fig. 6.23.

Basically, the same findings as for the synthetic spectra are also valid for the real measured data and long fitting intervals lead to smaller fit errors than shorter ones. In addition, also for the spectra measured during the MSM 68/2 ship cruise, intervals of the same lengths tend to have smaller fit errors with decreasing wavelengths of the lower boundaries. However, this effect is less pronounced for the real data especially for the longest intervals. Compared to the synthetic spectra, the absolute values of the fit error decrease stronger with increasing elevation for the measured spectra. This is consistent to the findings for the fit RMS for which the elevation dependence was also more pronounced for the real measured spectra. However, also for the measured spectra, the overall 2D-patterns of the fit errors are independent of the elevation angle.

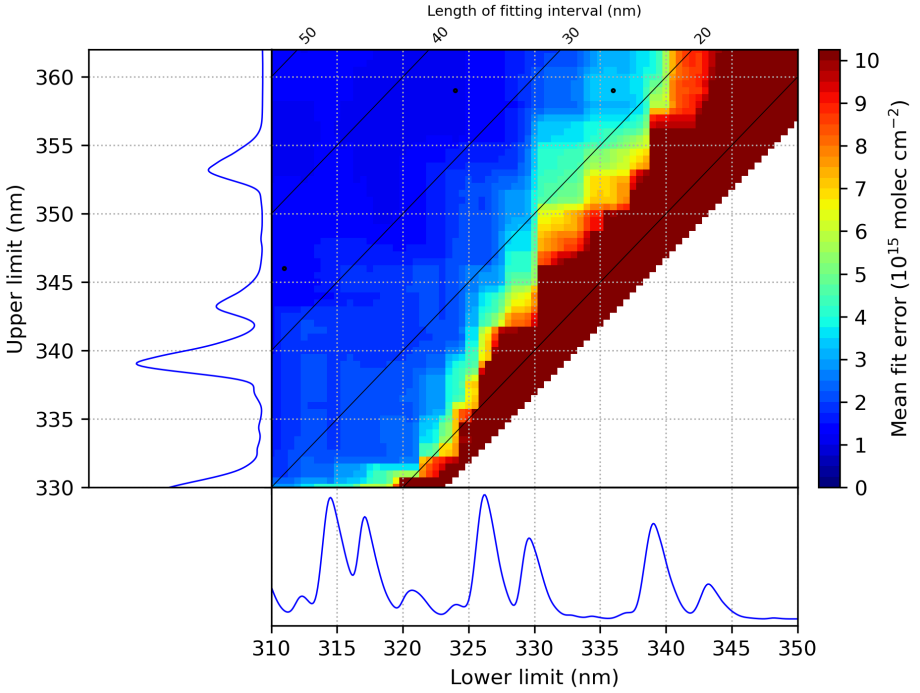


Figure 6.23: Same as Fig. 6.22 but for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017 at 1° elevation.

Variability of retrieved HCHO dSCDs

As outlined in Sect. 6.1.1, 100 spectra with different random noise spectra (all having the same RMS) added are available for each simulated scenario. Thus, 100 HCHO dSCDs are retrieved for each scenario, defined as a specific combination of viewing geometry and trace gas scenario, which in principle all should yield the same values. Like outlined above, here the standard deviation of these 100 dSCDs is used as measure for the variability of the retrieved HCHO dSCDs and allows to examine the stability of the results obtained with different fitting intervals. A well-suited interval should yield a small standard deviation of the 100 dSCDs indicating stable fit results for that specific interval. As an example, Fig. 6.24 presents the standard deviations of the 100 retrieved HCHO dSCDs obtained for different fitting intervals for the representative scenario as functions of the boundaries of these intervals.

Also the variability of the retrieved HCHO dSCDs is mainly determined by the length of the fitting interval, whereby long intervals lead to smaller standard deviations (compare parallel coloured contours and skewed black lines). This overall observation holds for all geometries and is independent of the NO_2 and HCHO scenarios. The absolute values slightly depend on the specific scenario and show a weak elevation dependence. For the latter, no clear behaviours can be identified. Nevertheless, the general 2D-patterns are not affected. Similar to the observations made for the fit errors, intervals of the same lengths tend to yield slightly smaller variabilities with decreasing wavelengths of the lower boundaries (e.g. intervals with a length of 35 nm in Fig. 6.24). Again, these findings for the

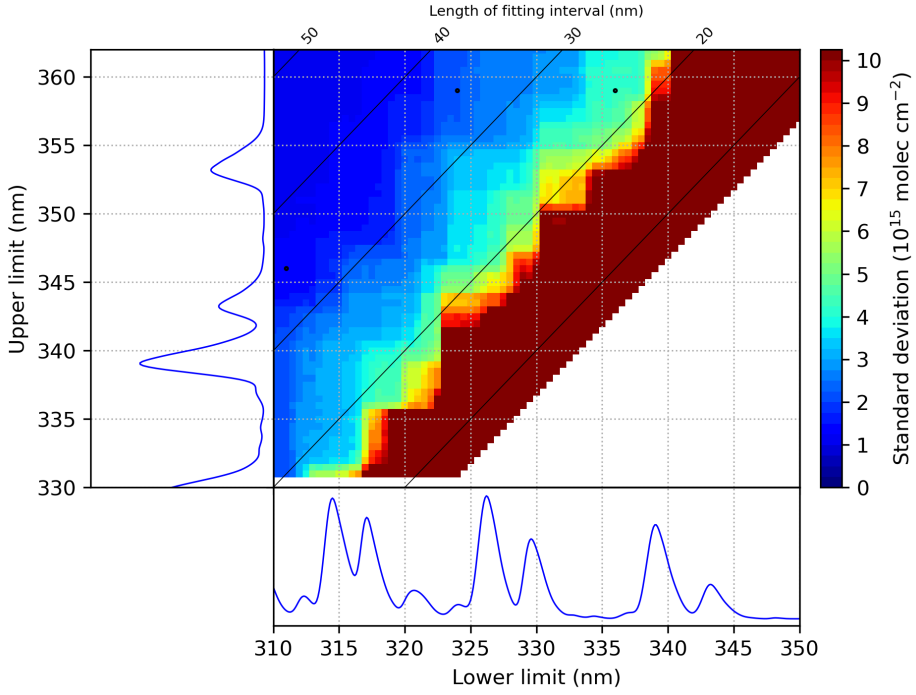


Figure 6.24: Standard deviation of the 100 retrieved HCHO dSCDs at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes.

synthetic spectra can mainly be explained by the spectral absorption structures of HCHO which are included in the different intervals, whereby the inclusion of more and stronger absorption bands at shorter wavelengths (and for longer fitting ranges) leads to more stable results.

Figure 6.25 depicts the standard deviations of the HCHO dSCDs retrieved for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017 at 1° elevation. As described below, red polygons mark spectral regions in which minima in the variability of the results are found.

Compared to the synthetic spectra, more complex patterns are found for the real measured spectra. Nevertheless, still the length of a given fitting interval is the dominant factor affecting the variability of the retrieved HCHO dSCDs, whereby long intervals generally lead to more stable results indicated by smaller standard deviations. Contrary to the synthetic data, for many, yet not all, intervals of the same lengths, lower interval boundaries at shorter wavelengths result in slightly larger variabilities of the dSCDs (e.g. intervals with a length of 25 nm). This finding can be explained by several aspects. Firstly, worse detector properties with a lower signal-to-noise ratio towards the edges of the CCD-chip as already explained above might contribute here. Secondly, the O_3 absorptions increase strongly towards short wavelengths which potentially has a twofold impact. In

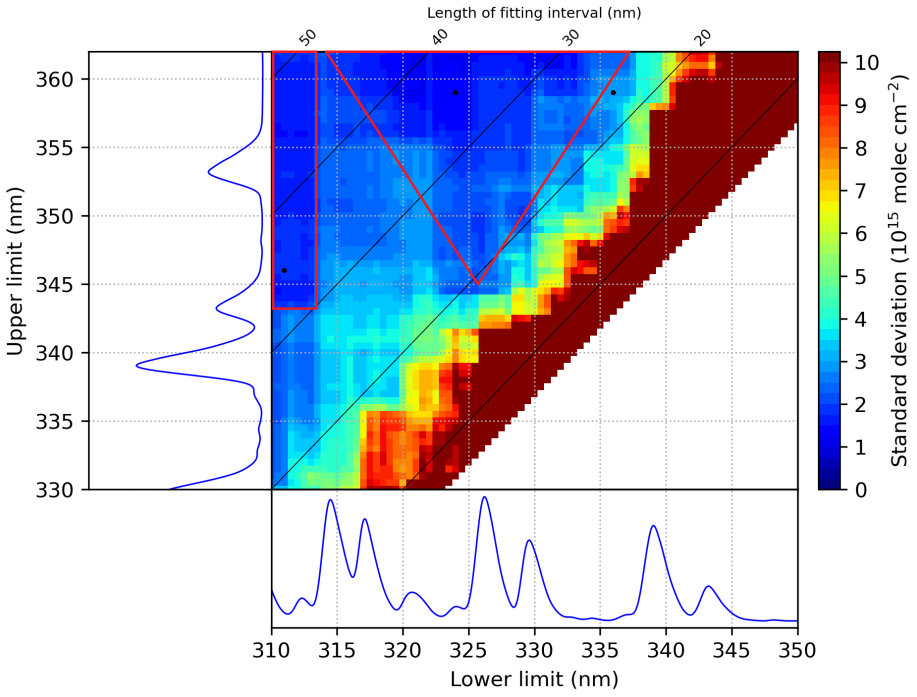


Figure 6.25: Same as Fig. 6.24 but for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017 at 1° elevation. Red polygons mark spectral regions in which minima in the variability of the results are found as described in the text.

particular, not all radiative transfer effects completely cancel out even though an interpolated FRS is used in the DOAS fits for the real data due to slightly varying sun geometries even within one elevation sequence. Further, the fitted O_3 cross-sections might not perfectly match the real atmospheric absorptions which can have a significant effect at short wavelengths, where its absorptions are very strong. In addition to this finding, minor local maxima exist in several spectral regions, e.g. for lower boundaries between ca. 314 and 322 nm in combination with upper boundaries between roughly 343 and 356 nm, which might be caused by spectral interferences with other cross-sections included in the DOAS fit.

Together these dependencies of the variability of the retrieved dSCDs on the fitting interval result in less smooth 2D-patterns in Fig. 6.25. This also holds for the other not presented elevations. Therefore, two red polygons mark spectral regions in which minima in the variability are found. For elevations above 3°, these spectral regions become smaller and at 30° the rectangular region is not visible anymore and also in this region rather high standard deviations are found. Possible reasons for the latter might be the fact that at higher elevation angles the sky is usually darker than for spectra measured at low elevations. Further, potentially existing horizontal variations of atmospheric properties, e.g. aerosols, can cause higher variabilities of the results at 30° which is probably less important at low elevations. Lastly, also the absolute values of the HCHO dSCDs are lower at 30° elevation and the noise probably affects the fit results (relatively) stronger. Besides that, the general 2D-patterns are comparable for the

different elevations, while the exact values of the standard deviations naturally differ slightly. For instance, at 10° elevation, the triangular shaped spectral region does not reach as far towards longer wavelengths for the lower boundary as for the low elevations and the local maximum of the variabilities between both polygons is less pronounced.

The differences between the findings for the synthetic and real data show that results obtained for synthetic spectra cannot always directly be transferred to spectra measured under real atmospheric conditions and stress the importance to also include real measurement data in the sensitivity study. The latter help to test whether conclusions based on simulated spectra are also valid under real measurement conditions. In particular, regarding the variability of the retrieved HCHO dSCDs, not all fitting intervals of a sufficient length, which was found to be the major dependence for the synthetic spectra, are well-suited under real atmospheric conditions, where more confined spectral regions were found to perform well. A last aspect that should be kept in mind for the spectra measured during the MSM 68/2 ship cruise is that the crucial assumptions are that the true atmospheric HCHO content is constant over the considered time period (11 a.m. to 3 p.m.) and that the variability of the retrieved dSCDs is solely caused by statistical fluctuations. As described earlier (compare Sect. 6.3.1), these assumptions are most likely reasonably well fulfilled.

Deviations

While for the real measured data only the three previously investigated quantities can be examined, the main advantage of synthetic spectra is that the HCHO dSCDs retrieved with varying fitting intervals can be compared with true reference dSCDs based on the VCDs which were used for the simulations of the spectra. This allows an identification of well-suited fitting intervals which lead to only small deviations from the true reference values. Therefore, absolute and relative deviations were calculated from the mean of the 100 retrieved HCHO dSCDs for each tested fitting interval using Eq. 6.1 and 6.2. As an example, the results for the previously used representative scenario are presented in Fig. 6.26 which shows the resulting relative (upper panel) and absolute (lower panel) deviations as functions of the lower and upper boundaries of the different tested fitting ranges.

Compared to the three previously discussed quantities, the interpretation of the results for the deviations from the true reference dSCDs is more complex as more detailed structures in the 2D-patterns occur. Moreover, the findings depend strongly on the sun geometry. Nevertheless, for all trace gas scenarios and viewing geometries, intervals shorter than around 20 nm lead to large deviations (mostly strong overestimations) and are thus no good options[2]. In addition, these intervals showed high variabilities of the retrieved HCHO dSCDs and large fit errors for synthetic and measured data. Besides this, mainly overestimations of the HCHO dSCDs (intervals with lower boundaries between ca. 323 and 335 nm and upper boundaries between about 345 and 360 nm) in the order of ca. 2×10^{15} to $9 \times$

2 Some exceptions exist, e.g. intervals with lower boundaries in the range of 320 nm.

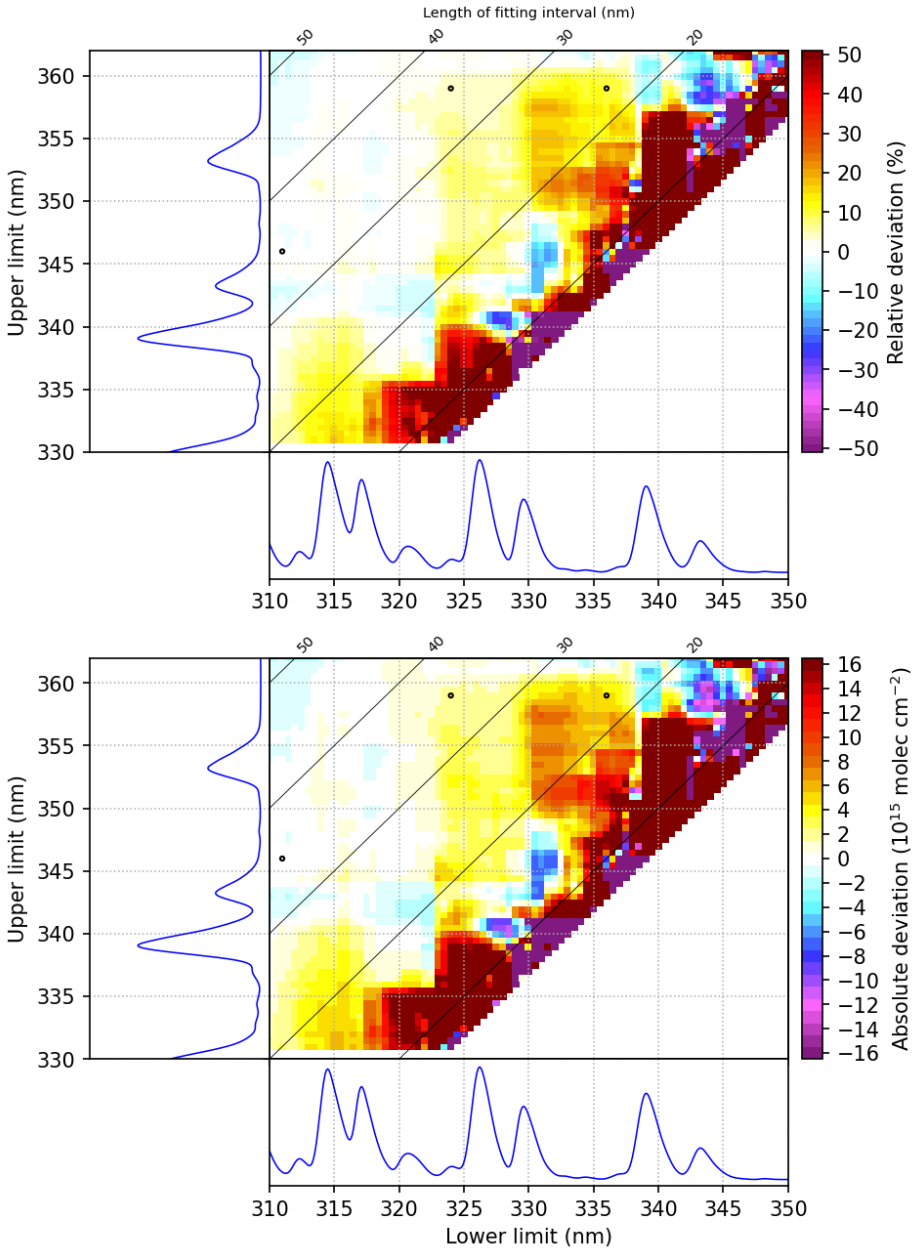


Figure 6.26: Relative (upper panel) and absolute (lower panel) deviations of the retrieved HCHO dSCDs from the true reference values at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as functions of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes.

10^{15} molec cm^{-2} and rather good agreements between retrievals and simulations with absolute deviations below 2×10^{15} molec cm^{-2} (lower boundaries between ca. 310 and 323 nm and upper boundaries between about 340 and 355 nm) are found for the representative example depicted in Fig. 6.26. In some smaller spectral

regions also slight underestimations with absolute deviations of approximately -2×10^{15} to -4×10^{15} molec cm^{-2} occur.

These general 2D-patterns of the (absolute) deviations from the reference dSCDs (lower panel) are similar for the other, not presented, trace gas scenarios and geometries, while the details in the 2D-patters can have (strong) dependencies. For instance, for some sun geometries, the detailed patterns and absolute values of the deviations show a pronounced dependence on the elevation angle. The strength of this elevation dependence differs between the different scenarios as well as between different spectral regions of the tested fitting intervals for a given sequence (defined as a set of trace gas scenario and sun geometry). More precisely, the latter means that for a given sequence some fitting intervals show a stronger elevation dependence of the deviations from the true reference dSCDs than others. Consistent to large parts of the here-presented sensitivity study, the (detailed) 2D-patterns of the deviations and their absolute values are almost independent of the HCHO scenario, while both depend slightly on the NO_2 scenario, whereby the general patterns are not largely affected.

Altogether, a large set of tested fitting intervals performs well for some scenarios but bad for others. Similar to the investigations on the NO_2 and O_3 cross-sections in the previous section, the difficulty is to get an overview of the results that goes beyond the very basic considerations outlined above since information from 144 plots has to be compiled. Hence, again the RMS of the absolute deviations from the true reference dSCDs (RMS_{abs}) and the corresponding standard deviations within one sequence (defined as a specific combination of trace gas scenario and sun geometry) were calculated for each tested fitting interval for all individual sequences using Eq. 6.4. It should be noted that here the standard deviations are calculated from the absolute deviations obtained for the four elevation angles within a given sequence and not over all simulated scenarios (trace gas scenarios and sun geometries) like done for the considerations on the NO_2 and O_3 cross-sections. Therefore, the standard deviations here are a measure for the elevation dependence of the absolute deviations within a sequence. Plotting these two quantities as functions of the boundaries of the tested intervals like above allows to characterise the results for one complete sequence with only one plot instead of four. The resulting plots of RMS_{abs} (upper panel) and the corresponding standard deviations (lower panel) for the representative scenario are presented in Fig. 6.27.

Most structures in the 2D-patterns of the absolute deviations presented in the lower panel of Fig. 6.26 are also represented by the RMS_{abs}-values in the upper panel of Fig. 6.27. Further, intervals leading to large deviations and thus large RMS_{abs}-values are associated with large standard deviations indicating strong elevation dependencies of the absolute deviations within the sequence, e.g. intervals with lower boundaries between ca. 330 and 338 nm and upper boundaries between about 350 and 360 nm. Another observation for the representative scenario is that overall long fitting intervals yield small RMS_{abs}-values indicating rather good agreements between DOAS retrievals and simulations. However, also here more detailed (weak) structures exist. For instance, local maxima of RMS_{abs} and the corresponding standard deviations are found for several intervals starting at wavelengths below around 317 nm.

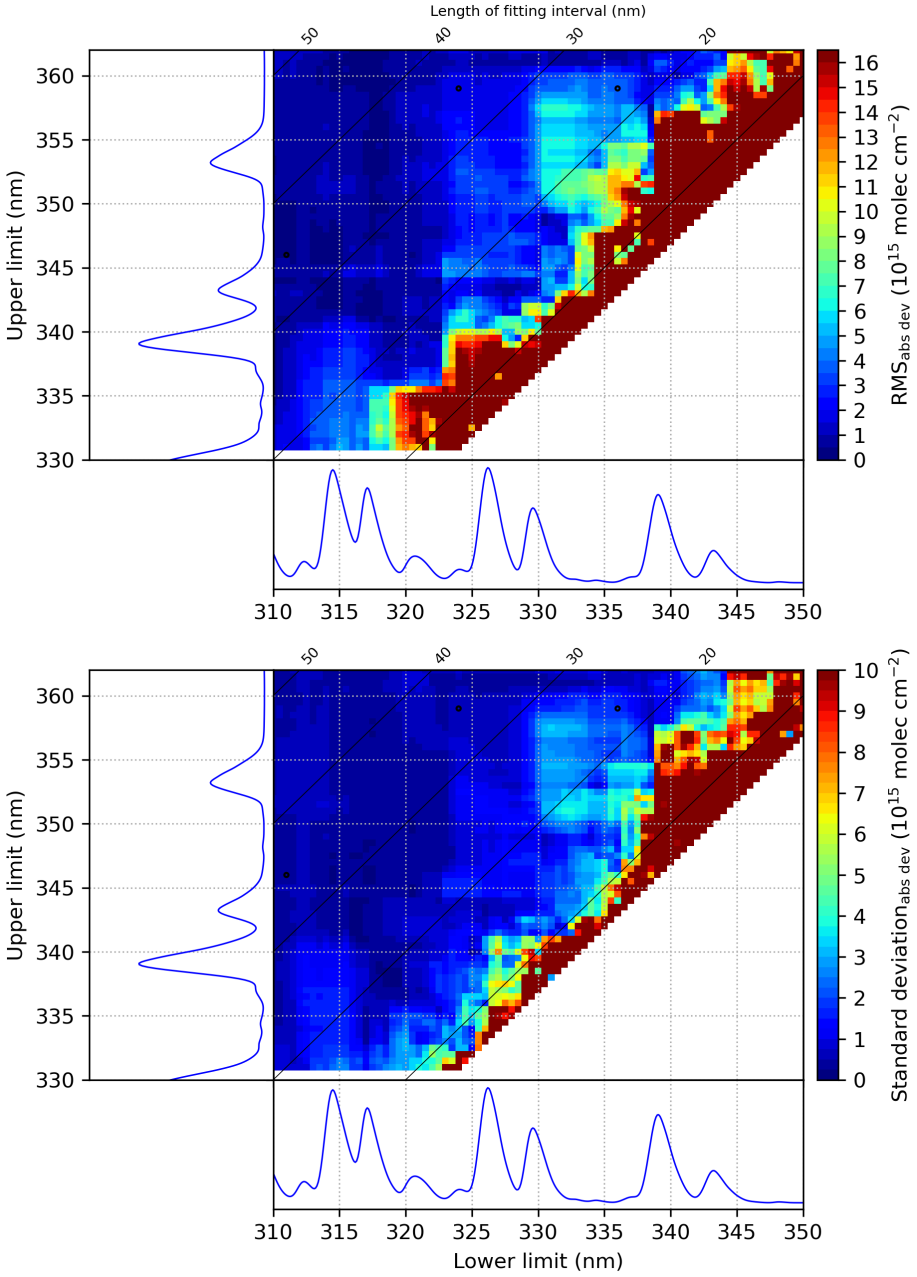


Figure 6.27: RMS of the absolute deviations of the retrieved HCHO dSCDs from the true reference dSCDs (upper panel) and corresponding standard deviations (lower panel) within the sequence for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as functions of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes.

Again, the general 2D-patterns as well as the absolute values of both quantities are almost independent of the HCHO scenario. Also the NO_2 scenario has only

a small influence on the 2D-patterns of RMS_{abs} and its absolute values for a given sun geometry. The effect of the NO_2 scenario on the standard deviations, i.e. elevation dependence within a sequence, is larger. Contrary to that, larger differences in the 2D-patterns and the absolute values of RMS_{abs} are observed for different sun geometries. Nevertheless, the overall patters are still only slightly affected. For instance, intervals with lower boundaries between ca. 330 and 338 nm and upper boundaries between about 350 and 360 nm consistently lead to large RMS_{abs}-values and, for many sun geometries, also to large standard deviations. Sun geometries for which large deviations from the reference dSCDs were found during the investigations regarding the NO_2 and O_3 cross-sections (compare Sect. 6.2.2.1 and 6.2.3.1), e.g. the scenarios with an SZA of 60° and an RAA of 30°, yield comparatively high RMS_{abs}-values, and thus large absolute deviations, in large parts of the spectral regions in the presented 2D-plots. Interestingly, such sun geometries are associated with relatively small standard deviations for many fitting ranges indicating only small elevation dependencies.

Based on the findings for the RMS_{abs} and the corresponding standard deviations for the individual scenarios (e.g. 2D-patterns depend only weakly on the trace gas scenario), the mean of the RMS_{abs} over all trace gas scenarios and sun geometries was calculated for each tested fitting interval. In addition, the standard deviations of the RMS_{abs} over all individual trace gas scenarios and sun geometries were calculated for all of these intervals. Note that here the standard deviations characterise the variability of RMS_{abs} amongst the different scenarios of synthetic spectra. Figure 6.28 shows the resulting mean RMS_{abs} (upper panel) and the corresponding standard deviations (lower panel) as functions of the boundaries of the tested fitting intervals. As described in detail below, red and green polygons mark spectral regions in which minima of both quantities are found.

As expected, the general 2D-patterns of the mean RMS_{abs} are similar to ones of the absolute deviations obtained for the individual scenarios and their corresponding RMS (within a sequence). Nonetheless, minima in the 2D-distributions of the mean RMS_{abs} and its standard deviations are found for fitting intervals with lower boundaries in the range of approximately 315 to 323 nm and upper boundaries between roughly 340 and 357 nm. However, many, yet not all, of these intervals have boundaries directly at strong HCHO absorption bands which can lead to problematic fit results for real measured spectra. In particular, fitting intervals with boundaries in this spectral region led to relatively large variabilities of the retrieved HCHO dSCDs for the measured spectra (compare Fig. 6.25). The observation that this effect occurs mainly for the measured spectra can likely be explained by an imperfect representation of the true atmospheric HCHO absorptions or other absorptions by the fitted reference absorption cross-sections. Such small imperfections can have a strong effect for cases were fitting intervals start and/or end directly at major absorption bands. Since the fitted reference absorption cross-sections were also used for the simulations of the synthetic spectra, this effect does not play a role for the synthetic data.

Including also information from the RMS_{abs}-values of the individual scenarios, a spectral region was marked with a red polygon in Fig. 6.28 in which

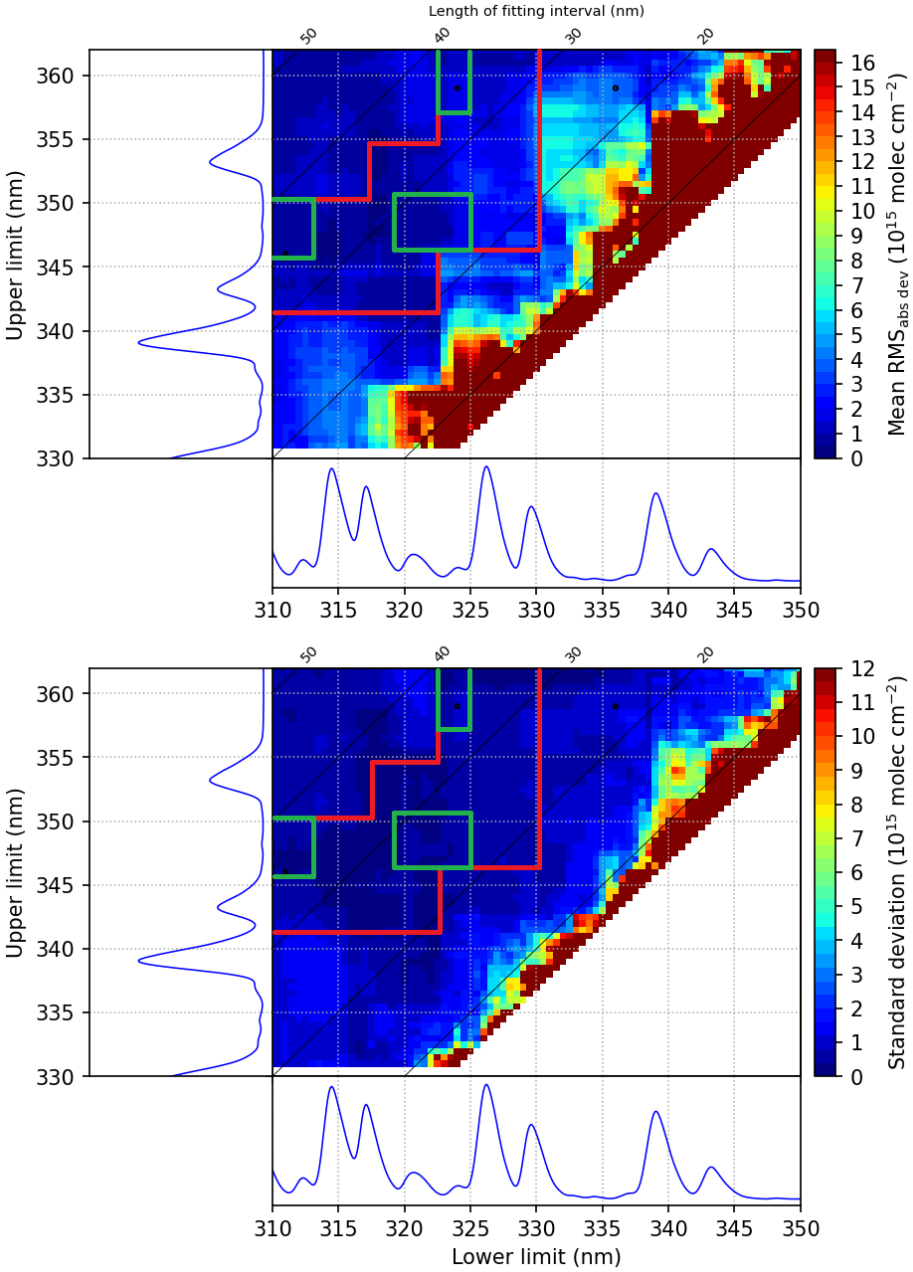


Figure 6.28: Mean RMS_{abs} (upper panel) and corresponding standard deviations (lower panel) calculated over all scenarios as functions of the lower and upper boundaries of the tested fitting intervals. The lengths of these intervals are represented by the skewed black lines with the corresponding values on the upper x-axis. Commonly used fitting intervals are marked by black circles. Additionally, the normalised HCHO absorption cross-section is shown on the x- and y-axes. Red and green polygons mark spectral regions in which minima of both quantities are found (for details refer to the text).

both RMS_{abs}-values and standard deviations (of the absolute deviations within one sequence) are small for many scenarios. Also, the mean RMS_{abs} and the standard deviations (amongst the individual RMS_{abs}) show low values for fitting

intervals with boundaries in the marked spectral region. It should be noted that also comparably long intervals (top left corners of 2D-plots) show low mean RMS_{abs}-values. However, for some sun geometries such intervals lead to large RMS_{abs}-values and thus large deviations from the true reference dSCDs probably caused by strong O_3 absorptions at such short wavelengths. This observation is weakly represented by faintly increased standard deviations in this spectral region (compare lower panel of Fig. 6.28). Contrary to that, no particularly large deviations from the true reference dSCDs are found for neither scenario for fitting intervals with boundaries in the marked spectral region.

If additionally intervals which start and/or end directly at major HCHO absorption bands and intervals that perform less well for real measured spectra are excluded, fitting ranges with upper and lower boundaries in three rather confined spectral regions remain which are marked by green rectangles in Fig. 6.28. Taking also the considerations on the fit RMS, the fit error, and the variability of the retrieved HCHO for both real and synthetic data from above into account, only the uppermost green rectangle which also includes the standard fitting interval (black circle in green uppermost rectangle) is left. Intervals with boundaries in this spectral region are not the best ones in neither aspect but lead to slight overestimations of the retrieved HCHO dSCDs for almost all scenarios (e.g. Fig. 6.26). Nevertheless, the results obtained with these fitting intervals are rather stable between the different trace gas scenarios and sun geometries indicated by overall small standard deviations. Also for the other investigated quantities (fit RMS, fit error and variability of retrieved dSCDs) such intervals are neither the best nor the worst, whereby especially for the variability of the retrieved dSCDs obtained for the measured spectra a minimum is found in this spectral region (compare Fig. 6.25).

Summarising the considerations in this section so far, fitting intervals with lower boundaries between ca. 322 and 325 nm and upper boundaries from ca. 357 to 362 nm are apparently reasonable choices.

One possible improvement in terms of agreement between retrievals and simulations compared to the standard fitting interval (324 to 359 nm) could be the extension of the lower interval boundary by 1 nm yielding an interval from 323 to 359 nm which leads to a faintly smaller mean RMS_{abs} (compare uppermost green rectangle in the upper panel of Fig. 6.28) and also the standard deviation is slightly smaller. However, this improvement is very small which becomes also evident by taking additionally the absolute deviations and the corresponding RMS_{abs}-values of the individual scenarios into account (not shown). Also the choice of the colour bar increments might contribute here. Another aspect that should be kept in mind is that the fit RMS increases strongly for intervals with lower boundaries below 324 nm especially for the measured spectra (compare Fig. 6.20 and 6.21). Besides this and the standard fitting range, also other intervals deserve specific attention.

For instance, fitting intervals in the range of ca. 319.5 to 350 nm (or similar ones) directly lie in the spectral region that shows minimum RMS_{abs}-values (compare centre green rectangle in Fig. 6.28). However, such intervals are about 4.5 nm shorter than the standard fitting range leading to larger fit errors and higher

variabilities of the retrieved HCHO dSCDs. In addition, also for the measured spectra most of such intervals show enhanced fluctuations in the resulting HCHO dSCDs (compare Fig. 6.25).

Similar intervals but with upper boundaries at longer wavelengths, e.g. 319 to 360 nm (outside red and green polygons), could also be considered. Ignoring the fit RMS which is especially for the real measured spectra significantly larger compared to the standard interval, this interval yields overall similar results as the standard fitting range and performs even better in terms of fit errors and variabilities of the results for the synthetic data. However, for some individual sun geometries, such intervals lead to notably larger RMS_{abs}-values indicating larger absolute deviations. In general, also for the measured spectra, such fitting intervals perform reasonable but are closer to the spectral region in which the variabilities of the retrieved HCHO dSCDs increase strongly (compare Fig. 6.25). In addition, the fit RMS is markedly larger for the real data using such intervals.

The results for the synthetic spectra also suggest that some intervals with lower boundaries at short wavelengths might be reasonable choices as well, e.g. an interval from 311 to 346 nm (compare left green rectangle in Fig. 6.28). All investigated quantities show good results for such intervals which are similar to the ones of the standard interval, whereby the variability of the HCHO results is even smaller. Also for the real data, most quantities show in general similar results for the standard interval and the one from 311 to 346 nm. However, the fit RMS for the latter fitting range is clearly larger compared to the standard interval (compare Fig. 6.21) and at 30° the variability is largely increased using the 311 to 346 nm interval for the measured spectra. These two findings can be explained by strong atmospheric O_3 absorptions and worse detector properties towards the edges of the CCD-chip as outlined in more detail above.

The commonly used shorter 336 to 359 nm fitting interval lies in a spectral region which shows substantially larger mean RMS_{abs}-values (compare upper panel in Fig. 6.28) than the previously discussed intervals indicating larger deviations from the true reference dSCDs. Also, the standard deviation of the absolute deviations within a given sequence of synthetic spectra is systematically larger for this interval (compare lower panel in Fig. 6.27), i.e. the absolute deviations show a strong elevation angle dependence. In addition, the fit error and the variability of the retrieved HCHO dSCDs are relatively high for this interval caused by its rather short length (compare Fig. 6.22 and 6.24). For the measured spectra, this interval is close to the spectral region in which the variabilities of the HCHO results increase strongly (compare Fig. 6.25). Despite these drawbacks, the 336 to 359 nm fitting interval is a reasonable choice for intervals with lower boundaries above approximately 330 nm and the usage of such fitting ranges can be justified if a fixed FRS is used for the DOAS analyses. In such cases, the usage of fitting ranges with lower boundaries at rather short wavelengths leads to strong influences of the strong O_3 absorptions and shorter fitting intervals with lower boundaries at longer wavelengths yield better results. For such applications, the 336 to 359 nm fitting range is still recommended here.

For passive DOAS applications using a sequential FRS, the advantages of longer fitting intervals are obvious. The final selection based on this sensitivity

study depends on how much weight is put on the different discussed aspects. In this work, a compromise is made between all of these different aspects and the standard fitting interval from 324 to 359 nm is still recommended since it is a reasonable choice in many aspects and yields overall very stable results. In Sect. 6.7, an estimation of the uncertainties which are related to that choice is presented.

6.4 REMAINING DEVIATIONS

Despite the optimisation efforts for the main DOAS fit settings presented in the previous sections, no perfect agreement between retrievals and simulations is achieved for the synthetic spectra. Only for few scenarios, the remaining deviations of the retrieved from the true reference HCHO dSCDs are very small and rather close to 0, e.g. in the case of high HCHO and NO_2 at an SZA of 20° and an RAA of 90° an RMS_{abs} of ca. 0.36×10^{15} molec cm^{-2} is obtained. For the larger fraction of the scenarios, however, the deviations which remain after the optimisations of the fit settings are significant. In particular, the mean RMS_{abs} over all scenarios is around 1.75×10^{15} molec cm^{-2} for the selected optimised combination of O_3 cross-sections which, in addition to the standard O_3 cross-sections, includes the linear and square Puķīte terms (compare "Std + both higher" in Fig. 6.17). For individual geometries, the deviations obtained with the optimised fit settings[3] can even reach up to 3.7×10^{15} molec cm^{-2} in the case of high HCHO, low NO_2 at an SZA of 60° and an RAA of 30°. Hence, this section provides an overview on several aspects that might contribute to these remaining deviations between retrievals and simulations or cause uncertainties leading to the observed deviations. While in the previous sections synthetic spectra with noise have been used to obtain results that are representative for real measurements, for the investigations outlined in this section, if not indicated differently, synthetic spectra without noise were used to exclude its effects as possible reason for the remaining deviations.

6.4.1 *Ring effect and noise*

As outlined in Sect. 3.1.5.1, inelastic Raman scattering causes the Ring effect which is accounted for by including a solar Ring spectrum (or even two Ring spectra) as pseudo absorber in the DOAS fit. Rotational Raman scattering was also considered for the simulations of the synthetic spectra and thus also the standard fit settings used for the sensitivity study include two such Ring spectra (compare Table 6.3). Different methods to correct for the Ring effect exist (compare Sect. 3.1.5.1). Here, the Ring spectra are calculated for the 90° spectrum for the case of low HCHO and NO_2, an SZA of 45° and an RAA of 90°. However, no radiative transfer effects besides the scattering angle are considered for this calculation and thus these Ring spectra can be regarded as approximation. Despite this limitation, this way of correcting for the Ring effect works typically astonishingly well. Nevertheless, an

3 These analyses were already performed using the finally selected fitting interval and combinations of NO_2 and O_3 cross-sections.

imperfect correction of the Ring effect in the DOAS retrieval leading to remaining influences of the Ring effect might be one possible reason for the remaining deviations between retrievals and simulations.

To investigate this potential effect, synthetic spectra which do not include Raman scattering and thus no Ring effect were simulated using the settings summarised in Sect. 6.1.1 and analysed using the optimised standard fit settings from above. The resulting HCHO dSCDs were then compared with the ones obtained for the standard spectra which include the Ring effect (and noise). This was done for the same 36 scenarios which were already used for the previous studies. It should again be stressed that contrary to the previous investigations, spectra without noise were used here to exclude the noise as potential reason for the remaining deviations. Further, this allows to also study the effect of noise on these remaining deviations.

Figure 6.29 shows the resulting HCHO dSCDs (uppermost panel), the corresponding absolute deviations from the true reference dSCDs (second panel), the effects of the different scenarios on the HCHO dSCDs (third panel) and the fit RMS (lowermost panel) as functions of elevation for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. The effect on the dSCDs (third panel) is defined as the difference between the HCHO dSCDs retrieved with the spectra without Ring effect and/or without noise and the ones obtained for the standard spectra[4]. The results obtained for the standard spectra containing noise and Ring effect, the spectra without noise but still including the Ring effect and the spectra without noise and Ring effect are depicted in different colours as indicated in the legend above the figure.

This specific sequence of synthetic spectra was chosen since large deviations from the reference dSCDs of up to roughly 3.7×10^{15} molec cm^{-2} at 10° elevation remain after the optimisations of the main DOAS fit settings for this combination of trace gas scenario and sun geometry[5]. The results obtained for the spectra without noise but still including the Ring effect (green curves in Fig. 6.29) suggest that parts of the remaining deviations can be attributed to the noise included in the standard spectra. The absolute effect on the HCHO dSCD at 10° elevation is a reduction by 6×10^{14} molec cm^{-2} for the spectra without noise (compare third panel) leading to a remaining deviation of ca. 3.1×10^{15} molec cm^{-2}. The latter is further reduced for the spectra without Ring effect yielding a remaining deviation of 2.66×10^{15} molec cm^{-2} (red curve in the second panel) which indicates an effect of approximately 4.4×10^{14} molec cm^{-2} caused solely by the Ring effect (dashed red line in the third panel) for this example scenario. Together, the deviations are reduced by 1.04×10^{15} molec cm^{-2}. Nevertheless, the remaining deviation at 10° is still ca. 70 % of the one found for the noisy spectra including the Ring effect and the HCHO dSCD is only reduced by 3.6 %. For the other elevations of the presented sequence, similar findings are true (only at 1° the effect of the noise is smaller). Naturally, the fit RMS is largely reduced for the noise-free spectra and becomes even smaller for the spectra without Ring effect.

4 In the following, this quantity describes always the difference between dSCDs retrieved with modifications compared to the standard spectra and/or settings.

5 A detailed look shows that these deviations are only marginally smaller than for the scenario which shows overall the largest remaining deviations (compare above).

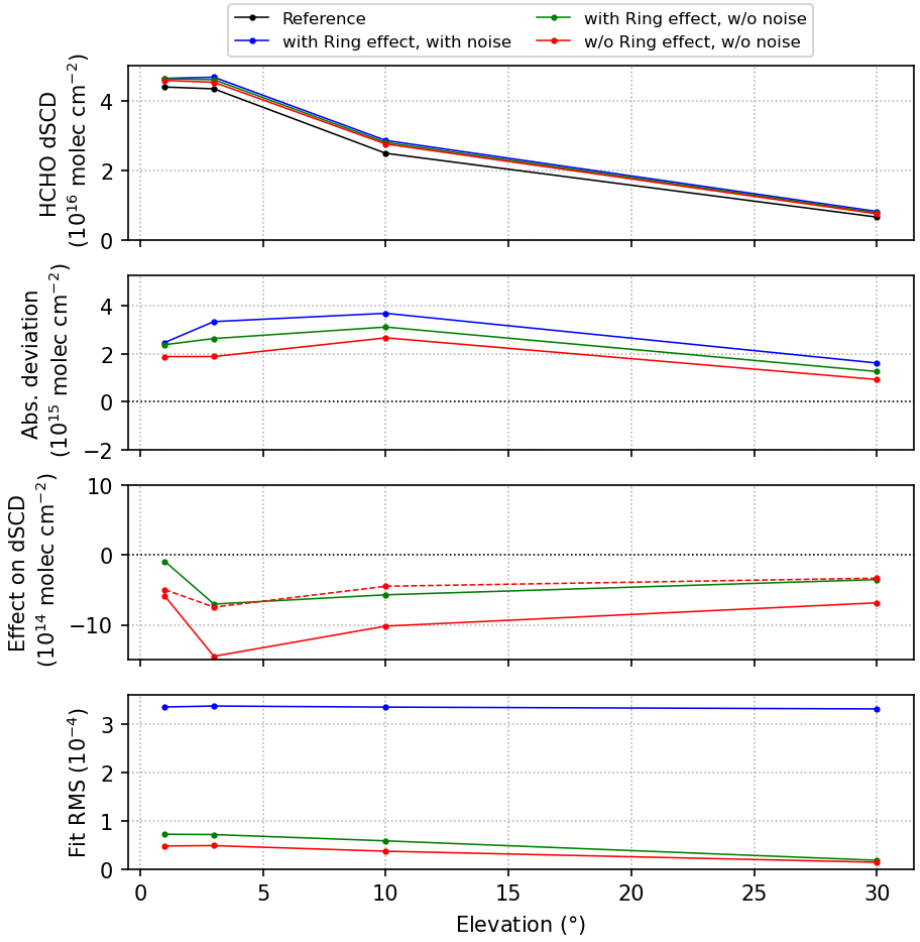


Figure 6.29: Mean retrieved HCHO dSCDs, absolute deviations from the true reference values, effects of the different scenarios on the HCHO dSCDs and fit RMS as functions of elevation for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. Results for spectra with and without Ring effect as well as noise are indicated by different colours as described in the text.

Altogether, parts of the remaining deviations can be explained by the noise and the Ring effect for the presented example, whereby the absolute influences on the HCHO results are similar for both effects. For some of the other (not presented) scenarios, however, the findings can be different and more subtle. Some aspects of this are outlined below.

For a given NO_2 scenario and sun geometry, the effect of the noise is independent of the HCHO scenario, whereas a small dependence on the HCHO abundance is found for the influence of the Ring effect for which the impact at the two lowest elevations is larger for the high HCHO scenario. In general, the effects of both noise and Ring effect are slightly larger for the low NO_2 scenario of most sun geometries.

Besides detailed dependencies for a few geometries, the individual effects of noise and the Ring effect are comparable and have similar magnitudes for almost all sun geometries, while the absolute values of the effects on the HCHO dSCDs differ. The spectra without noise but including the Ring effect lead for almost

all geometries (elevation angle and sun geometry) to smaller HCHO dSCDs compared to the standard spectra with the absolute differences ranging from almost 0 to around -6×10^{14} molec cm^{-2}.

The influence of the Ring effect is more complex and leads for most viewing geometries (like for the presented example) to further reductions of the HCHO dSCDs. However, also enhancements which are for some geometries partly compensated by the effect of the noise can occur, e.g. for the scenarios with an SZA of 60° and an RAA of 30°. For the latter geometry, the spectra without noise and Ring effect lead to larger deviations compared to the standard spectra. This effect is particularly strong for the low elevation angles in the case of high NO_2. These findings for the Ring effect indicate complex compensation effects and interferences which are further discussed at the end of this section. The absolute effects of the Ring effect range between $\pm 7 \times 10^{14}$ molec cm^{-2} for the different scenarios ignoring the odd scenario from before. Nevertheless, the influences of both effects on the HCHO results are almost two orders of magnitude smaller than the dSCDs (at low elevations) themselves.

The finding that for most sequences slightly lower deviations from the true reference HCHO dSCDs are obtained for the spectra without noise and Ring effect leads for some sun geometries to almost perfect agreements between retrievals and simulations. This applies mainly to scenarios which already showed relatively small deviations for the standard spectra, e.g. spectra at an SZA of 20° and an RAA of 30°. However, also slight underestimations of the reference dSCDs occur for some elevations of such sequences (e.g. for spectra at an SZA of 20° and an RAA of 180°).

The fact that the influence of the Ring effect differs notably between the sun geometries shows that the strength of the Ring effect is closely related to the viewing geometry of a measurement. This becomes also evident by the finding that for several sun geometries the elevation angle dependence of the absolute deviations from the reference dSCD is markedly reduced for the spectra without Ring effect. Nevertheless, still for many scenarios the largest deviations are found at 10° elevation and the Ring effect is apparently not the only cause for this (compare Fig. 6.29).

Altogether, the findings in this section suggest that there is no fundamental problem with the Ring effect and the way it is accounted for in the DOAS analyses in this work. Nevertheless, parts of the deviations which remain after the optimisations of the main fit settings for the standard spectra can be attributed to the Ring effect and the noise included in the synthetic spectra. Thereby, the absolute influences of both effects are comparable for most scenarios. The actual combined effects and behaviours differ amongst the different scenarios, whereby for scenarios which showed relatively small remaining deviations already for the standard spectra only very small deviations remain for the spectra without noise and Ring effect. However, especially for scenarios with large remaining deviations for the standard spectra, both effects cannot entirely explain the remaining deviations between retrievals and simulations (e.g. sequence in Fig. 6.29). Thus, other effects must contribute significantly to the remaining deviations.

6.4.2 *Intensity weighted convolution for other absorbers*

As outlined in Sect. 3.1.5.3, besides other dependencies, the strength of the solar I_0-effect for a given absorber depends on its absorption strength and is therefore particularly important for strong absorbers like NO_2 and O_3. Therefore, it is typically neglected for weak absorbers. In theory, however, the solar I_0-effect occurs for any absorber. Since in the optimised standard fit settings (compare Table 6.3 + previous optimisations) this effect is only considered for the NO_2 and O_3 cross-sections and the corresponding Puḳīte terms, at least parts of the remaining deviations between retrieved and true reference HCHO dSCDs might be caused by missing I_0-corrections for the other absorbers included in the fit. As outlined in Sect. 3.1.5.3 and 6.2.1, instead of a classical I_0-correction (compare Eq. 3.21) an intensity weighted convolution (compare Eq. 3.25) is used in this work which fulfils the main aim of the I_0-correction as it weights the absorptions at the individual wavelengths with the corresponding intensities of I_0.

To study the effects of missing I_0-corrections for the other (weak) absorbers included in the DOAS fit for HCHO besides NO_2 and O_3, namely HCHO, BrO and O_4, the absorption cross-sections of these trace gases were convolved using intensity weighted convolutions and included in the fit instead of the cross-sections convolved with standard convolutions. Then, the HCHO dSCDs obtained with these intensity weighted cross-sections were compared with the ones obtained using the optimised standard fit settings. Here, both individual and combined effects were studied for the same 36 previously used scenarios.

The resulting HCHO dSCDs (uppermost panel), absolute deviations from the true reference dSCDs (second panel), the effects of the intensity weighted convolutions on the HCHO dSCDs (third panel) and the fit RMS (lowermost panel) as functions of elevation for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180° are presented in Fig. 6.30. The results obtained with the optimised standard settings and intensity weighted convolutions for HCHO, BrO and O_4 are shown in different colours as indicated by the legend above the figure. Again, the results for this specific scenario are presented as an example since it shows rather large remaining deviations for the optimised fit settings, reaching up to 3.1×10^{15} molec cm^{-2} at 10° elevation, even for spectra without noise (compare previous section).

Obviously, the usage of an intensity weighted convolution for the BrO cross-section has no significant effect on the retrieved HCHO dSCDs (absolute effect below 2×10^{13} molec cm^{-2}) and does not reduce the remaining deviations (blue curves in the second and third panels). This is true for all investigated scenarios, whereby the absolute effects vary slightly around 0.

For the HCHO and O_4 cross-sections, however, the usage of an intensity weighted convolution has an influence on the retrieved HCHO dSCDs. As expected from theory, this effect is directly related to the absorption strength and is thus strongest pronounced at low elevations, where the retrieved HCHO dSCDs are approximately 3.4×10^{14} molec cm^{-2} smaller compared to the optimised standard settings when intensity weighted convolutions are used for either of the two trace gases. With increasing elevation the effects become smaller and lead

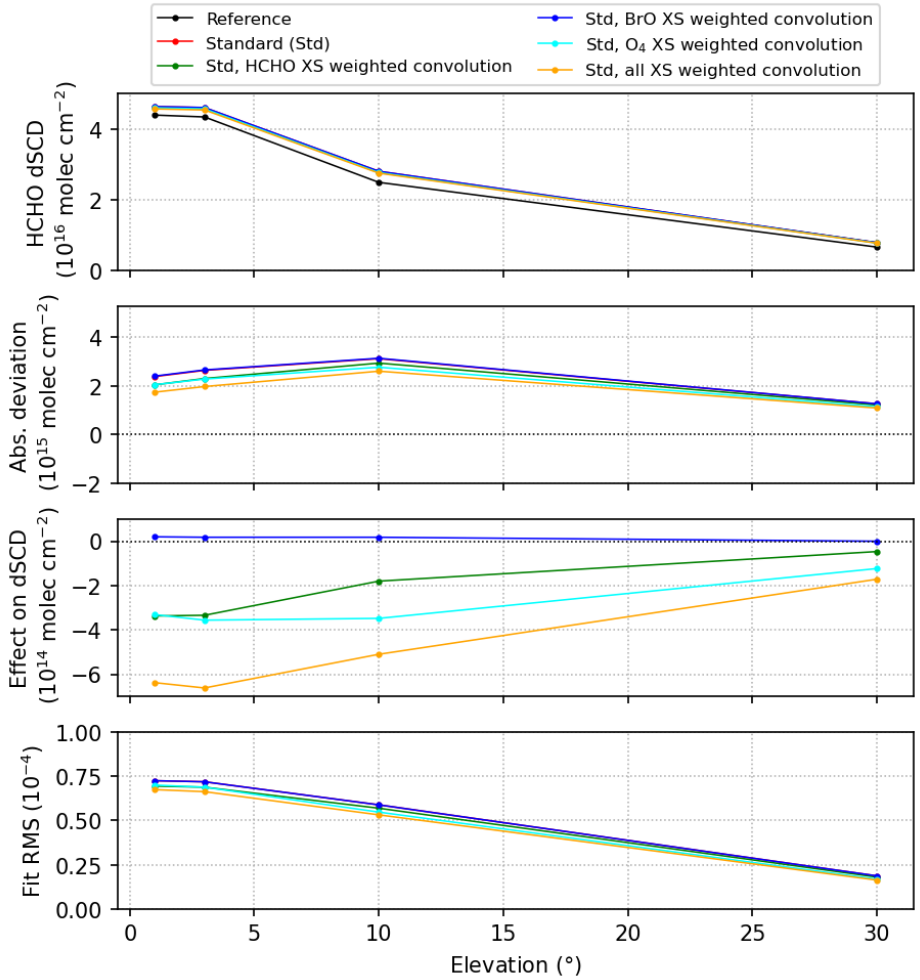


Figure 6.30: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effects of the intensity weighted convolutions on the HCHO dSCDs and fit RMS as functions of elevation for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. Results obtained with the optimised standard settings and different sets of cross-sections included in the DOAS fit are shown in different colours.

to ca. 1.2×10^{14} molec cm^{-2} and 0.5×10^{14} molec cm^{-2} smaller HCHO dSCDs at 30° if intensity weighted convolutions are applied for the O_4 and HCHO cross-section, respectively. The combined effect of intensity weighted convolutions for the cross-sections of HCHO, BrO and O_4 is basically the sum of the individual effects being ca. -6.5×10^{14} molec cm^{-2} at low elevations and decreases to around -1.7×10^{14} molec cm^{-2} at 30° elevation. Consequently, at low elevations the remaining deviations can be reduced by ca. 25 %, whereby the effect on the HCHO dSCDs themselves is only a few percent. For the presented example, at low elevations these values are of the same order as the influences of the Ring effect and the noise on the HCHO dSCDs. Besides the effects on the dSCDs, small effects on the fit RMS, which is slightly reduced, are found for the HCHO and O_4 cross-sections.

For a given sun geometry and NO_2 scenario, the influence of the intensity weighted convolution for the O_4 cross-section does almost not depend on the HCHO scenario. As the effect of the I_0-correction is closely linked to the absorption strength, for the HCHO cross-section, the effect on the dSCDs is notably smaller for the low HCHO scenarios for which the effect is at maximum 1×10^{14} molec cm^{-2} at low elevations. This is roughly three times smaller than for the high HCHO scenarios which is consistent with a three times smaller HCHO VCD. Thus, compared to the effect of the O_4 cross-section, in low HCHO cases the individual effect for the HCHO absorption cross-section is markedly smaller.

The overall patterns and absolute values of the effects of the intensity weighted convolutions for the cross-sections of HCHO and O_4 on the HCHO results do not strongly depend on the NO_2 scenario. Nevertheless, for O_4, at low elevations the effect is consistently larger for the low NO_2 scenarios probably caused by slightly longer average light paths for such scenarios.

Further, the general effects on the HCHO results are similar for different sun geometries and also the absolute changes of the dSCDs show only small variations. Generally, compared to the HCHO cross-section, the differences between the different sun geometries are larger for the effect of the intensity weighted convolution of the O_4 cross-section. These different findings for the O_4 and HCHO cross-sections can be explained by the fact that the O_4 absorptions are closely related to the average length of the atmospheric light path which is strongly influenced by the measurement geometry. However, the differences of the effects on the HCHO results between the different sun geometries is at maximum only 1.5×10^{14} molec cm^{-2}.

To summarise, for all investigated scenarios, the intensity weighted convolutions of the HCHO and O_4 cross-sections and thus also the combined effect[6] lead to systematic reductions of the retrieved HCHO dSCDs, whereby the effects are clearly stronger at low elevations. These reductions show relatively small variations between the different scenarios and only the effect of an intensity weighted convolution of the HCHO cross-section has a significant dependence on the HCHO scenario. Together, this leads to slightly better agreements for scenarios which showed substantial deviations from the true reference HCHO dSCDs after the optimisation of the fit settings (compare example in Fig. 6.30). However, for scenarios which showed only small over- or underestimations and thus small absolute deviations from the true reference dSCDs (e.g. spectra with an SZA of 20° and an RAA of 30° in the case of high NO_2), this systematic reduction of the retrieved HCHO dSCDs leads to underestimations. Therefore, the agreements between retrievals and simulations are worsened in such cases. The latter was different for the influence of the Ring effect where changes of the HCHO dSCDs were observed in both directions leading at least for some of such scenarios to improved agreements. Further, the effects of intensity weighted convolutions for the different cross-sections are largest at low elevations and the deviations at 10° and 30° are only slightly reduced. Hence, also the solar I_0-effect which is neglected for weak absorbers in the standard fit settings cannot explain the rather large remaining deviations between retrievals and simulations.

6 The effect of an intensity weighted convolution of the BrO cross-section can be neglected.

6.4.3 *Further higher order terms*

So far, Puķīte terms of NO_2 and O_3 accounting for first-order dependencies of the SCDs on absorption and scattering, namely the linear and square Puķīte terms, were introduced (compare Sect. 3.1.5.2 and Sect. 6.2) and their effects on the DOAS retrievals for HCHO tested (compare Sect. 6.2.2 and 6.2.3). For both absorbers, the combined usage of both terms led to improved agreements between retrievals and simulations. While for NO_2 the improvements mainly occurred for the high NO_2 scenarios and at low elevations, the agreements were mainly improved at 10° and 30° elevation for O_3. Thus, further Puķīte terms of NO_2 and O_3 accounting for additional higher order effects might potentially reduce the rather large remaining deviations between retrieved and true reference HCHO dSCDs. Therefore, the following additional Puķīte terms were tested:

- The cross-correlation term of NO_2 and O_3 absorptions which is referred to as $\sigma_{NO_2}\sigma_{O_3}$ in the following.
- Higher (second) order Puķīte terms for NO_2 and O_3: $\lambda^2\sigma$, $\lambda\sigma^2$ and $(\lambda\sigma)^2$.

It should be noted that in principle also further higher order terms could have been tested, however, here a selection based on expectations which terms might be reasonable was made. For further details on such higher order Puķīte terms, refer to Puķīte et al. (2010) and Puķīte and Wagner (2016).

Only the low aerosol scenarios (compare Table 6.1) were considered for the studies presented so far. For these scenarios in combination with low HCHO abundance, the O_4 absorptions can be roughly twice as high as the HCHO absorptions in terms of optical depths. Thus, also the usage of first-order Puķīte terms for O_4 might be justified and is thus also tested in this section.

To study the individual effects of the outlined higher order terms on the HCHO retrieval, they are separately included in the DOAS fit which was, besides that, performed using the optimised standard settings. Thereby, for one specific higher order Puķīte term, e.g. $\lambda^2\sigma$, the corresponding terms for NO_2 and O_3 were tested simultaneously to reduce the amount of fit scenarios.

Figure 6.31 depicts the resulting HCHO dSCDs (uppermost panel), the corresponding absolute deviations from the true reference dSCDs (second panel), the effects of the different additional Puķīte terms for NO_2, O_3 and O_4 on the HCHO dSCDs (third panel) and the fit RMS (lowermost panel) as functions of elevation. Like in the previous two sections, results for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°, which shows large remaining deviations after the optimisation of the DOAS fit settings, are shown. The results obtained with the different combinations of cross-sections included in the fit are shown in different colours as indicated in the legend above the figure.

The inclusion of the linear or square Puķīte term of O_4 as well as the combination of both in the DOAS retrieval of HCHO does not significantly affect the resulting HCHO dSCDs (corresponding lines are almost identical with the red curve in all panels). For the presented example, the effects on the dSCDs are below 2.5×10^{13} molec cm^{-2} and the fit RMS is reduced by less than 1 %. In fact,

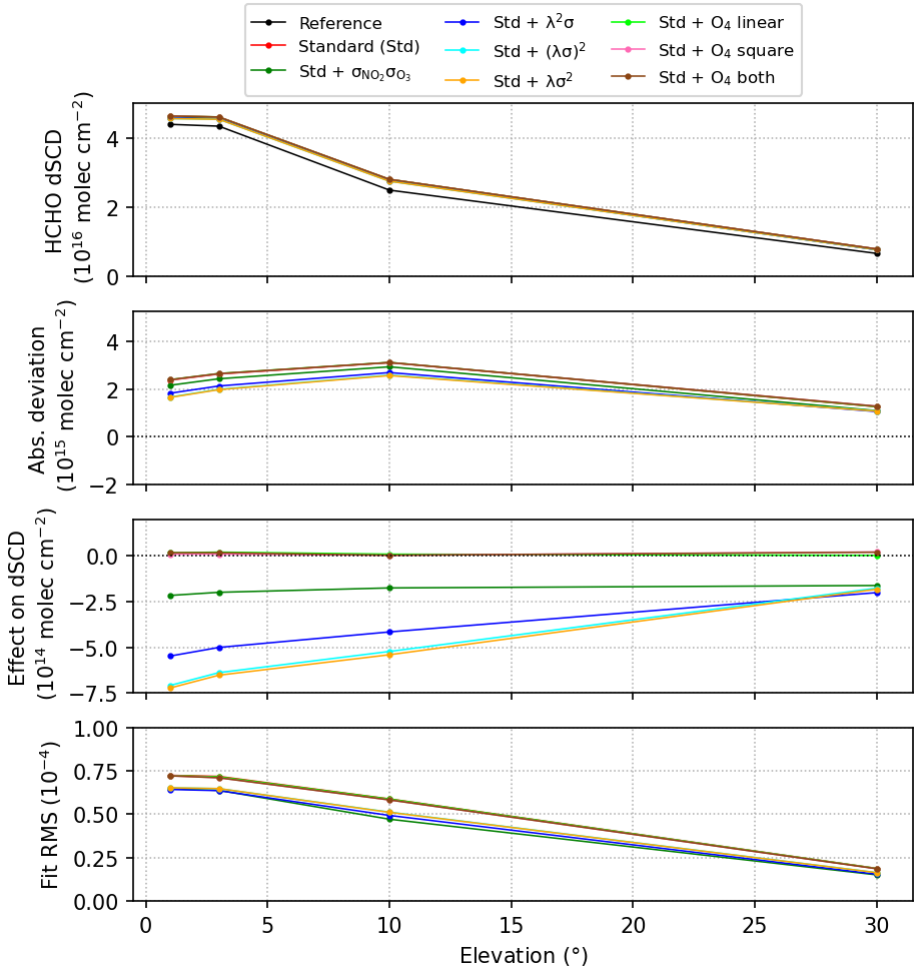


Figure 6.31: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effects of the further Puḳīte terms of NO_2, O_3 and O_4 on the HCHO dSCDs and fit RMS as functions of elevation for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. Results obtained with the optimised standard settings and different sets of Puḳīte terms included in the DOAS fit are shown in different colours.

the HCHO dSCDs are marginally increased for the fit scenarios including Puḳīte terms for O_4 leading to worse agreements between retrievals and simulations at all elevations. These findings also hold for the other not presented scenarios, whereby the absolute values of the effect on the HCHO dSCDs are higher for some of the other scenarios but stay below ca. 2×10^{14} molec cm^{-2} and are mostly positive. Also the fit RMS-values are not significantly improved for any scenario. These negligible effects of the O_4 Puḳīte terms on the HCHO results can probably be explained by the spectrally rather broad and, compared to NO_2 and O_3, relatively weak absorptions of O_4. Hence, the usage of Puḳīte terms for O_4 cannot explain the remaining deviations between retrieved and true reference HCHO dSCDs.

Similar to the effects of the intensity weighted convolutions in the previous section, the inclusion of additional (higher order) Puḳīte terms for NO_2 and O_3

reduces the retrieved HCHO dSCDs. As expected, the absolute values of these reductions depend on the absorption strength and are therefore higher at low elevations, where the retrieved HCHO dSCDs can be up to 7.2×10^{14} molec cm^{-2} smaller for the presented sequence depending on the specific additionally included terms (compare different lines in the third panel of Fig. 6.31). The largest effects are found for the $\lambda\sigma^2$ and $(\lambda\sigma)^2$ terms for which the effects are very similar. At 30°, the effects are smaller and similar for all additional Puḳīte terms. Here, around 2×10^{14} molec cm^{-2} smaller HCHO dSCDs are obtained. For the presented example, the effects on the dSCDs are of the same order as for the Ring effect, noise and intensity weighted convolutions for the remaining absorbers. Besides the effects on the dSCDs, also the fit RMS is affected. In particular, small reductions are achieved which are similar for all additionally tested Puḳīte terms for NO_2 and O_3. For the presented scenario, the findings are rather clear and the remaining deviations are reduced by roughly 30 % at low elevations, whereby the effect on the corresponding dSCDs is again in the percent range. However, more complex patterns are found for some of the other geometries which are briefly summarised below.

For a given sun geometry and NO_2 scenario, the effects of the additional higher order Puḳīte terms of NO_2 and O_3 show no strong dependence on the HCHO scenario. While the general patterns are very similar, the absolute effects on the HCHO dSCDs differ slightly between both HCHO scenarios (maximum difference: ca. 2×10^{14} molec cm^{-2}), where the effects are slightly larger for the low HCHO scenarios at low elevations.

Further, the general patterns of the effects of the further higher order terms do not strongly depend on the NO_2 scenario. Also here the maximum differences between both NO_2 scenarios are in the order of 2×10^{14} molec cm^{-2}. These only small differences between the low and high NO_2 scenarios suggest that the impacts of such higher order effects on the HCHO fit are not large. This was different for the first-order Puḳīte terms of NO_2 which showed clearly larger effects for the high NO_2 scenarios (compare Sect. 6.2.2).

In general, the effects of the further higher order terms on the HCHO dSCDs are also similar for different sun geometries. However, the differences between the geometries can exceed 3×10^{14} molec cm^{-2} and are thus larger than the ones between the HCHO and NO_2 scenarios. At low elevations, the additional Puḳīte terms of NO_2 and O_3 lead to approximately 2×10^{14} to 8×10^{14} molec cm^{-2} smaller HCHO dSCDs depending on the specific term and sun geometry. Also at elevations of 10° and 30°, mainly small reductions of the retrieved HCHO dSCDs are obtained with the additional higher order terms. For a few of the investigated sun geometries (e.g. an SZA of 20° at an RAA of 180°), however, slightly larger HCHO dSCDs are obtained with all additional cross-sections at 30° elevation. Further, there is a clear exception to the otherwise overall consistent findings, namely spectra with an SZA of 60° and an RAA of 30° which also showed different behaviours for some of the previous studies (e.g. influence of the Ring effect). Here, the results obtained when the cross-correlation term $\sigma_{NO_2}\sigma_{O_3}$ is included in the DOAS fit show a different elevation dependence (for

all trace gas scenarios) with the largest effects on the retrieved HCHO dSCDs occurring at 30° elevation and not at the low elevations.

Altogether, the additional higher order Puḳīte terms of NO_2 and O_3 lead to a systematic reduction of the retrieved HCHO dSCDs for almost all investigated scenarios (for exceptions compare above). Moreover, only small differences between the different scenarios are found and the absolute effects on the dSCDs are similar. These reductions lead to slightly better agreements between retrievals and simulations especially for scenarios which showed substantial remaining deviations after the optimisations of the DOAS fit settings, e.g. the scenario presented in Fig. 6.31. For scenarios for which rather small deviations from the reference dSCDs are obtained with the optimised standard fit settings, e.g. spectra with an SZA of 20° at an RAA of 30° for the low NO_2 case, the observed reductions of the retrieved HCHO dSCDs lead to rather good agreements between retrievals and simulations especially at the lowest two elevations. Similar to the findings for the intensity weighted convolutions in the previous section, a worsening of the agreements caused by the systematic reductions of the retrieved HCHO dSCDs is observed for scenarios which showed only small over- or underestimations of the reference dSCDs and thus small deviations for the optimised standard settings (e.g. spectra with an SZA of 20° at an RAA of 180° for the low NO_2 scenario). Lastly, it should be mentioned that a view on the spectral fits themselves shows that the different higher order terms of a given absorber (NO_2 or O_3) strongly correlate and interfere amongst each other. Thus, in summary, also the usage of either additional higher order Puḳīte term of NO_2 and O_3 cannot explain the remaining deviations as the systematic reductions of the dSCDs occur similarly for all sequences and not in an appropriate way specific for each scenario.

The investigations presented in this section were also briefly performed for the spectra which do not include the Ring effect. It is interesting to note that for these spectra very different patterns in the results obtained with additional higher order terms are found. For instance, the inclusion of (first-order) Puḳīte terms for O_4 in the DOAS fit leads to HCHO dSCDs that are larger by up to 8×10^{14} molec cm^{-2} compared to the ones obtained with the optimised standard settings for many (but not all) scenarios. This is an effect in the "wrong" direction as it worsens the agreement between retrievals and simulations. Also the additional higher order Puḳīte terms of NO_2 and O_3 show different effects on the HCHO results compared to the standard spectra (without noise), whereby for many (but not all) sequences enhanced (up to ca. 5×10^{14} molec cm^{-2}) dSCDs are obtained. In particular, this leads to larger deviations from the true reference HCHO dSCDs also for scenarios which showed relatively small deviations for the optimised standard settings, e.g. spectra with an SZA of 20° at an RAA of 30° for the low NO_2 case. A last observation is that the differences between the different sun geometries are larger for the spectra without Ring effect. These rather contradictory findings for spectra with and without Ring effect show that the usage of additional (likely too many) higher order Puḳīte terms is probably associated with subtle spectral interferences which limits their additional value.

6.4.4 *Wavelength dependence and accuracy of the AMFs*

The investigations on the remaining deviations of the HCHO dSCDs retrieved using the optimised fit settings from the true reference dSCDs presented in the three previous sections focused on uncertainties caused by the retrieval and its settings. However, also uncertainties associated with the calculation of the reference HCHO dSCDs from the VCDs used for the simulations of the synthetic spectra (compare Sect. 6.1.1) might contribute to the remaining deviations. Hence, two effects related to the AMFs which are used to convert the input VCDs to the reference HCHO dSCDs[7] (compare Sect. 3.2.3 and 6.1.2) are considered here. Firstly, also the simulations of the AMFs underlie approximations and are thus associated with uncertainties which in consequence potentially lead to inaccurate reference dSCDs to which the retrieval results are compared. Secondly, the AMFs are wavelength dependent (compare Sect. 3.2.3) and the choice of the wavelength to which the AMFs are interpolated might contribute to the remaining deviations between retrievals and simulations. The influences of both effects on the comparison results are examined in this section.

The AMFs used in this sensitivity study were simulated using SCIATRAN (compare Sect. 6.1.1). Typically, the standard accuracy and precision settings are used which can be considered as a reasonable trade-off between computing time and sufficiently accurate results. However, these standard simulation settings do not yield the highest possible accuracy and can be modified in a way to guarantee the latter. To investigate whether the limited accuracy of the AMFs simulated with the standard accuracy and precision settings can explain the remaining deviations between retrieval and simulations, AMFs were simulated with different enhanced accuracy and precision settings for an SZA of 45° and the RAAs specified in Table 6.1 for the case of high NO_2. The resulting AMFs at 10° elevation and an RAA of 180° (example sun geometry used in Sect. 6.2) are compared in Fig. 6.32. The upper left panel shows the wavelength dependencies of the AMFs obtained with different accuracy and precision settings, while the relative differences of the AMFs for the different settings with respect to the standard settings are shown in the upper right panel. Green lines represent the results for simulations performed with double precision compared to the standard settings (blue) to solve the DOM (discrete ordinate method) eigenvalue problem. Red lines indicate the results of simulations for which in addition to the double precision twice as many discrete angles were used for the interpolation on the local solar angles along the line of sight (Dbl + UD20). Lastly, the results of simulations for which twice as many streams (discrete zenith angles) were used to approximate the integrals in the simulations are shown in cyan (Dbl + 32 streams). The lower panel depicts the (convolved) HCHO absorption cross-section and the wavelengths for which the AMFs have been simulated (red dots).

For most wavelengths, the relative differences are below ca. ± 0.2 %, while larger differences occur for wavelengths outside of HCHO absorption bands, namely at 335, 350 and 360 nm. No significant differences between the results obtained

7 In fact, dAMFs are used to perform this conversion. However, these are calculated from the AMFs as described in Sect. 3.2.3 and 6.1.2.

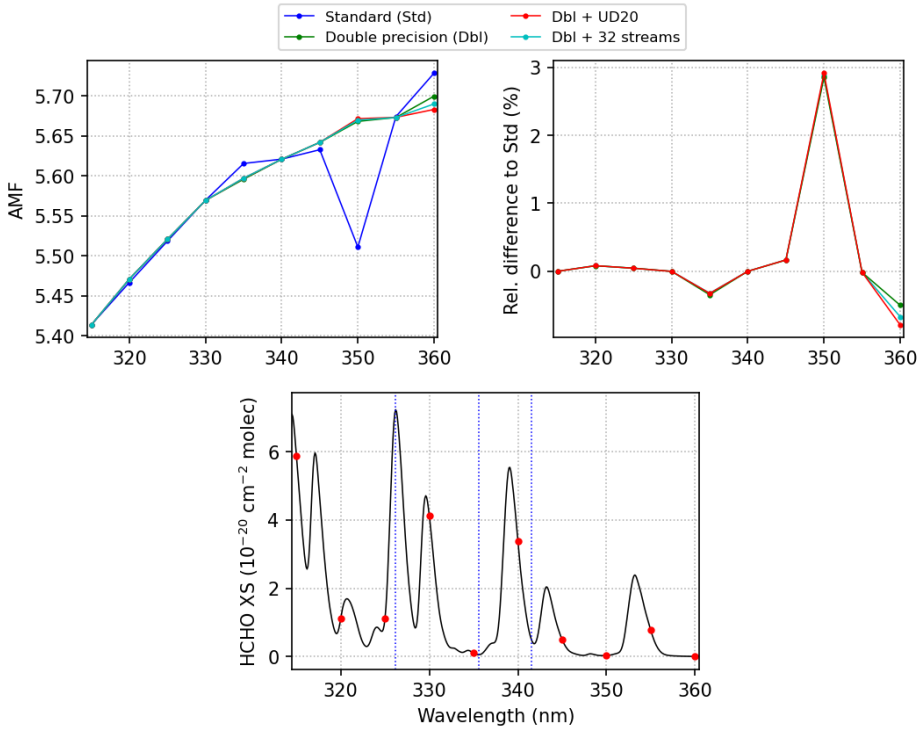


Figure 6.32: The upper left panel compares the wavelength dependencies of the AMFs obtained with different accuracy and precision settings as described in the text for the case of high NO_2, an SZA of 45° and an RAA of 180° at 10° elevation. The corresponding relative differences with respect to the standard settings are presented in the upper right panel. The (convolved) HCHO absorption cross-section and the wavelengths for which the AMFs have been simulated (red dots) are depicted in the lower panel, where blue vertical dashed lines additionally indicate different choices of interpolation wavelengths as described in the text.

with the different enhanced accuracy and precision settings are found. Overall, these findings also hold for the other (not presented) geometries (RAAs and elevations) for which only small variations in the (relative) differences are observed. Despite these findings, for many scenarios, the remaining deviations from the reference dSCDs are larger than the here observed small effects of the enhanced simulation settings on the AMFs. Thus, also the accuracy and precision of the AMF simulations cannot explain the remaining differences between retrieved and reference HCHO dSCDs. Nevertheless, one important finding here is that only AMFs at wavelengths which show significant HCHO absorptions should be used for the interpolation of the AMFs to the wavelengths that are used to calculate the true reference dSCDs since for such wavelengths the potential errors of the simulations of the AMFs are small. In particular, AMFs at wavelengths of 335, 350 and 360 nm should be excluded from the interpolation to keep the effect of the simulation uncertainties of the AMFs considerably below 0.5 %. It should be stressed that this procedure was already applied for all previous studies (compare Sect. 6.1.2).

As described earlier (compare Sect. 3.2.3), the AMFs are wavelength dependent. The upper left panel of Fig. 6.32 shows that for the presented viewing geometry the AMFs vary by around 3 % within the wavelength range from 324 to 359 nm (standard fitting interval). The strength of this wavelength dependence depends strongly on the elevation angle leading to a variation of the AMF of up to 15 % (not shown) in this wavelength range at 1° elevation. This observed wavelength dependence of the AMFs is mainly caused by the increasing Rayleigh scattering probability towards smaller wavelengths. Given these observations, the reference dSCDs which are used for the comparisons of the retrieved HCHO dSCDs with the simulations depend on the choice of the wavelength to which the AMFs are interpolated. So far, the centre wavelengths of the respective fitting intervals were selected leading for example to 341.5 nm for the standard fitting interval. However, also other choices are justified:

- Interpolation to the wavelength of the strongest absorption band within the applied fitting interval (326.2 nm for the standard interval).
- Interpolation to the "effective" wavelength of the applied fitting interval which is defined as the weighted mean wavelength within the applied fitting interval (335.6 nm for the standard interval).

All three wavelength choices are marked by blue vertical dashed lines in the lower panel of Fig. 6.32.

To examine the influence of the wavelength dependence of the AMFs, and consequently also the reference dSCDs, on the comparison results, reference HCHO dSCDs were calculated using the three described wavelengths to which the AMFs were interpolated for each of the previously used scenarios (compare Sect. 6.2.2.1). Then, absolute deviations of the HCHO dSCDs retrieved with the optimised fit settings with respect to the different reference dSCDs were calculated. Again, results for spectra without noise but including the Ring effect were used for these investigations.

Figure 6.33 shows the retrieved HCHO dSCDs using the optimised fit settings and reference dSCDs (upper panel), the corresponding absolute deviations from the reference dSCDs obtained with the three different wavelengths to which the AMFs were interpolated (middle panel) as well as the effects on the reference dSCDs with respect to the standard interpolation wavelength (lower panel) as functions of elevation for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180° (same sun geometry as in Fig. 6.32). The results for different wavelength choices are shown in different colours as indicated by the legend above the figure.

For the presented scenario, the retrieval using the optimised fit settings overestimates the HCHO dSCDs at all elevations when the AMFs interpolated to the centre wavelength of the fitting interval are used to calculate the reference dSCDs (standard approach). Given the observed wavelength dependence of the AMFs (compare upper left panel in Fig. 6.32) and the fact that both alternative wavelengths are shorter than the centre wavelength of the used fitting interval (341.5 nm), even larger overestimations of the HCHO dSCDs are obtained when the alternative wavelengths are used for the interpolation of the AMFs

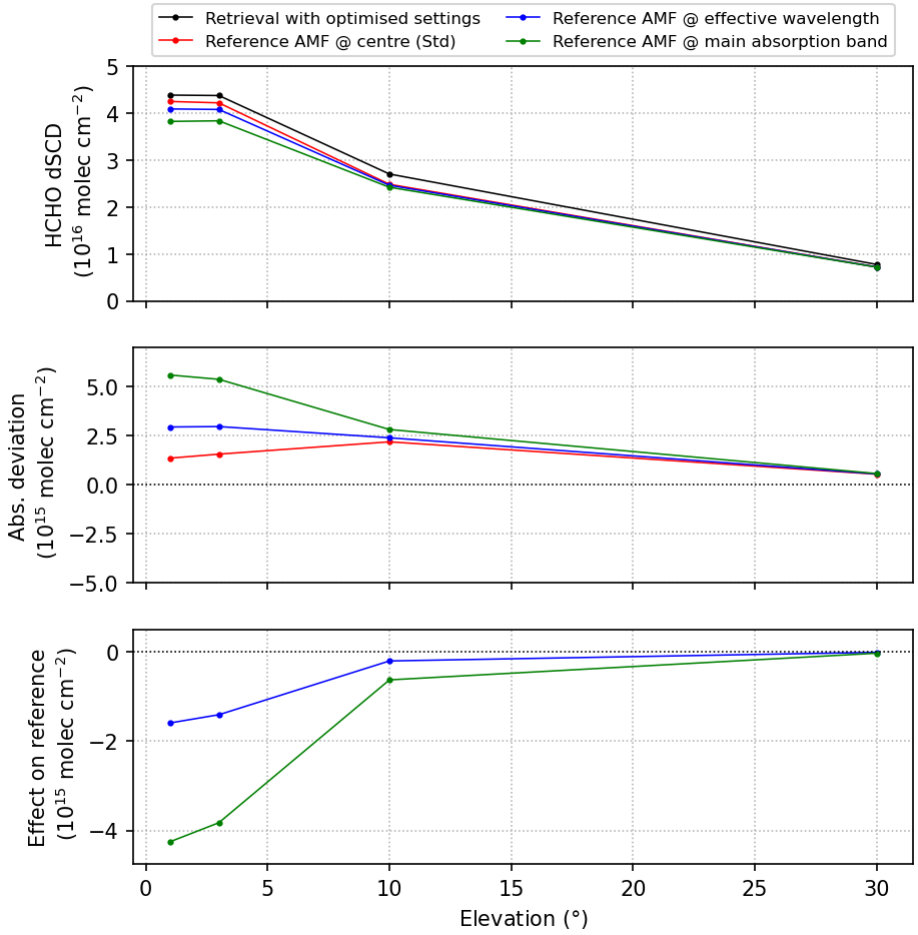


Figure 6.33: Retrieved and reference HCHO dSCDs, absolute deviations from the reference dSCDs obtained with the three different interpolation wavelengths and corresponding effects on the reference dSCDs with respect to the standard interpolation wavelength as functions of elevation for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. Results obtained with the three different interpolation wavelengths as explained in the text are indicated by different colours.

and the subsequent calculation of the reference dSCDs (blue and green lines in Fig. 6.33). Consistent to the elevation dependence of the wavelength dependence of the AMFs, this effect is largest at low elevations, where the reference dSCDs are around 4×10^{15} molec cm^{-2} and 1.5×10^{15} molec cm^{-2} smaller for the wavelength of the main absorption band and the effective wavelength, respectively (compare lower panel). Accordingly, the overestimations are larger (compare middle panel). Since the wavelength dependence of the AMFs is weaker at higher elevations, this effect becomes smaller with increasing elevation and at 30° the results of the three different scenarios are barely distinguishable. At 1° elevation, the effect on the reference dSCD is around 10 % for the wavelength of the main absorption band and agrees reasonably well with the change of the AMF between these two wavelengths. Altogether, worse agreements between retrievals and simulations are obtained using different interpolation wavelengths for the presented example.

In general, the wavelength dependence of the AMFs is similar for all (not presented) scenarios since the main effect is the increasing scattering probability towards shorter wavelengths leading to shorter average light paths and thus smaller AMFs. Consequently, the reference dSCDs obtained with the two alternative interpolation wavelengths are smaller compared to the ones obtained with the centre wavelength of the fitting interval for almost all scenarios[8] (trace gas scenario and sun geometries). Further, the absolute effects on the reference dSCDs are comparable for all NO_2 scenarios and sun geometries, whereas the absolute differences are smaller for the low HCHO scenario caused by the lower dSCDs in that case. These findings lead to worse agreements compared to the standard approach for all scenarios which show overestimations of the HCHO dSCDs like the presented example scenario. However, for scenarios for which slight underestimations are obtained with the optimised fit settings using the standard interpolation wavelength, the smaller reference dSCDs can lead to better agreements (e.g. for spectra with an SZA of 20°, an RAA of 180° and high NO_2 abundance). Nevertheless, also the wavelength dependence of the AMFs and thus of the reference HCHO dSCDs cannot explain the remaining deviations between DOAS retrievals and simulations since the agreements are worsened especially for the problematic scenarios.

During the studies presented in this section, also the wavelength dependencies of the AMFs for the high aerosol scenario (compare Table 6.1) which was not included in the sensitivity studies so far were investigated. Thereby, it was found that for many sun geometries notably better agreements between retrieved and reference HCHO dSCDs are found for the high aerosol scenario. This becomes also evident in Fig. 6.34 which shows the same quantities for the same trace gas scenario and sun geometry as Fig. 6.33 but for the case of high AOD (i.e. an AOD of 0.5).

Obviously in the case of high AOD, the retrieved HCHO dSCDs as well as the deviations between retrievals and simulations are smaller. Moreover, also the elevation dependence of the absolute deviations with typically largest deviations at 10° elevation is significantly weaker for the high AOD scenario. Although in theory complex relations between AODs and AMFs exist, these two findings can be explained by smaller AMFs at low elevations and slightly larger AMFs at high elevations (30° and 90°) in the case of high AOD. The lower AMFs at low elevations are caused by generally shorter average light paths in the atmospheric layers which are relevant for the HCHO absorptions and are a consequence of enhanced scattering on aerosol particles. For high elevations, this effect appears to be the opposite. These lower AMFs at low elevations are related with weaker trace gas absorptions leading to smaller deviations. In view of these rather good agreements between retrieved and referenced dSCDs for the high AOD case, the optimised fit settings are apparently well justified. However, it should be mentioned that for a few geometries larger deviations are found for the high AOD scenario especially for scenarios which showed good agreements for the low AOD case (e.g. spectra with an SZA of 20°, an RAA of 180° and low NO_2 abundance). Lastly, also the wavelength dependence of the AMFs is smaller for

8 A few exceptions exist at 30° elevation, e.g. in case of an SZA of 60°, an RAA of 30° and high NO_2.

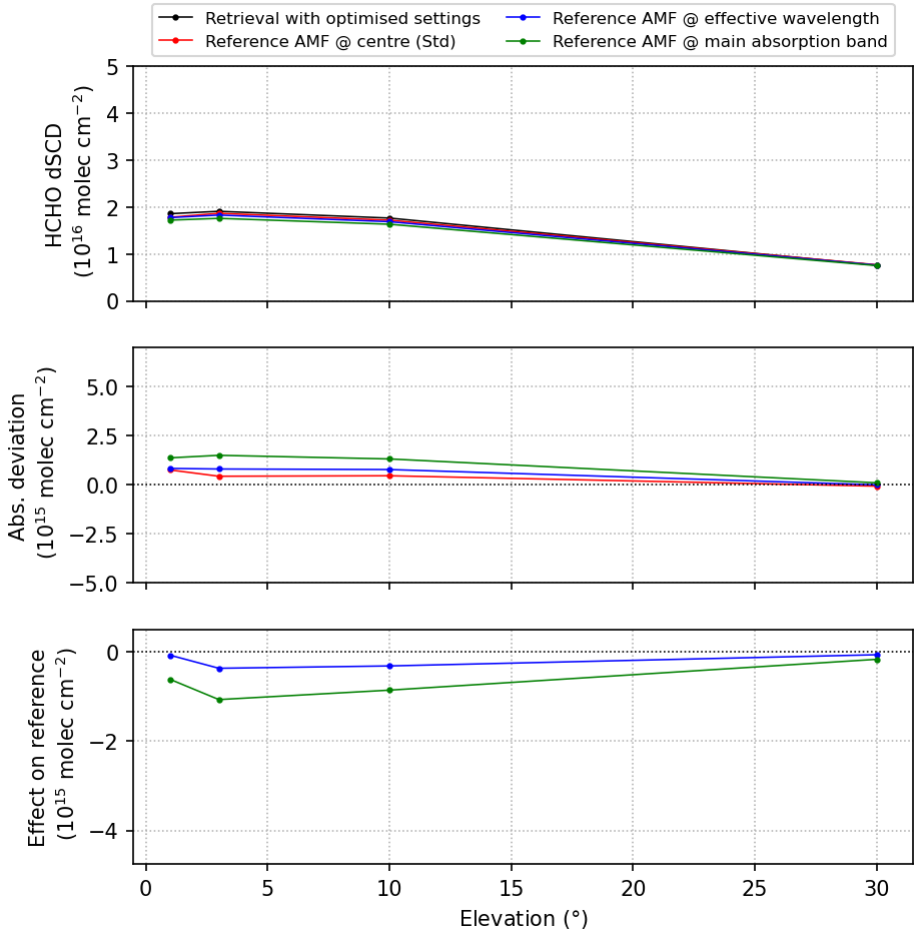


Figure 6.34: Same as Fig. 6.33 but for the case of high AOD.

the high AODs which is represented in Fig. 6.34 by the smaller spread of the reference dSCDs obtained with different interpolation wavelengths (compare middle and lower panels).

In conclusion, neither the accuracy and precision of the AMF simulations nor the wavelength dependence of the AMFs can explain the remaining deviations. The standard approach of interpolating the AMFs to the centre of the used fitting interval appears to be well suited at least for the here-proposed interval. In theory, this might be different for other intervals for which the effective wavelengths and/or the wavelengths of the strongest absorption band are longer than the centre wavelengths. For such intervals, however, also the absolute deviations might be different (compare Sect. 6.3.2).

6.4.5 *Accuracy and precision of the simulated spectra*

As none of the previously examined effects can (completely) explain the remaining deviations of the retrieved HCHO dSCDs from the reference dSCDs, another possible cause of these deviations are the accuracy and precision of the simulations of the synthetic spectra themselves. More precisely, the question is how well the

assumed trace gas abundances are represented in the simulated spectra. In order to study to which extent the accuracy and precision settings of the simulations affect the agreement between DOAS retrieval and simulations, spectra for two selected sequences were simulated with enhanced accuracy and precision settings. In particular, the selected scenarios are:

- High HCHO, high NO_2, low AOD at an SZA of 45° and an RAA of 180° which can be considered as standard sun geometry since it was used for many investigations of this sensitivity study.
- High HCHO, high NO_2, low AOD at an SZA of 60° and an RAA of 30° which is a sequence for which despite all optimisation efforts still high remaining deviations are observed.

Spectra without noise were simulated using two sets of enhanced accuracy and precision settings which are similar to the ones for the simulations of the AMFs presented in the previous section. One simulation was performed with higher accuracy (including double precision) and 16 streams as for the standard simulations. The second simulation used the same improved accuracy settings but was performed with 32 streams which is supposed to yield an even higher accuracy. For further details on the enhanced accuracy and precision settings as well as an explanation of the term *streams*, refer to Sect. 6.4.4. Given the findings for the influence of the Ring effect on the remaining deviations (compare Sect. 6.4.1), spectra with Ring effect were simulated, whereby the corresponding simulation settings were not changed for the sake of calculation time.

First, the spectra obtained with the enhanced accuracy and precision settings are compared to the standard spectra. Since the DOAS fit is performed in optical depth space (compare Sect. 3.1), differences in terms of optical depths are relevant. Therefore, the natural logarithm of the ratio between the spectra with enhanced accuracy and precision and the standard spectra is calculated. Figure 6.35 shows these logarithmic ratios for the case of high HCHO and NO_2, low AOD, an SZA of 60°, an RAA of 30° and 3° elevation. The logarithmic ratios for the 16 and 32 streams scenarios are represented by green and blue lines, respectively, while vertical black dashed lines mark the HCHO fitting interval used in this work. This specific geometry was chosen since it shows the largest effects of the enhanced accuracy and precision settings.

For the presented example, the logarithmic ratios show for both improved simulation settings pronounced broad-band wavelength dependencies[9] which are stronger for the 16 streams scenario (compare green line). In the wavelength range relevant for the HCHO retrieval, relatively weak wavelength dependencies of the logarithmic ratios are found and the maximum values are around 4.1×10^{-4} for the 32 streams scenario. For the 16 streams scenario, the logarithmic ratios are below 3.5×10^{-4} in the relevant spectral range. The absolute values of the logarithmic ratios as well as the differences between the 16 and 32 streams scenarios become smaller with increasing elevation. At 90°, no differences between

9 Besides these, also high frequency variations are observed whose amplitudes are around 0.2×10^{-4} and thus irrelevant.

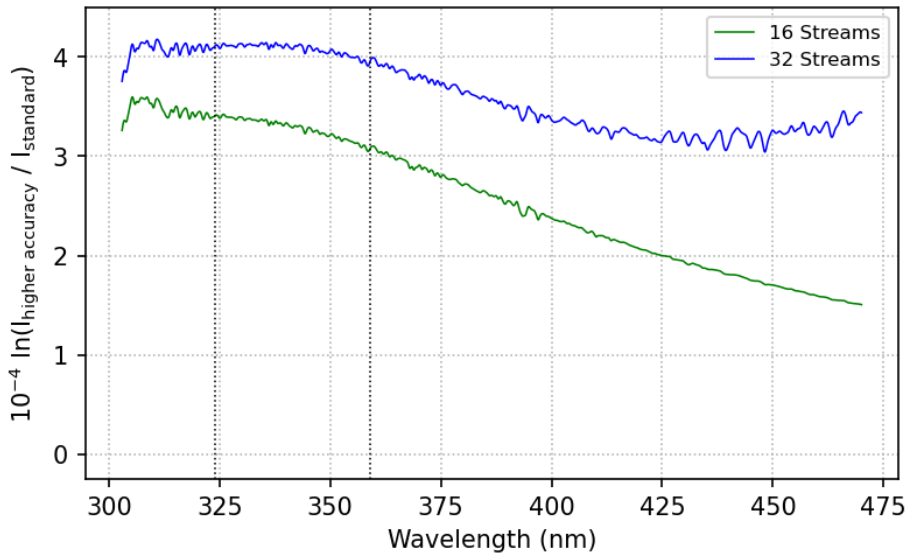


Figure 6.35: Logarithmic ratios of spectra obtained with enhanced accuracy and precision settings with respect to the standard spectrum for the case of high HCHO and NO_2, low AOD, an SZA of 60° and an RAA of 30° at 3° elevation. Results for the 16 and 32 streams scenarios are shown in green and blue, respectively. The used HCHO fitting interval is marked by vertical black dashed lines.

both enhanced accuracy settings are observable and also the absolute values of the logarithmic ratios are almost 0 indicating no effect of the enhanced accuracy and precision settings anymore.

No significant effects of the improved simulations settings are found in the 16 streams case at all elevations inside the relevant wavelength range for the second simulated (not presented) sun geometry. Also outside this spectral region, only negligible fluctuations occur. The 32 streams scenario exhibits largest logarithmic ratios of up to ca. -0.75×10^{-4} at 1° elevation inside the HCHO fitting range. Like for the previous sun geometry, the differences between both simulation settings as well as their absolute values decrease with increasing elevation and practically no effects occur at 90°.

In summary, the observed effects of the enhanced accuracy and precision settings of the simulations on the synthetic spectra are in the order of (at maximum) a few 10^{-4} in terms of optical depths. Although these effects correspond to roughly 10 % of the typical optical depths of the HCHO absorptions (in the order of a few 10^{-3} for a dSCD of a few 10^{16}), no spectral structures which are similar to the absorption structures of HCHO are found (compare Fig. 6.35 and lower panel of Fig. 6.32). Hence, no spectral interferences are expected and the effects on the retrieved HCHO dSCDs are presumed to be small.

To prove this expectation, the spectra simulated with enhanced accuracy and precision settings were analysed using the optimised DOAS fit settings. The resulting HCHO dSCDs are compared to the ones obtained with the standard spectra in Fig. 6.36 which shows the retrieved HCHO dSCDs, the corresponding absolute deviations from the reference dSCDs, the effects of the improved simulations settings on the dSCDs and the fit RMS as functions of elevation for

the case of high HCHO and NO_2, low AOD, an SZA of 60° and an RAA of 30°. The results for the different simulation settings are shown in different colours as indicated by the legend.

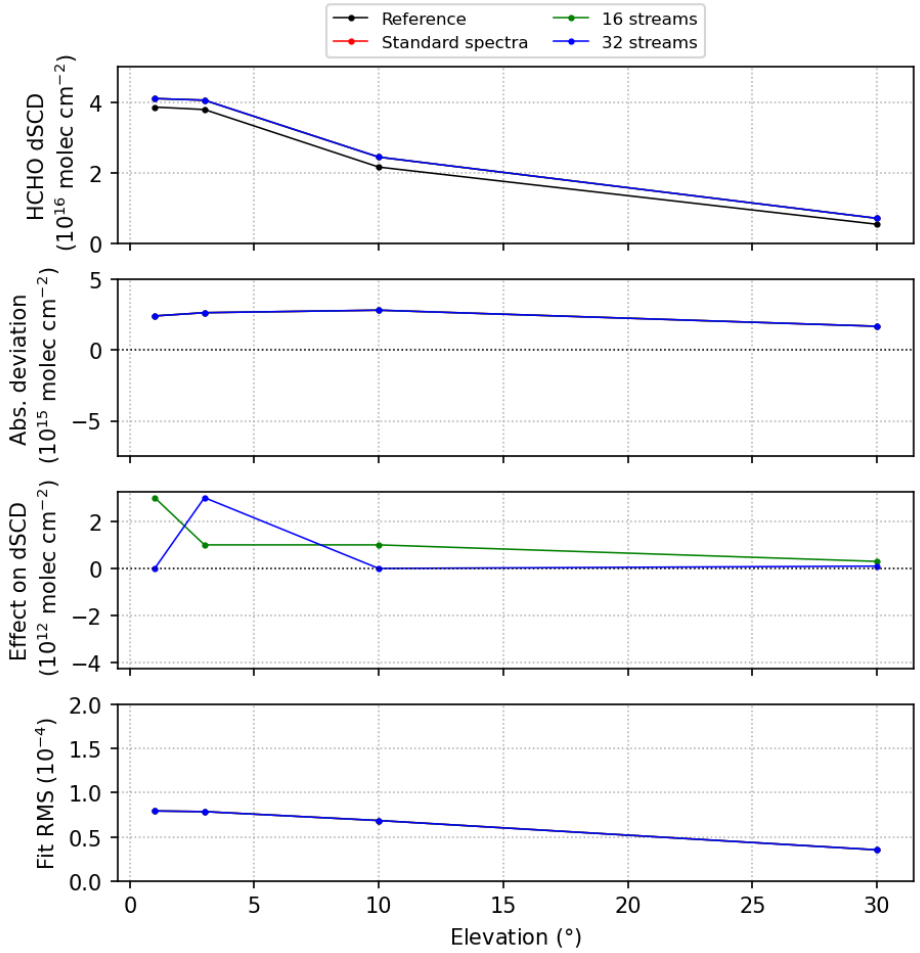


Figure 6.36: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effects of the enhanced accuracy and precision settings for the simulations of the synthetic spectra on the HCHO dSCDs and fit RMS as functions of elevation for the case of high HCHO and NO_2, an SZA of 60° and an RAA of 30°. Results obtained with the optimised standard settings for the different sets of synthetic spectra are shown in different colours.

For dSCDs, absolute deviations from the reference dSCDs and fit RMS, the results for the spectra obtained with the different simulation settings are almost identical and the corresponding lines are not distinguishable (compare uppermost, second and lowermost panels) indicating in general negligible effects. The absolute effects on the retrieved HCHO dSCDs are at maximum ca. 3×10^{12} molec cm^{-2} at low elevations (compare third panel) and thus, roughly four orders of magnitude smaller than the dSCDs themselves[10]. With increasing

10 Note that *QDOAS* provides only four decimal digits in the output files corresponding to 10^{12} molec cm^{-2} for dSCDs in the order of 10^{16} molec cm^{-2} leading to a rather discrete pattern in the third panel.

elevation, the effects become even smaller and are almost 0 at 30°. For the second additionally simulated (not presented) sequence, the same findings and similar values are valid but the absolute effects on the dSCDs are in the opposite direction (marginally smaller columns). Also for this scenario, the fit RMS is practically not affected.

In conclusion, the accuracy and precision of the simulated synthetic spectra do almost not contribute to the observed remaining deviations of the retrieved HCHO dSCDs from the reference dSCDs. In view of these findings, the standard accuracy and precision settings of the simulations of the synthetic spectra are sufficient.

6.4.6 *Low resolution simulations*

Besides the accuracy and precision of the synthetic spectra, two further implications regarding the spectra and their analyses using the DOAS method exist.

The first one is related to the temperature dependence of the trace gas absorptions which is relevant for species occurring at different altitudes, and thus at different temperatures, in the atmosphere. These temperature dependencies were also taken into account for the simulations of the synthetic spectra by assuming typical atmospheric profiles for such trace gases and including absorption cross-sections at several temperatures (compare Sect. 6.1.1 as well as Table 6.1 and 6.2). For instance, O_3 cross-sections at temperatures ranging from 193 to 293 K (in steps of 10 K) and O_4 cross-sections at 203 and 293 K were included in the simulations (compare Table 6.2). In a typical DOAS analysis, however, not all absorption cross-sections can be included to avoid over-fitting. Therefore, typically only selected cross-sections at representative temperatures are included in the DOAS fits, e.g. cross-sections at 223 and 243 K for O_3 or at 293 K for O_4 (compare Table 6.3). For some specific scenarios, these limitations might be relevant and contribute to the remaining deviations.

The second implication is more fundamental and related to the DOAS method itself. As described in Sect. 3.1.4, during the DOAS analysis the logarithmic ratio of the spectrum to be analysed and the FRS is calculated. However, both spectra were measured at the rather low spectral resolution of the instrument (or convolved before the analysis for the synthetic spectra). Consequently, the convolution by the instrument's slit function is performed before the logarithmic ratio is calculated. In the model assumed for the DOAS analysis, however, the logarithmic ratio should be convolved (compare Sect. 3.1.4 and 3.1.5.3). Since convolution and logarithm as well as calculating the ratio and convolution do not commute, inaccuracies are introduced which might lead to differences between retrieved and reference dSCDs. These effects were already described in Sect. 3.1.5.3 and are known as saturation and solar I_0-effect, respectively. While the inaccuracies related to the saturation effect increase with increasing absorption strength, the I_0-effect can be accounted for by applying I_0-corrected or intensity weighted convolutions for strong absorbers. Since the effect is only small for weak absorbers, it is typically neglected for these (compare Sect. 6.4.2). As the described correction methods are approximations, higher order effects might remain and still cause inaccuracies

leading to the observed remaining deviations of the retrieved HCHO dSCDs from the true reference values.

To estimate the influence of the described effects on the HCHO results, low resolution simulations at the output resolution of 0.1 nm instead of 0.01 nm were performed for the two selected sequences used in the previous section (compare Sect. 6.4.5). For these simulations, already convolved trace gas absorption cross-sections obtained with standard convolutions using a Gaussian function with a FWHM of 0.65 nm (compare Sect. 6.1.1) were used as input instead of the high resolution cross-sections. In addition, also the solar reference spectrum was already convolved. Spectra with and without Ring effect were simulated. These additionally simulated spectra cover a slightly smaller wavelength range, spanning from 307 to 465 nm, caused by limitations of the convolution tool of the *QDOAS* software package and the fact that the simulation of Raman spectra requires a certain wavelength margin at the edges of the considered wavelength range.

The first set of simulations is called *first principles* (FP) in the following and includes only one low resolution cross-section per absorber, whereby water vapour absorptions were not considered at all. For these simulations, a different O_3 cross-section at 213 K (instead 223 K like in the standard analysis settings) was used[11]. The spectra resulting from the FP simulations were analysed including only exactly the same set of absorption cross-sections which was used for the simulations to examine the effects of the first implication mentioned above. In particular, these cross-sections are HCHO, BrO, NO_2 at 294 K (and corresponding Puķīte terms), O_3 at 213 K (and corresponding Puķīte terms) and O_4 at 293 K.

The second set of low resolution simulations also used the already convolved (low resolution) cross-sections (and solar reference spectrum) as described above but included all cross-sections listed in Table 6.2. In particular, the NO_2 cross-section at 220 K, the O_4 cross-section at 203 K and O_3 cross-sections at various additional temperatures were additionally included for this second set of low resolution simulations as well as a water vapour cross-section. In that way, only saturation and I_0-effect do not occur and potential inaccuracies introduced solely by them can be estimated. These simulations are called *low resolution* (LR) in the following.

Besides the changes described above, the optimised fit settings were used. The HCHO results obtained for the two additional sets of synthetic spectra are compared to the ones obtained for the standard spectra in Fig. 6.37 which shows the retrieved HCHO dSCDs (uppermost panel), the corresponding absolute deviations from the reference dSCDs (second panel), the differences between the HCHO dSCDs obtained for the spectra simulated at low spectral resolution and the standard spectra (third panel) and the fit RMS (lowermost panel) as functions of elevation for the case of high HCHO and NO_2, an SZA of 60° and an RAA of 30°. The results for the different sets of synthetic spectra are shown in different colours as indicated in the legend.

11 This was not intended. However, since this choice is expected to be not critical for this specific purpose, the simulations were not repeated for the sake of time.

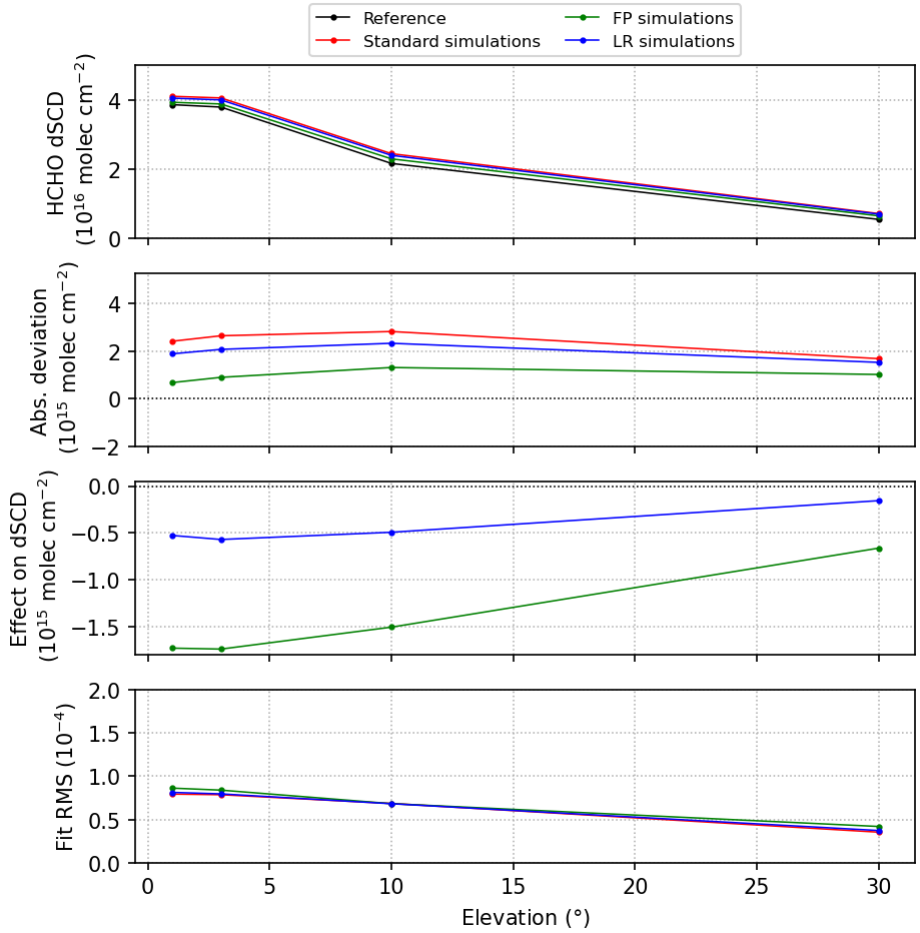


Figure 6.37: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, differences between the HCHO dSCDs obtained for the spectra simulated at low spectral resolution and the standard spectra and fit RMS as functions of elevation for the case of high HCHO and NO_2, an SZA of 60° and an RAA of 30°. Results obtained for the different sets of synthetic spectra are shown in different colours.

For the presented example, both sets of synthetic spectra simulated at low spectral resolution lead to smaller retrieved HCHO dSCDs than the standard spectra and thus to smaller deviations from the reference dSCDs at all elevations. The effect is largest at low elevations with an absolute difference with respect to the results of the standard spectra of up to ca. -1.7×10^{15} molec cm^{-2} for the FP simulations, while smaller differences are found at high elevations where the absolute difference is only -0.66×10^{15} molec cm^{-2} at 30°. The influence on the HCHO dSCDs which can be attributed to saturation and I_0-effect (LR simulations) is around -0.55×10^{15} and -0.15×10^{15} molec cm^{-2} at low elevations and 30°, respectively, and thus roughly one quarter to one third of the total effect of the low resolution simulations. Consequently, the inaccuracies caused by the fact that not all cross-sections are included in a typical DOAS fit are larger. The fit RMS is only marginally affected with negligibly larger values for the spectra simulated at low resolution.

Similar overall findings are valid for the second additionally simulated (not presented) sequence. The absolute effects, however, are slightly smaller and differences in the retrieved HCHO dSCDs with respect to the results obtained for the standard spectra of -1.25×10^{15} and -0.3×10^{15} molec cm^{-2} are found at low elevations for the FP and LR simulations, respectively. For this geometry, these lower HCHO dSCDs lead to worse agreements at low elevations compared to the standard spectra for the FP simulations, while the deviations are almost 0 at 30°. Further, the RMS is almost not affected showing marginally smaller values for both low resolution simulations of this sequence.

The observed changes in the retrieved HCHO dSCDs for the FP simulations (both effects together; green lines) are larger than the other effects discussed in this section so far[12], while the individual effect of the LR simulations is comparable (or only slightly larger) to the influences of the Ring effect, noise, intensity weighted convolutions for the weak absorbers or some of the further higher order terms of NO_2 and O_3. For details, refer to the corresponding subsections above. However, for both additionally simulated sequences, smaller HCHO dSCDs are retrieved which leads to larger remaining deviations at low elevations for the sequence with an SZA of 45°.

For the here presented example sequence, the standard spectra without Ring effect yield larger deviations from the reference dSCDs than the spectra with Ring effect (compare Sect. 6.4.1). This becomes also evident by comparing the red curves in the two upper panels of Fig. 6.37 with the corresponding ones in Fig. 6.38 which shows the same quantities for the same scenario but for the spectra without Ring effect.

Also for both sequences of additionally simulated spectra not including the Ring effect, smaller HCHO dSCDs are retrieved. However, the absolute values of the differences with respect to the results for the standard spectra are smaller than for the spectra with Ring effect (compare third panels of Fig. 6.37 and 6.38) and values of -0.8×10^{15} and -0.45×10^{15} molec cm^{-2} are found at low elevations for the FP and LR simulations, respectively. While for the LR simulations the reductions are only slightly smaller and still comparable to the ones observed for the spectra with Raman scattering, the reductions are notably smaller for the FP simulations. At 30°, the differences are almost 0 for both scenarios of low resolution simulations. Together with the finding from above that the remaining deviations from the reference dSCDs are higher for the standard spectra without Ring effect, these comparably small reductions of the HCHO dSCDs retrieved for the spectra simulated at low resolution lead to remaining deviations which are still larger than the ones obtained for the standard spectra including the Ring effect (compare red and green lines in the second panels of Fig. 6.37 and 6.38, respectively). Again, the fit RMS is only slightly affected with marginally smaller values for both low resolution simulations.

Smaller HCHO dSCDs compared to the standard spectra without Ring effect are also retrieved for the second additionally simulated (not presented) sequence. Contrary to the spectra which include the Ring effect, the reductions are similar

12 Except for the considerations regarding the AMFs. However, for these, worse agreements were observed for all alternative approaches.

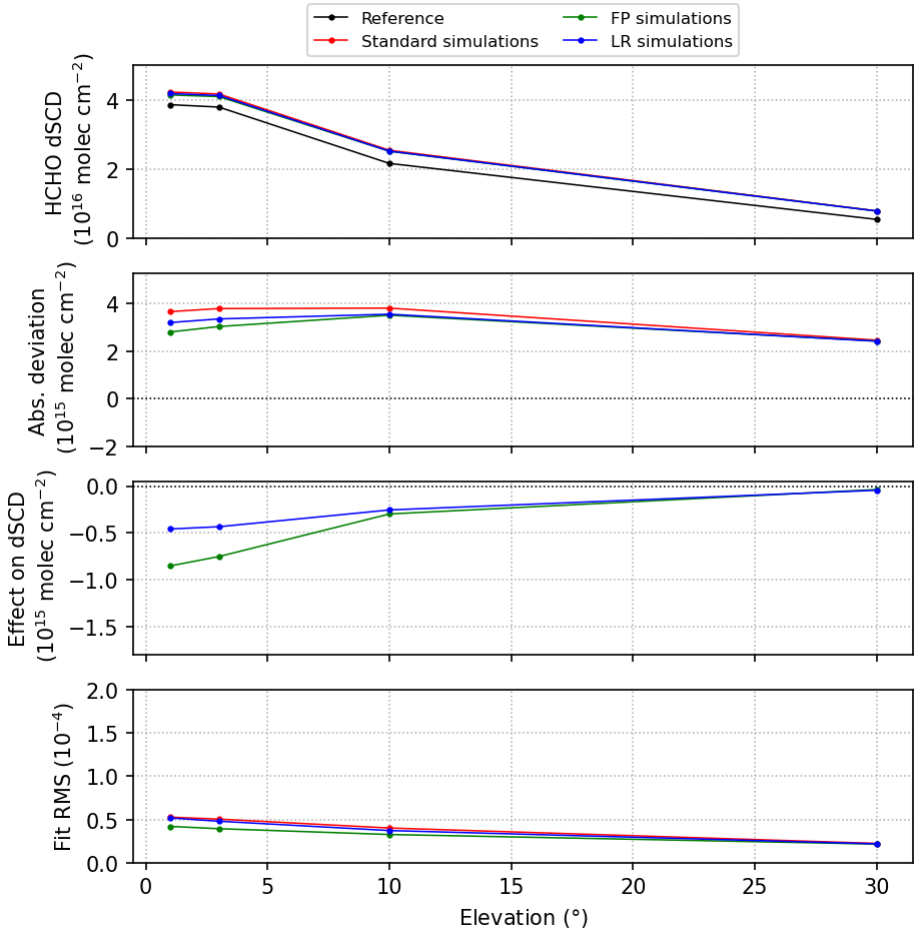


Figure 6.38: Same as Fig. 6.37 but for spectra without Ring effect.

(even marginally larger) to the ones for the scenario presented in Fig. 6.38. Since for the standard spectra without Ring effect at an SZA of 45° and an RAA of 180° lower absolute deviations than for the standard spectra with Ring effect were observed, these reductions in the HCHO dSCDs lead to improved agreements between retrievals and simulations. Especially the HCHO results for the FP simulations show almost perfect agreement at low elevations and 30°, while still non-zero deviations are observed at 10°. Compared to the presented sun geometry, the effects on the fit RMS are larger and smaller values are observed especially for the spectra obtained with the FP simulations.

Summarising the observations in this section, it can be concluded that a certain fraction of the remaining deviations of the HCHO dSCDs retrieved with the optimised fit settings can be attributed to the limitation that a typical DOAS fit does not include all absorption cross-sections (at all relevant temperatures) which are included in the simulations (rather large effect on dSCDs for the FP simulations). This effect applies also to real atmospheric measurements since trace gas absorptions occur at multiple temperatures and only cross-sections at representative temperatures are typically included in the DOAS fits. Parts, but not all, of this effect are compensated by the usage of a sequential FRS. Another,

yet smaller, part of the remaining deviations can be explained by the wrong order of convolution and calculation of the logarithmic ratio between the spectrum to be analysed and the FRS (effect on dSCDs caused by LR simulations is smaller than the one for the FP simulations). Nevertheless, it should be mentioned that only two selected sequences of synthetic spectra were considered here. A last aspect is that the effects on the HCHO dSCDs for the spectra obtained with the low resolution simulations are smaller for the case without Ring effect, especially for the FP simulations. Together, this leads to larger remaining deviations than for the spectra with Ring effect for the presented sequence similar to the findings in Sect. 6.4.1. These findings for the spectra without Ring effect indicate complex compensation effects and interferences which are further discussed in the next section.

6.4.7 *Discussion of remaining deviations*

Several effects potentially causing the observed notable deviations of up to 3.7×10^{15} molec cm^{-2} (mean RMS_{abs} over all scenarios is ca. 1.75×10^{15} molec cm^{-2}) of the retrieved HCHO dSCDs from the corresponding reference dSCDs, which remain for the analysis of the synthetic spectra despite the optimisation efforts for the DOAS fit settings of HCHO, were examined in this section.

These investigations showed that the accuracy and precision settings for the simulations of both synthetic spectra and AMFs (needed for the conversion of the input trace gas VCDs to dSCDs) are sufficient and irrelevant for the remaining discrepancies. Also the wavelength dependence of the AMFs was shown not to be the cause and the approach to interpolate them to the centre of the relevant fitting interval is apparently not a bad choice.

Further, additional Puḳīte terms were tested. Thereby, it was shown that first-order Puḳīte terms for O_4 have almost no effect on the HCHO results (below ca. 2×10^{14} molec cm^{-2}) and for the few cases for which the largest effects were found these even led to slightly worse agreements between retrievals and simulations. At low elevations, additional (higher order) Puḳīte terms for O_3 and NO_2 lead overall to ca. 2×10^{14} to 8×10^{14} molec cm^{-2} smaller HCHO dSCDs except for specific sun geometries which are also problematic for other investigations, e.g. an SZA of 60° with an RAA of 30°. For some sun geometries, this systematic reduction of the HCHO dSCDs leads to improved agreements between retrievals and simulations, while for others a worsening is observed by these unspecific reductions. Especially for spectra not including Raman scattering, strong interferences between the additional Puḳīte terms are observed. Therefore, it was concluded that the additional value of probably too many of such higher order terms is limited and, more importantly, they are not the major cause of the remaining deviations.

Also I_0-corrections (intensity weighted convolutions) for the HCHO, O_4 and BrO cross-sections alone cannot explain the rather large remaining discrepancies. The effect of an I_0-correction of the BrO cross-section on the HCHO results is negligible. For the HCHO and O_4 cross-sections, reduced HCHO dSCDs by up to approximately 7×10^{14} molec cm^{-2} at low elevations are obtained for all sun

geometries which is again an unspecific effect and therefore does not lead to an improvement for all cases. Here, especially for scenarios with small remaining deviations, underestimations occur which is not the desired effect. Further, these effects decrease with increasing elevation, although the largest improvements are "required" at 10° as the deviations are largest at this elevation for many geometries.

For the spectra without noise but still including the Ring effect, again consistently smaller HCHO dSCDs are retrieved. Here, the effects vary from almost 0 to approximately -6×10^{14} molec cm^{-2} and consequently slightly worse agreements between retrievals and simulations than for the standard spectra are possible for specific sun geometries. However, these are smaller than for the I_0-corrections for the HCHO and O_4 cross-sections or the additional Puḳīte terms.

Parts of these reduced HCHO dSCDs are compensated for the spectra which in addition do not include the Ring effect for some geometries. For these spectra, the findings are more complex and show a stronger dependence on the viewing geometry. The absolute values of the effects on the retrieved HCHO dSCDs are comparable to the ones of the noise albeit increases and reductions of the dSCDs can occur with values of $\pm 7 \times 10^{14}$ molec cm^{-2}. Another finding is that for the spectra without Ring effect the elevation dependence of the remaining deviations is reduced for several geometries. Nevertheless, for most geometries still the largest deviations occur at 10° and thus the Ring effect and a possibly insufficient correction during the fit is not the major reason for that. In conclusion, no fundamental issues with the Ring effect and its correction could be identified. Therefore, both the noise and the Ring effect only have a small contribution to the remaining deviations.

The largest effects on the retrieved HCHO dSCDs are found for the low resolution simulations for the spectra which include Raman scattering. The individual effect which can be attributed to the wrong order of convolution and calculation of the logarithmic ratio of the spectrum to be analysed and the FRS, namely remaining inaccuracies of saturation and I_0-effect (LR simulations), is comparable to the effects of noise, Ring effect, I_0-corrections for the HCHO and O_4 cross-sections as well as some additional Puḳīte terms. The effect of the FP simulations, however, is considerably larger. These only include absorption cross-sections which are also used in the DOAS fit and yield an individual absolute effect (excluding the influence of the LR simulations) of up to -1.15×10^{15} molec cm^{-2}. However, only two sequences were considered here and not for both a clear improvement of the agreement between retrieval and simulation was obtained. Further, for the low resolution spectra without Ring effect larger deviations than for the spectra with Ring effect were obtained.

In summary, the following conclusions for the remaining deviations between the HCHO dSCDs retrieved from the synthetic spectra using the optimised fit settings and the corresponding reference values can be drawn:

1. Besides small contributions from noise, the Ring effect and its correction, missing I_0-corrections (intensity weighted convolutions) for the HCHO and O_4 cross-sections as well as remaining non-linear absorptions of NO_2 and O_3, a notable part of the remaining deviations is caused by the fact that

not all absorption cross-sections which are used for the simulations are included in the DOAS fit (large individual effect of the FP simulations). This is similar to real atmospheric measurements and their typical DOAS analyses for which not all atmospheric absorptions and effects can be accounted for.

2. Another, yet smaller part (smaller individual effect of the LR simulations), can be attributed to the wrong order of convolution and calculation of the logarithmic ratio in the DOAS analysis. Also this is similar to real atmospheric DOAS applications and apparently a limitation of the DOAS method.

3. Since for the low resolution simulations the case without Raman scattering leads to worse results for a sun geometry with small RAA, it can be concluded that for such cases, which are dominated by aerosol scattering and thus have large differences of the broad-band spectral dependencies between measurement and FRS, higher order effects become more important, which are not completely corrected. For such cases, also spectral interferences and compensating effects (between different fitted cross-sections and Ring spectra) might thus play a larger role which can lead to better or worse agreements. This is also supported by the other findings regarding the Ring effect. A possible improvement here could be the calculation of Ring spectra for each individual measurement sequence.

Altogether, the remaining deviations are to some extent caused by limitations of the DOAS method itself and the fact that even by considering multiple correction terms not all non-linearities of the inversion problem are corrected. As already mentioned in Sect. 6.2.2.1, the observed remaining deviations occur mostly independently from the trace gas scenario (e.g. high/low NO_2) and thus justify the approach used in that section to find the best-suited combination of NO_2 cross-sections included in the DOAS fits for HCHO. Lastly, it should be stressed that for the high AOD scenarios for many sun geometries rather good agreements between retrieved and reference HCHO dSCDs were found. For such cases, the AMFs are smaller. Hence, the trace gas absorptions are weaker and the described limitations and non-linearities have smaller impacts on the DOAS results for HCHO. Also the effect of rotational Raman scattering (Ring effect) is smaller than for the cases without aerosol scattering. In view of this, it can be concluded that the optimised fit settings are evidently well-suited. An overview of the uncertainties associated with the choice of the optimised fit settings is provided in Sect. 6.7.

6.5 FURTHER FIT PARAMETERS

Besides the main DOAS fit settings described so far, also the influences of further fit parameters on the HCHO results were briefly investigated. These studies were less detailed and included the effects of the fitted intensity offset, the spectral shift and stretch (Sect. 6.5.1), the degree of the broad-band closure polynomial

(Sect. 6.5.2) as well as possible interferences of the absorption cross-sections of HCHO and BrO (Sect. 6.5.3). A short summary of the major findings is presented in this section.

6.5.1 *Intensity offset, spectral shift and stretch*

Typical DOAS analyses include an intensity offset correcting for possible instrumental effects such as detector non-linearities or spectrometer stray-light. Also errors of the offset and dark current correction as well as remaining influences of the Ring effect might contribute here. In the DOAS analyses presented in this work using the *QDOAS* software, the intensity offset is fitted as a non-linear parameter (Danckaert et al., 2017), whereby constant and first-order terms are considered in the standard fit settings (compare Table 6.3). In addition, *QDOAS* offers the possibility to correct for potential misalignments of wavelength dependent quantities in the analysis process such as the involved spectra or the absorption cross-sections (Danckaert et al., 2017). The standard settings allow a shift and a first-order stretch of the wavelength calibration of the spectrum to be analysed but not for other quantities included in the fit (compare Table 6.3). Both fit parameters are in theory not needed for the analysis of synthetic spectra since the intensities of the spectra (except small remaining differences caused by an imperfect Ring correction) and the wavelength calibration are exactly correct for these spectra. Nevertheless, tests showed that the fitting of an intensity offset as well as the correction of slight misalignments in the wavelength calibrations are essential for real measured spectra. For instance, the used AVANTES spectrometers are known to have a non-linear detector response (compare Sect 4.2.1.2) and the analysis of long time series using one fixed wavelength calibration requires the possibility to account for slight changes in the spectral properties of the instrument. Also the so-called tilt effect can cause an effective wavelength shift between the low elevation spectra and the FRS (Lampel et al., 2017b). Therefore, an intensity offset as well as a spectral shift and stretch are considered in the standard fit settings summarised in Table 6.3 which are also used for the analyses of the synthetic spectra in the here-presented sensitivity study.

In this section, the effects of these parameters on the HCHO results are studied and uncertainties caused by them are examined. For that, besides the analysis using the optimised fit settings, HCHO retrievals without fitting an intensity offset, not allowing a spectral shift and stretch as well as a retrieval not including any of the mentioned parameters were performed. This was done for spectra with and without noise for the 36 selected sequences from above (compare Sect. 6.2.2.1). The HCHO results obtained using each of the three described fit scenarios are compared to the results of the standard fit scenario in Fig. 6.39 which presents the retrieved HCHO dSCDs (uppermost panel), the corresponding deviations from the reference dSCDs (second panel), the effects of the different modified fit settings on the HCHO dSCDs (third panel) as well as the fit RMS (lowermost panel) as functions of elevation for spectra without noise in the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°.

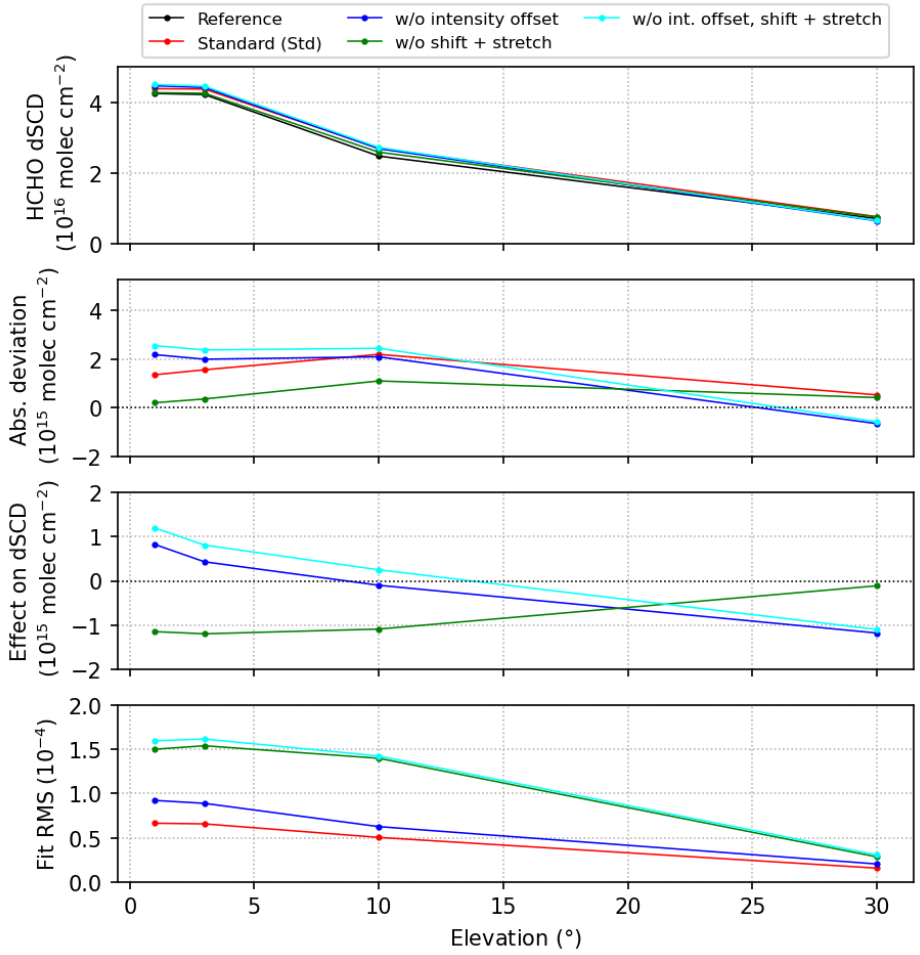


Figure 6.39: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effects of the different modified fit settings (see text) on the HCHO dSCDs and fit RMS as functions of elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. Results obtained for the different fit scenarios are indicated by different colours.

For the presented example, the fit scenario not including an intensity offset (blue lines) yields larger HCHO dSCDs at low elevations and smaller dSCDs at 30° leading to larger deviations from the reference dSCDs at these elevations, while almost no difference with respect to the optimised fit settings is found at 10°. The absolute effect on the HCHO dSCDs varies between roughly $\pm 1 \times 10^{15}$ molec cm^{-2}. Another finding is that the fit RMS is notably enhanced for this fit scenario at all elevations indicating worse spectral fits (compare lowermost panel). These general patters are true for almost all other (not presented) sequences, whereby for some geometries larger dSCDs are obtained at all elevations. The absolute effect on the retrieved HCHO dSCDs varies rather strongly between the different viewing geometries (specific combination of SZA, RAA and elevation angle) with values ranging from approximately -2×10^{15} to 4.5×10^{15} molec cm^{-2}. For a given sun geometry, the effect on the HCHO results differs only weakly between the different trace gas scenarios. While the absolute effect at low elevations is

slightly larger for the low HCHO case, the effect is slightly larger for the high NO_2 scenario.

On the first view, the fit scenario not allowing a spectral shift and stretch (compare green lines) yields HCHO dSCDs which deviate clearly less from the reference dSCDs. This reduction of the dSCDs is largest at low elevations where roughly 1.2×10^{15} molec cm^{-2} smaller dSCDs are retrieved and decreases to almost 0 at an elevation of 30°. However, at the same time, the fit RMS is significantly larger for this fit scenario and around twice as high as for the standard fit scenario (compare lowermost panel). Since the comparisons in Fig. 6.39 are based on spectra without noise, these larger fit RMS-values are caused by systematic spectral structures in the residuals of the DOAS fits introduced by the missing spectral shift and stretch. This becomes also evident in Fig. 6.40 which compares the fit residuals (left column) and the corresponding fits of the HCHO cross-section (right column) of both fit scenarios obtained for the spectrum at 1° of the example sequence presented in Fig. 6.39.

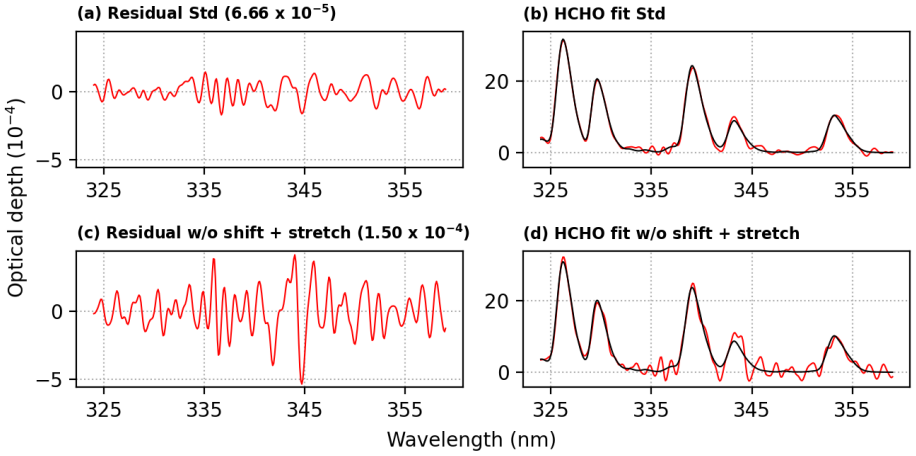


Figure 6.40: Residuals (left column) and fitted HCHO cross-section (right column) for the 1° spectrum of the example sequence shown in Fig. 6.39. Panel (a) shows the fit residual for the standard fit scenario, while panel (c) presents the fit residual for the case when no spectral shift and stretch are allowed. Panels (b) and (d) show the corresponding fits of the HCHO cross-section, whereby black and red lines indicate the fitted cross-section and the fitted cross-section plus residual, respectively.

Obviously, the amplitude of the spectral structures in the residual spectrum is markedly larger for the fit scenario without spectral shift and stretch. In consequence, these large spectral residual structures lead to a worse fit of the HCHO absorption cross-section which in turn results in a smaller HCHO dSCD which is 4.38×10^{16} molec cm^{-2} and 4.27×10^{16} molec cm^{-2} for the optimised fit settings and the fit settings without shift and stretch, respectively. This reduction finally leads to a smaller deviation from the reference dSCD which is 4.25×10^{16} molec cm^{-2} for the presented example. Most likely, these prominent spectral residual structures are caused by the tilt effect leading to an effective wavelength shift between the low elevation spectra and FRS which does not occur if a spectral shift and stretch is allowed in the DOAS retrieval (Lampel et al., 2017b). In

conclusion, the better agreement between retrieved and reference dSCDs is likely not caused by real improvements of the spectral retrieval but rather pronounced spectral residual structures leading to fits representing the simulated HCHO absorptions less well. This conclusion is also supported by the finding that with the fit settings not allowing a spectral shift and stretch smaller HCHO dSCDs than with the optimised standard settings are retrieved for all sequences independent of the specific deviations from the reference dSCDs. Similar to some of the findings in Sect. 6.4, this unspecific reduction leads to worse agreements between retrievals and simulations for sequences showing only small deviations (slight over- or underestimations) for the optimised fit settings, e.g. spectra with an SZA of 20° at an RAA of 90°. Like for many of the other effects, the general patterns and absolute values of the effects on the HCHO results are quasi independent of the trace gas scenario for a given sun geometry, whereas especially the absolute values vary notably between the different viewing geometries defined as specific set of SZA, RAA and elevation angle (from ca. 0 to almost -3×10^{15} molec cm^{-2}).

The HCHO results for the fit scenario which includes neither of the two parameters are comparable to the ones obtained with the scenario without intensity offset, while the changes in the fit RMS are dominated by the effect of the missing spectral shift and stretch (cyan lines in Fig. 6.39). Together, worse agreements compared to the optimised fit settings are obtained with this fit scenario. Again, these overall tendencies are true for all other (not presented) sequences with the absolute effects mainly varying between different viewing geometries, while only small fluctuations between the HCHO and NO_2 scenarios are found.

The findings described in this section also hold for the spectra containing noise. More precisely, for a given sequence of synthetic spectra similar effects of the different modified fit scenarios on the HCHO results are obtained for spectra with and without noise. Naturally, the absolute values of the effects differ slightly and especially the relative increase of the fit RMS (compare above) is, with a factor below ca. 1.5, smaller for the spectra containing noise since already the noise elevates the RMS-values significantly.

As already mentioned above, both intensity offset as well as spectral shift and stretch should in theory not be needed for the analysis of synthetic spectra (except to account for the tilt effect), while they are indispensable for the analysis of real measured spectra. However, given the findings in this section, the usage of an intensity offset is apparently also important for the analysis of synthetic spectra and worse agreements between retrieved and reference HCHO dSCDs are obtained if no intensity offset is fitted. Although the deviations from the true reference dSCDs are clearly reduced for many scenarios if a spectral shift and stretch of the spectrum to be analysed is not allowed in the fit, for measured spectra, a shift and stretch has to be allowed (even if a tilt correction spectrum is included in the fit). Therefore, it was decided to include intensity offsets and spectral shifts and stretches in both the analysis of synthetic spectra for the here-presented sensitivity study[13] and the optimised fit settings for HCHO used

13 Since this sensitivity study was an iterative process, similar investigations on the intensity offset and the spectral shift and stretch were already performed before the optimisation steps presented above. Here, these "older" findings are just confirmed using the final optimised fit settings.

in the following chapters. The influence of the tilt effect on the HCHO analyses should be investigated in further studies.

6.5.2 *Polynomial degree*

To account for broad-band effects in the spectra of scattered sunlight, e.g. caused by Rayleigh and Mie scattering on molecules and aerosol, respectively, the DOAS fit typically includes a low order closure polynomial[14] (compare Sect. 3.1.3 and 3.1.4). For that, different choices of the polynomial degree are possible. On the one hand, the degree should be high enough to correct the broad-band effects sufficiently well. On the other hand, a too high polynomial degree could already interfere with or account for parts of the highly structured narrow-band trace gas absorptions. The here-proposed optimised fit settings for the HCHO retrieval include a fifth-order polynomial (compare Table 6.3).

To examine in how far this choice is justified, again spectra with and without noise for the 36 selected scenarios from above were analysed using the optimised fit settings and varying the polynomial degree between three and six. The HCHO dSCDs retrieved with these different fit scenarios are compared in Fig. 6.41 which shows the retrieved HCHO dSCDs (uppermost panel), the corresponding deviations from the reference dSCDs (second panel), the effects of the three tested polynomial degrees on the retrieved HCHO dSCDs (third panel) as well as the fit RMS (lowermost panel) as functions of elevation for spectra without noise in the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°.

For the presented example, the increase of the polynomial degree to six leads to larger HCHO dSCDs and thus larger deviations from the reference dSCDs at all elevations (cyan lines). Although the fit RMS is marginally reduced, this higher polynomial degree leads overall to worse HCHO results. This is also true for almost all other not presented geometries, whereby the absolute effects on the retrieved HCHO dSCDs vary from almost 0 to ca. 1.6×10^{15} molec cm^{-2}. Only two exceptions from this general behaviour exist for spectra at SZAs of 45° and 60° with an RAA of 30° for which slight reductions of the retrieved HCHO dSCDs (below ca. -0.5×10^{15} molec cm^{-2}) and thus improved agreements between retrievals and simulations are observed. Further, almost no fluctuations between the different trace gas scenarios are found for a given sun geometry.

At elevations up to 10°, also the usage of a third-order polynomial leads to larger HCHO dSCDs and thus larger deviations from the reference dSCDs (blue lines). At 30°, smaller dSCDs are retrieved with a third-order polynomial. However, since the reduction is around 0.7×10^{15} molec cm^{-2}, the dSCDs are underestimated in that case by a slightly smaller amount than they are overestimated with the standard fifth-order polynomial (compare blue and red lines in the second panel). At the same time, the fit RMS is increased by ca. 10 to 15 % indicating overall worse HCHO results for the example sequence using a third-order polynomial. Moreover, large variability of the effect on the retrieved HCHO dSCDs and the fit RMS between the different sun geometries is found. For instance, up to 5.3×10^{15} molec cm^{-2} larger dSCDs corresponding to absolute

14 Also instrumental effects can have broad-band dependencies.

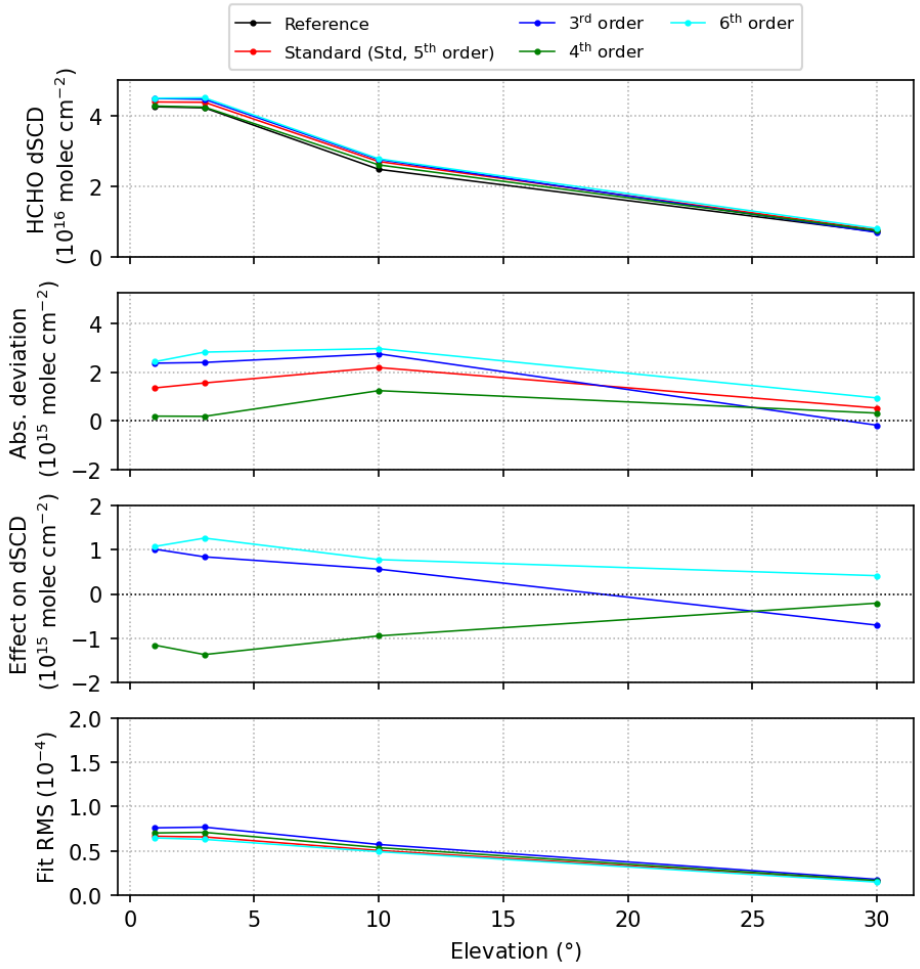


Figure 6.41: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effects of the different polynomial degrees on the HCHO dSCDs and fit RMS as functions of elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. Results obtained for the different fit scenarios are indicated by different colours.

deviations of approximately 7.5×10^{15} molec cm^{-2} (for the low HCHO scenario) are obtained at low elevations for spectra at an SZA of 60° and an RAA of 30°. In contrast to that, up to roughly 2×10^{15} molec cm^{-2} smaller dSCD are retrieved at 30° for spectra with an SZA of 20° and an RAA of 180° yielding an underestimation of the reference dSCDs by 2.6×10^{15} molec cm^{-2}. Almost no effect at all is found at low elevations for spectra with an SZA of 20° and an RAA of 30°. Overall, for neither sun geometry improved agreements between retrievals and simulations are found. Similar to the findings for the usage of a sixth-order polynomial, only small differences between the different trace gas scenarios are found for a given sun geometry, whereby the effects tend to be slightly larger at low elevations for the high NO_2 and low HCHO scenarios.

A decrease of the polynomial degree to four results in reduced HCHO dSCDs at all elevations, whereby the effect is largest at low and smallest at high elevations. Consequently, notably smaller deviations of the retrieved dSCDs from

the reference values are obtained for the example sequence. Another finding is that the elevation dependence of the absolute deviations is larger for the results obtained with the fourth-order polynomial. Like for the third-order polynomial, the fit RMS is increased, however, as expected the worsening is smaller for the fourth-order polynomial (below roughly 10 %). Despite the overall promising results for the presented and some of the other not presented sun geometries, worse results or almost no effects on the HCHO dSCDs are obtained with the reduced polynomial degree for several sun geometries. In particular, the reduction of the retrieved dSCDs is again rather unspecific (compare some of the effects in Sect. 6.4) leading to underestimations for sequences showing only small over- or underestimations for the optimised fit settings, e.g. spectra with an SZA of 20° at an RAA of 180°. For spectra with an SZA of 60° at an RAA of 30° for which the largest deviations from the reference dSCDs are obtained with the optimised fit settings, the effect is practically 0 at all elevations. Finally, the fluctuations between the different trace gas scenarios are again comparably small for a given sun geometry. Altogether, the usage of a fourth-order polynomial leads to better or worse results depending on the specific scenario.

The described behaviours of the HCHO results obtained using different polynomial degrees are also true for the spectra with noise. Naturally, the absolute values of the effects differ slightly and especially the relative effects on the fit RMS are smaller since already the noise elevates the RMS-values significantly.

In conclusion, the usage of a fifth-order polynomial as suggested in the optimised fit settings is apparently a reasonable choice. While the increase of the polynomial degree to six results in overall worse results for almost all sequences of synthetic spectra, large variability of the effects on the retrieved HCHO dSCDs between the different sun geometries is found if a third-order polynomial is used. For the latter, no improved agreement between retrievals and simulations is found for any scenario. Further, for spectra with an SZA of 60° at a small RAA of 30°, which is probably dominated by aerosol scattering and thus shows large differences of the broad-band spectral dependencies between the spectrum to be analysed and the FRS, largely increased deviations are found. This indicates that a polynomial degree of three is probably too low to sufficiently account for such broad-band spectral features[15]. Lastly, for some sun geometries promising improvements of the agreements between retrieved and reference dSCDs are obtained using a fourth-order polynomial. However, the observed reduction of the retrieved dSCDs is apparently rather unspecific leading to worse agreements, in terms of underestimations, for scenarios which show only small over- or underestimations for the optimised fit settings. In conclusion, still the usage of a fifth-order polynomial is recommended in this work. Although this fit scenario yields not the lowest deviations from the reference dSCDs for some sequences (compare Fig. 6.41), it leads to rather small fluctuations between the different sun geometries and thus to rather stable results in terms of mostly overestimations. As consistency is a major goal of this sensitivity study, this is a strong argument in favour of this choice.

15 Similar findings hold for an SZA of 45°, whereby the absolute effect is smaller.

6.5.3 *BrO absorption*

In order to account for (mostly stratospheric) BrO absorptions, the standard fit settings include a BrO cross-section (compare Table 6.3) and also the simulations of the synthetic spectra considered a stratospheric BrO column (compare Table 6.1). However, in the spectral region relevant for the HCHO retrieval, the absorption cross-sections of HCHO and BrO (at typical spectral resolutions of MAX-DOAS instruments) show rather similar spectral features. The normalised cross-sections of both absorbers in that spectral region are shown in Fig. 6.42.

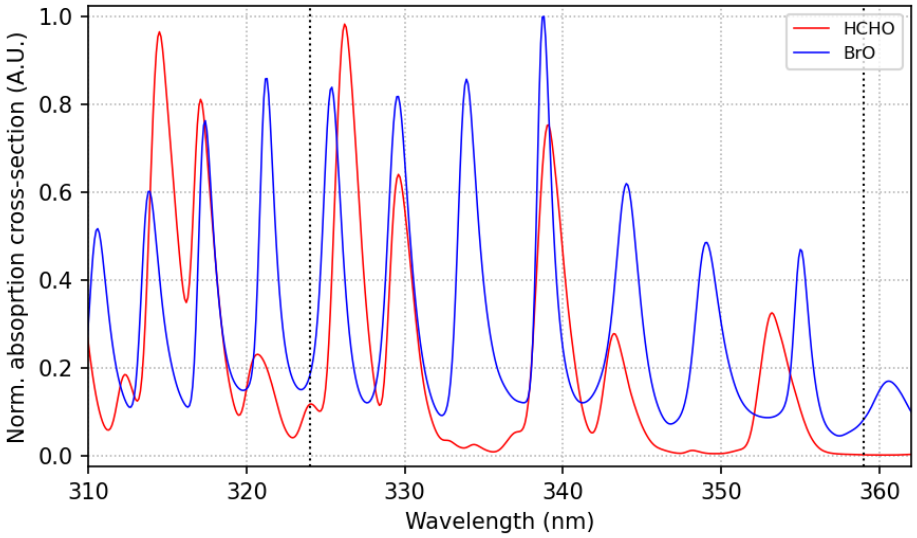


Figure 6.42: Absorption cross-sections of HCHO and BrO normalised to their respective maxima in the spectral range relevant for the HCHO retrieval at typical spectral resolution of MAX-DOAS instruments. The fitting range proposed in this work is indicated by vertical black dashed lines.

Obviously, several absorption bands of both species overlap which is also true for the fitting interval proposed in this work (compare vertical black dashed lines). This overlap potentially leads to spectral interferences affecting the HCHO results. To examine the impact of such potential interferences on the retrieved HCHO dSCDs, retrievals with and without BrO cross-section were performed for the same 36 selected sequences from above (compare Sect. 6.2.2.1) using the optimised fit settings besides this modification. Again, this was done for spectra with and without noise. The HCHO results obtained with both fit scenarios are compared in Fig. 6.43 which depicts the retrieved HCHO dSCDs (uppermost panel), the corresponding deviations from the reference dSCDs (second panel), the effect on the HCHO dSCDs (third panel) as well as the fit RMS (lowermost panel) as functions of elevation for spectra without noise in the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°.

While slightly smaller HCHO dSCDs, and thus smaller deviations from the reference dSCDs, are obtained at 1° and 3° elevation for the scenario without BrO cross-section included in the fit, larger dSCDs are retrieved at 30° corresponding to larger deviations. Almost no effect is found at 10°. These observations suggest that

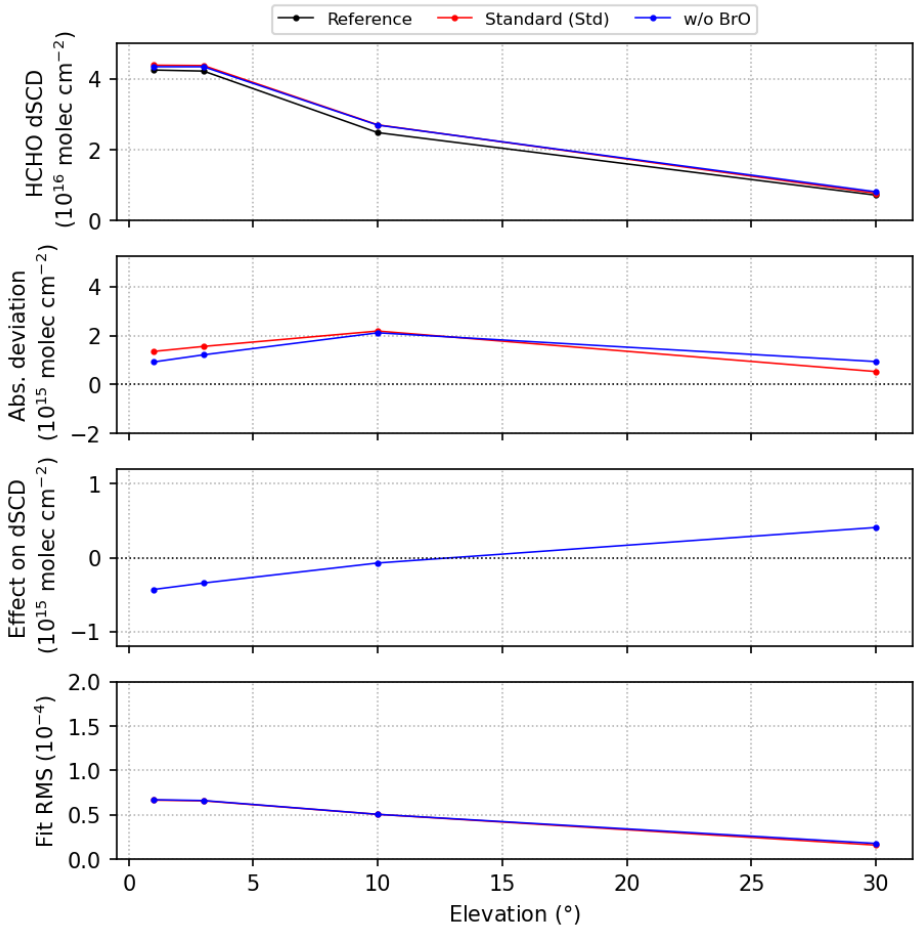


Figure 6.43: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, effect on the HCHO dSCDs and fit RMS as functions of elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. Results obtained with and without BrO cross-section included in the fit are indicated by different colours.

at low elevations the (simulated) stratospheric BrO absorptions are not important and the inclusion of the BrO cross-section leads to small spectral interferences with the HCHO cross-section yielding slightly higher HCHO dSCDs. Contrary to that, the usage of the BrO absorption cross-section is well justified at 30° and worse results are obtained if it is not included in the fit. For the presented example, the absolute effects are around $\pm 0.4 \times 10^{15}$ molec cm^{-2} and are of similar magnitude as most of the effects discussed in Sect. 6.4. The influence on the fit RMS is very small and can be neglected at low elevations, while the RMS is ca. 10 % larger for the scenario without BrO cross-section at 30°.

These findings are practically independent of the trace gas scenario and only the absolute values vary faintly. Larger differences are found between the different sun geometries. Here, for low SZAs (20°) mainly larger HCHO dSCDs and thus larger deviations from the reference dSCDs are obtained if the BrO cross-sections is not included in the DOAS fit, whereas at an SZA of 60° smaller dSCDs leading to slightly improved agreements are retrieved in that case for many

scenarios. At maximum, the absolute effects on the HCHO dSCDs vary between ca. $\pm 1 \times 10^{15}$ molec cm^{-2} depending on the specific sun geometry corresponding to relative effects of ca. 2.5 % and 6 % at low elevations for the high and low HCHO scenario, respectively. Basically the same findings are true for the spectra containing noise.

Given these findings and the fact that the observed effects on the HCHO results do not indicate major artefacts caused by the inclusion of the BrO cross-section in the DOAS retrieval of HCHO, it was decided to stay with the usage of the BrO cross-section. This might also be justified since also real measured spectra can contain (mostly stratospheric) BrO absorptions under specific measurement conditions. Nevertheless, it must be noted that viewing geometries exist for which the retrieved HCHO dSCDs have a (small) systematic high bias caused by the BrO cross-section which can also occur for real atmospheric measurements.

6.6 COMPARISON WITH OTHER COMMON FIT SETTINGS

In this section, the HCHO results obtained with the optimised fit settings based on the presented sensitivity study are compared to the results obtained with fit settings used by other groups in the DOAS community, namely the DOAS groups at the Royal Belgian Institute for Space Aeronomy (BIRA) and the Institute for Environmental Physics Bremen (IUP-Bremen) at the University of Bremen. Table 6.4 provides an overview of the retrieval settings suggested in this work and the ones used at IUP-Bremen (Alvarado, 2020) and BIRA (Pinardi, 2020)[16]. Besides small differences in the fitting intervals and the utilised absorption cross-sections, the main differences are the usage of Puḳīte terms for NO_2 and O_3 as well as the approach to account for the Ring effect. As described above (compare Sect. 3.1.5.1 and 6.1.2), for the latter two pseudo cross-sections derived from the 90° spectrum of one selected sequence are included in the fits in this work while both other groups apply a different approach using the ratio of a synthetic simulated Raman spectrum and a solar spectrum after convolving both spectra with the instrumental slit function.

To compare the performances of the different fit settings, noise-free spectra of the 36 selected sequences of synthetic spectra from above were analysed using the DOAS fit settings from Table 6.4 and visualised in the same way as before. As an example, Fig. 6.44 presents the resulting HCHO dSCDs (uppermost panel), the corresponding deviations from the true reference dSCDs (second panel), the differences of the results using the IUP-Bremen and BIRA settings with respect to the ones from the optimised fit settings (third panel) and the fit RMS (lowermost panel) for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°.

As seen earlier, for the presented scenario the optimised fit settings lead to an overestimation of the reference HCHO dSCDs by ca. 1.5×10^{15} molec cm^{-2} at low elevations. In contrast to that, the IUP-Bremen and BIRA settings yield around 3.3×10^{15} molec cm^{-2} smaller dSCDs corresponding to an underestimation of the reference dSCDs by roughly 1.8×10^{15} molec cm^{-2}. At 10° elevation, only

16 This comparison was performed at an early stage of this study in 2020. Hence, the fit settings used by both groups might have changed in the meantime.

Table 6.4: Comparison of the HCHO DOAS fit settings suggested in this work with the ones used by the DOAS groups at IUP-Bremen and BIRA.

PARAMETER	THIS WORK	IUP-BREMEN	BIRA
Fitting interval	324 – 359 nm	323.5 – 361 nm	324.5 – 359 nm
Trace gas			
HCHO	Meller and Moortgat (2000), 298 K		Meller (1992), 297 K
BrO	Wilmouth et al. (1999), 228 K	Fleischmann et al. (2004), 223 K	
NO_2	Vandaele et al. (1998), 294 K and 220 K	Vandaele et al. (1998), 220 K	Vandaele et al. (1996), 298 K
	plus linear and square Puķīte terms	–	–
O_3	Serdyuchenko et al. (2014)		
	223 K and 243 K	223 K and 293 K	223 K and 243 K
	plus linear and square Puķīte terms		–
O_4	Thalman and Volkamer (2013), 293 K		
Ring effect	Two pseudo cross-sections derived from universal FRS	Ratio of simulated Raman spectrum and solar spectrum, both convolved to instrumental resolution	
FRS	90° spectrum of corresponding sequence/scenario		
Polynomial	5^{th} order		
Intensity offset	Constant and 1^{st} order term		
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS		

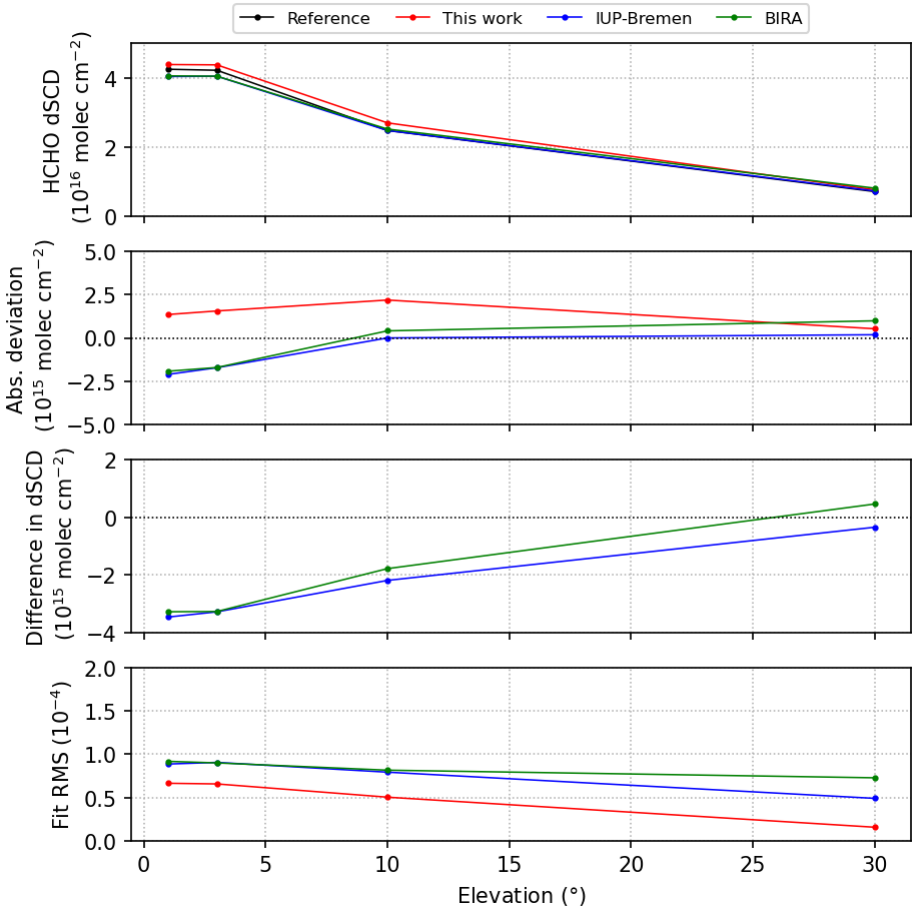


Figure 6.44: Retrieved HCHO dSCDs, absolute deviations from the true reference dSCDs, differences of the HCHO dSCDs obtained with the different fit settings (see text) with respect to the ones from the optimised fit settings and fit RMS as functions of elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. Results obtained with the different settings are indicated by different colours.

small deviations from the reference values are found for the IUP-Bremen and BIRA settings, whereas still a significant overestimation is obtained with the optimised settings. At 30°, all three fit settings agree reasonably well within ca. 1×10^{15} molec cm^{-2} with a maximum deviation of about 1×10^{15} molec cm^{-2} for the BIRA settings. Apparently, none of the fit settings can reproduce the reference dSCDs at low elevations well, while the IUP-Bremen and BIRA settings perform better already at 10°. However, their elevation angle dependencies of the absolute deviations from the reference dSCDs are stronger compared to the optimised fit settings. In addition, the dSCDs are underestimated at low and slightly overestimated at high elevations. As described earlier, such a behaviour not only affects the resulting VCDs but also the retrieved profile shapes. In that aspect, the optimised fit settings perform better since overestimations occur at all elevations, most likely only affecting the retrieved VCDs and the absolute values of the HCHO profile. Another observation is that the fit RMS is notably

smaller for the fit settings proposed in this work with a factor of up to 4.5 at 30° compared to the BIRA settings (green line in lowermost panel).

The finding that notably smaller HCHO dSCDs are retrieved at low elevations with the IUP-Bremen and BIRA fit settings as well as the observed elevation dependencies of the absolute deviations (second panel) are valid for all investigated sun geometries, whereby generally smallest dSCDs are retrieved with the IUP-Bremen settings. Naturally, the absolute values of these effects differ amongst the different sun geometries. Although the differences between the different sun geometries are rather small for all three fit settings, the variations are larger for the two alternative DOAS fit settings. Since for all sun geometries notably smaller dSCDs are retrieved at low elevations with the IUP-Bremen and BIRA settings, rather strong underestimations are obtained with these settings for sequences for which relatively good agreements between retrievals and simulations are found with the optimised fit settings, e.g. in the case of an SZA of 20° and an RAA of 180°.

While for a given sun geometry the findings are almost independent of the HCHO scenario, significant differences are found between the low and high NO_2 scenarios. To illustrate this, Fig. 6.45 shows the same quantities as Fig. 6.44 but for the corresponding high NO_2 scenario.

Obviously, the differences between the HCHO dSCDs retrieved with the different fit settings at low elevations are markedly larger than for the low NO_2 scenario. In particular, the absolute deviations observed for the optimised fit settings are comparable for both NO_2 scenarios while the underestimations are around 14×10^{15} and 5×10^{15} molec cm^{-2} at 1° for the IUP-Bremen and BIRA settings, respectively, and thus ca. 7 and 2.5 times larger than for the low NO_2 case. At 10°, the absolute deviations are comparable to the ones of the low NO_2 scenario for the BIRA settings, whereas for the IUP-Bremen settings still notably larger deviations are obtained. Similar to Fig. 6.44, the results of all three fit scenarios agree reasonably well at 30°. Together, this leads to even stronger elevation dependencies of the absolute deviations for the IUP-Bremen and BIRA settings compared to the low NO_2 scenario. Consistent to the preceding investigations on the different combinations of cross-sections (compare patterns in Fig. 6.5 in Sect. 6.2.2.1), it can be concluded that the general patterns of the results obtained with the two alternative fit settings for the high NO_2 scenario can be explained by missing Puķīte terms for NO_2 (compare Table 6.4). While for the BIRA settings the absolute deviations are comparable to the ones from this sensitivity study (compare Fig. 6.5 and 6.45), significantly larger values are obtained for the IUP-Bremen settings. The latter can be explained by the fact that these settings include an NO_2 cross-section at 220 K instead of cross-sections at 294 K and 298 K for the optimised and the BIRA settings, respectively. Obviously, this cross-section does not represent the tropospheric NO_2 absorptions, which occur at temperatures higher than 220 K, sufficiently well. Besides the choice of the correct temperature for the NO_2 cross-section, these findings demonstrate that the usage of NO_2 Puķīte terms is essential under high NO_2 conditions. This is also reflected by the strongly increased fit RMS-values at low elevations for the IUP-Bremen and BIRA settings caused by strong NO_2 residual structures (lowermost panel of Fig. 6.45).

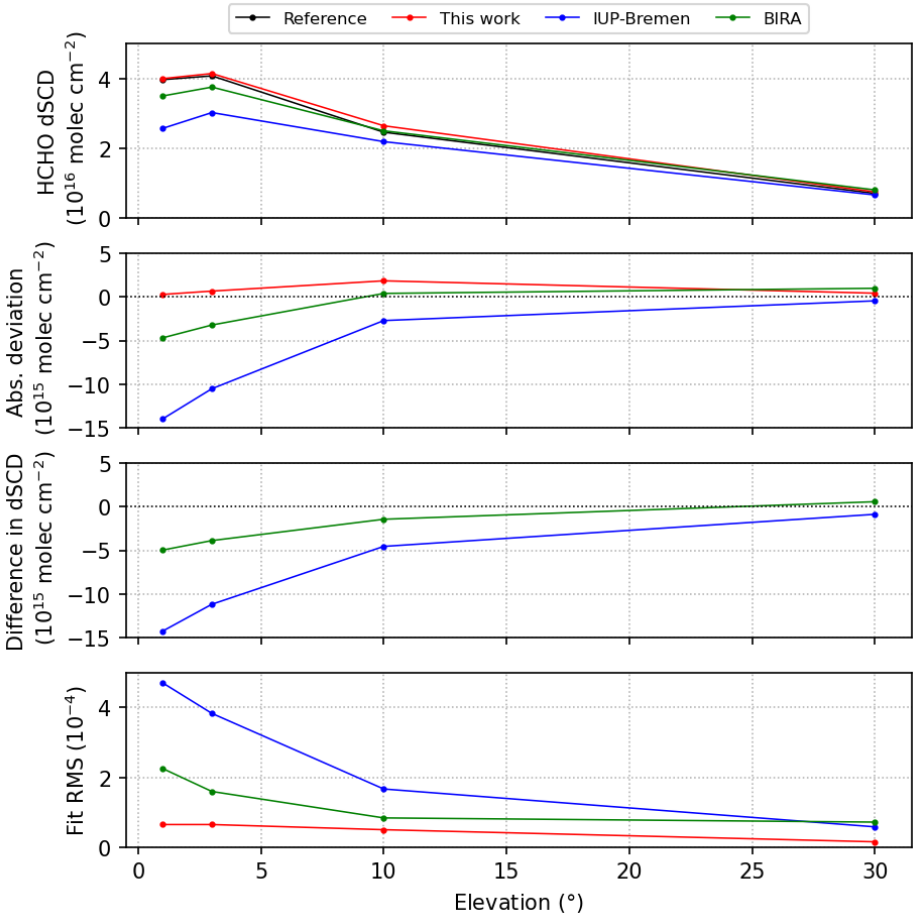


Figure 6.45: Same as Fig. 6.44 but for the case of high NO_2. Note the different y-axes compared to Fig. 6.44 in the second, third and lowermost panels.

Again, these findings are practically independent of the HCHO scenario, while the results differ more between the different sun geometries.

Another interesting aspect is that for some sun geometries the BIRA settings yield the largest fit RMS and partly the largest absolute deviations of all three settings at 30° (compare Fig. 6.44). For the low NO_2 scenario, this is also true for the fit RMS at low elevations. This is probably caused by missing Puḳīte terms for O_3 in the BIRA settings as these lead to the largest effects at high elevations (compare Sect. 6.2.3).

As mentioned above, a major difference between the fit settings based on this work and the settings used at IUP-Bremen and BIRA is the way how the Ring effect is accounted for. To investigate the influence of these different approaches on the comparison results, spectra without Ring effect were analysed. Figure 6.46 shows the corresponding comparisons for the same sequence as above in the case of low NO_2.

Apparently, the different ways of accounting for the Ring effect explain large parts of the discrepancies between the HCHO results obtained with the different fit settings since at all elevations overestimations of the HCHO dSCDs are observed for all three fit settings. Especially, at low elevations the HCHO

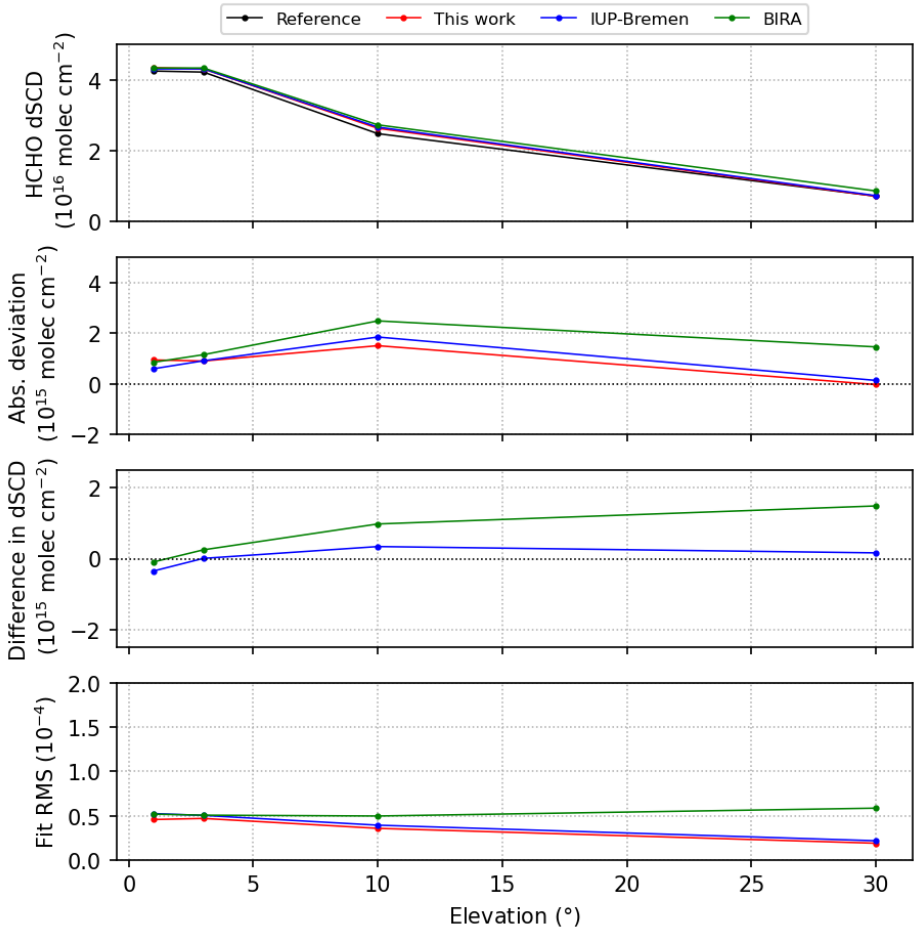


Figure 6.46: Same as Fig. 6.44 but for spectra without Ring effect. Note the different y-axes compared to Fig. 6.44 in the second and third panel.

dSCDs retrieved with the different DOAS settings agree well within ca. 0.5×10^{15} molec cm^{-2}. At 10°, a larger spread between the different dSCDs is found which is yet smaller than for the spectra with Ring effect. While the results obtained with the optimised and the IUP-Bremen fit settings agree reasonably well also at 30°, the dSCDs retrieved with the BIRA settings deviate stronger by about 1.5×10^{15} molec cm^{-2} which is a similar magnitude as for the spectra with Ring effect. The latter can be again attributed to missing O_3 Puķīte terms in the BIRA settings which is also represented by comparably large fit RMS-values at 10° and 30° elevation (compare green line in lowermost panel). These findings are generally also true for the other not presented scenarios except for SZAs of 45° and 60° at an RAA of 30° (compare above), whereby again for the high NO_2 scenario large negative deviations are found at low elevations for the IUP-Bremen and BIRA settings caused by missing Puķīte terms for NO_2. A last finding is that for the spectra without Ring effect the variations of the absolute deviations between the different sun geometries are similar for all three sets of fit settings with the BIRA settings being most stable for the spectra without Raman scattering.

Altogether, none of the different sets of fit settings yields perfect agreement between retrievals and simulations. The comparison results stress the importance of the inclusion of Puķīte terms for NO_2 and the choice of an NO_2 cross-section at a representative temperature for high NO_2 conditions. In this regard, the fit settings proposed in this work perform clearly better than the IUP-Bremen and BIRA settings especially at low elevations. Further, the comparisons confirm that for several sun geometries the additional usage of Puķīte terms for O_3 can help to improve the fit RMS and the agreement with the reference dSCDs at 10° and 30° elevation. Also, the elevation dependence of the absolute deviations within a given sequence and the variation of them between the different sun geometries are smaller for the optimised fit settings indicating overall rather stable results in terms of acceptable overestimations of the reference dSCDs[17]. For the other two sets of fit settings, under- and overestimations occur mostly at low and high elevations, respectively, which in consequence affects both retrieved VCDs and profile shapes, whereas for the optimised fit settings probably mainly the absolute values of the VCDs are influenced. In conclusion, the comparison results show that one major aim of the presented sensitivity study is fulfilled as the optimised fit setting perform equally well for many different atmospheric conditions. Lastly, the comparison results for the spectra without Ring effect suggest that the different approaches to account for the Ring effect are, besides the usage of Puķīte terms for NO_2 and O_3, the main reason for the discrepancies between the HCHO results obtained with the settings suggested in this work and the two sets of fit settings used by the DOAS groups at IUP-Bremen and BIRA.

6.7 MAIN FINDINGS AND UNCERTAINTY ESTIMATION

Many different aspects have been evaluated and discussed in the here-presented sensitivity study for the DOAS fit settings of HCHO. To provide an overview of these discussions, a brief summary of the main findings is given in this section. This includes an estimate of the uncertainties of the HCHO results based on the chosen optimised fit settings and recommendations for fit settings for specific measurement conditions or specific requirements on the results.

The first addressed aspect was the correct and consistent treatment of the Puķīte terms (and Raman cross-sections) for NO_2 and O_3 absorptions and their corresponding "normal" cross-sections (compare Sect. 6.2.1). Here, for the order of calculation and convolution, the recommendation is to calculate the Puķīte terms from the HR cross-sections before the convolution since also their derivation is performed in HR space (see also Puķīte et al., 2010). Further, the Puķīte terms and their corresponding "normal" cross-section have to be treated in a consistent way, whereby the usage of an intensity weighted convolution according to Puķīte et al. (2021) for both is recommended (compare Eq. 3.25). A classical I_0-corrected convolution for both is only possible after the typical SCDs and the cross-sections are scaled in an appropriate way. This approach, however, is mathematically not correct and therefore not recommended. Lastly, in this work the usage of non-normalised linear Puķīte terms for NO_2 and O_3 (compare Sect. 6.2.1 and

17 Only for a few sun geometries slight underestimations are found as outlined earlier.

Eq. 3.16 in Sect. 3.1.5.2) is suggested since no arbitrary normalisation is needed for these. Moreover, highly consistent results are obtained with normalised and non-normalised linear Puḳīte terms. These choices were already considered for the further optimisation steps.

Naturally, the main aspect was the optimisation of the DOAS fit settings of HCHO. For that, mainly the included cross-sections of NO_2 and O_3 as well as the applied fitting interval were examined starting at standard fit settings summarised in Table 6.3. Thereby, synthetic as well as measured spectra (compare Sect. 6.1.1) were used. This combination allows to minimise the discrepancy between retrieved and reference dSCDs and to test the optimised fit settings under real measurement conditions. For the final selections of the different fit settings, a trade-off between two important aspects was made. Firstly, the main aim is of course to optimise the fit settings in a way that the deviations of the retrieved HCHO dSCDs from their respective true reference values are minimised. Secondly, however, a major aim of this study was to optimise the fit settings in a manner that consistent results are obtained for many different atmospheric conditions and viewing geometries. Additional secondary criteria were the minimisation of the fit RMS, the fit error and the variability of the results.

The investigations on the NO_2 cross-sections included in the DOAS fit (compare Sect. 6.2.2) showed that the additional usage of (first-order) Puḳīte terms is essential under high NO_2 conditions since large systematic underestimations of the true reference dSCDs were obtained at low elevations with the standard fit settings. As also for the low NO_2 scenario, all combinations of NO_2 cross-sections led to systematic deviations from the reference dSCDs, whereby the differences between the different fit scenarios were smaller than the deviations themselves, it was concluded that perfect agreement between retrievals and simulations is probably not achievable. Thus, large focus was put on consistent results for both NO_2 scenarios besides the minimisation of the deviations. Consequently, the decision was to add the NO_2 cross-section at 220 K as well as the linear and square Puḳīte terms to the HCHO fit settings. This set of NO_2 cross-sections led to significantly improved agreements between retrieved and reference HCHO dSCDs at low elevations for the high NO_2 cases and to highly consistent results for both NO_2 scenarios for a given sun geometry. The results obtained with the real measured spectra, for which no true dSCDs are known, showed that this choice leads to notable improvements of the fit RMS.

This set of NO_2 cross-sections was then used for the studies regarding the selection of the O_3 cross-sections. Here, it was shown that the usage of two cross-sections at 223 K and 243 K is essential since worse agreements and increased fit RMS-values were observed for fit scenarios not including both cross-sections. This was also confirmed by the results for the real measured spectra. Further, it was found that despite the observation that the additional O_3 cross-sections increase the variability of the results between the different sun geometries, many of the tested combinations including additional O_3 cross-sections lead to improvements of the absolute deviations at 10° and 30° elevation. Contrary to that, mainly a slight worsening of the agreements was observed at low elevations together leading to smoother elevation dependencies of the deviations from the reference dSCDs

and thus to smaller effects on the retrieved profile shapes. In addition, many sets of additional cross-sections for O_3 led to smaller fit RMS-values especially at high elevations. Altogether, it was concluded that the advantages of additional O_3 cross-sections outweigh the disadvantages which was also confirmed by the comparisons with the results obtained with the BIRA fit settings (compare Sect. 6.6). Hence, the combination which additionally includes the linear and square Puķīte terms of O_3 was selected since it yielded the overall smallest deviations from the reference HCHO dSCDs based on the results for the synthetic spectra.

Using the final selection of cross-sections, the choice of the fitting interval was examined by investigating the dependencies of the fit RMS, the fit error, the variability of the retrieved HCHO dSCDs and the deviations from the reference dSCDs on the used fitting intervals. Thereby, it was found that the first three quantities mainly depend on the length of the applied fitting interval with the RMS-values increasing with increasing lengths, while the opposite is true for the fit error and the variability. For more details, refer to Sect. 6.3. The interpretation of the deviations from the true reference HCHO dSCDs was more complex which showed more variations between the different sun geometries. Although the standard fitting interval from 324 to 359 nm is not the best choice in neither individual aspect, it yields acceptable overestimations of the reference dSCDs for most sequences of synthetic spectra and thus consistent results for the different sun geometries. Again making a trade-off between deviations from the reference values, consistency as well as performance in terms of fit RMS, fit error and the variability of the resulting dSCDs, it was decided to stay with the standard fitting interval. This choice is also justified based on the results for the real measured spectra which show a minimum in the variability of the results in this spectral region. Other choices are also possible if the focus is put on other aspects (see below). However, as consistency is the central aspect in this work, the considerations were made in this regard here. A last aspect here was that findings based on synthetic spectra cannot always directly be transferred to real measurements. For instance, the interpretation of the variability of the results was more complex for the real measured spectra probably caused (amongst other factors) by detector properties.

Altogether, these optimisations led to the optimised standard fit settings for formaldehyde summarised in Table 9.1 which additionally include the linear and square Puķīte terms for O_3 and NO_2 compared to the standard settings (compare Table 6.3). These fit settings are used to retrieve HCHO dSCDs from the MAX-DOAS measurements presented in this work (compare Chapt. 9 to 13). A major advantage of the usage of synthetic spectra for the here-presented sensitivity study is the possibility to estimate a lower threshold for the uncertainties of these dSCDs which are related to the choice of these specific fit settings. Here, it is important to note that measurement errors are an additional source of uncertainties for real atmospheric measurements.

Despite all optimisation efforts discussed in the paragraphs above, systematic deviations from the true reference dSCDs in terms of mostly overestimations exist for all scenarios (compare Sect. 6.4). Since the optimisations were performed in

a way that consistent results are obtained for all trace gas scenarios, only comparably small differences between these are found and the remaining deviations mainly vary between the different sun geometries. For the latter, the RMS_{abs}-values describing the absolute deviations from the reference dSCDs within a given sequence differ by almost 3×10^{15} molec cm^{-2} (peak-to-peak). Nevertheless, small (systematic) differences are also observed between the low and high NO_2 scenarios for which the mean RMS_{abs} differ by approximately 4.5×10^{14} molec cm^{-2} (larger for low NO_2 scenario), while the differences are negligible between both HCHO scenarios where the mean RMS_{abs} differ by only 6×10^{12} molec cm^{-2}. These remaining deviations are used to estimate the uncertainties (or at least their order of magnitude) of the HCHO dSCDs retrieved with the optimised fit settings.

The absolutely largest deviation from the reference HCHO dSCDs obtained with the optimised fit settings occurs in the case of high HCHO, low NO_2, an SZA of 60° and an RAA of 30° at 3° elevation where a value of 3.7×10^{15} molec cm^{-2} is found (compare Sect. 6.4). The lowest absolute deviation is 1.6×10^{12} molec cm^{-2} and occurs at 1° for the case of low HCHO, high NO_2, an SZA of 20° and an RAA of 90°. As discussed earlier, mostly systematic overestimations of the reference dSCDs are obtained. Further, the absolute deviations show a systematic elevation dependence within a given sequence, whereby for many sequences the largest deviations occur at 10°. However, it should be stressed that this elevation dependence was already significantly reduced by the inclusion of the linear and square Puḳīte terms for O_3 (compare Sect. 6.2.3). Slight underestimations of the reference dSCDs are only found for a few sequences with an SZA of 20°.

To provide a reasonable estimate of the typical uncertainty of the HCHO dSCDs retrieved with the here-suggested optimised fit settings, the mean over all RMS_{abs}-values of the different sequences is calculated and a value of 1.75×10^{15} molec cm^{-2} is obtained. In case of high HCHO abundance, this corresponds to uncertainties in the order of 5 % and 25 % at low and high elevations, respectively, while they reach up to 12 % and 70 % in case of low HCHO abundance. It has to be stressed again that the remaining deviations vary notably between the different sun geometries. Nevertheless, based on the here-presented study and the investigated scenarios, the upper limit of the uncertainties related to the choice of the suggested fit settings is estimated to be an overestimation of the true dSCDs by 3.7×10^{15} molec cm^{-2} at low elevations, while the upper limit is around 2.5×10^{15} molec cm^{-2} at 30°. Further, it should be noted that this estimate only describes the expected uncertainties of the retrieved HCHO dSCDs related to the choice of the fit settings and to the fitting procedure itself but does for instance not include uncertainties related to inaccuracies of the absorption cross-sections since the analyses were performed using the cross-sections which were also used for the simulations of the synthetic spectra. A last aspect is that these estimated uncertainties are based on synthetic data and that the results might not be directly transferable to real data. As shown above, slight differences in the findings obtained with the synthetic spectra and the real measured spectra occurred for some investigations. However, no fundamentally different behaviours were identified.

The investigations presented in Sect. 6.4 showed that none of the considered effects alone can explain the remaining deviations between retrieved and reference HCHO dSCDs. Further, some effects could even be excluded, e.g. the accuracy and precision of the simulations. Therefore, it was concluded that limitations of the DOAS method itself, i.e. the fact that not all absorptions which occur can be accounted for and the wrong order of convolution and calculation of the logarithmic ratio of the spectrum to be analysed and the FRS, are the main contributors to these remaining deviations and thus to the uncertainties of the retrieved HCHO dSCDs. Further, remaining inaccuracies introduced by the Ring effect and spectral noise were identified to slightly contribute here. In addition, the studies in Sect. 6.5 suggested that also the choice of the polynomial degree and the decision whether the BrO cross-section is included in the fit or not contribute to some extend and other choices can lead to slight improvements or a worsening (both at maximum ca. 1×10^{15} molec cm^{-2}) depending on the specific sun geometry. However, based on these rather brief investigations, the corresponding choices used in the optimised fit settings are found to be reasonable leading to rather stable results. These aspects are already included in the uncertainties estimated above.

Besides systematic uncertainties, also the variability of the retrieved dSCDs was examined for both synthetic and real measured spectra (compare Sect. 6.3.2). In particular, the standard deviation over the dSCDs retrieved for the 100 different spectra containing noise (or in the considered 4-hour time period for the real data) was used to estimate the statistical fluctuations of the retrieved dSCDs. For the synthetic spectra, the standard deviations are in the order of 2.5×10^{15} molec cm^{-2} for the finally chosen fit settings and vary only slightly between the different viewing geometries (combinations of elevation angle and sun geometry) and trace gas scenarios. Thereby, the peak-to-peak difference is ca. 5×10^{14} molec cm^{-2}. As described in Sect. 6.3.2, some of the findings obtained for the synthetic spectra cannot directly be transferred to real data. In particular, the findings for the dependence of the fit RMS on the fitting interval suggested that the noise level of the synthetic spectra was chosen slightly too high. This is also reflected by the finding that the standard deviations are notably smaller for the real measured spectra. Here, values of around 1.2×10^{15} molec cm^{-2} are found, varying only by less than 3×10^{14} molec cm^{-2} between the different elevations. One aspect contributing to the observed difference between the findings for the synthetic and real data might be the fact that 7 times less measurements were included for the real data. Together, the variability related to the choice of the fit settings can be estimated to be notably below 2.5×10^{15} molec cm^{-2} corresponding to relative fluctuations of below 6 % and 30 % at low and high elevations, respectively, for high HCHO conditions and 15 % and 90 % at low HCHO abundances. This statistical component is expected to be different for other instruments as it depends on the noise level of the used spectrometer unit.

In conclusion, the results of this sensitivity study suggest for most measurement conditions a systematic overestimation of the true HCHO dSCDs in the order of 1.75×10^{15} molec cm^{-2} and a statistical uncertainty below 2.5×10^{15} molec cm^{-2}. Note that this does not include other uncertainties, e.g. related to the used absorp-

tion cross-sections or shortcomings of the measurements or the data processing. Although not all remaining deviations and uncertainties could be clarified in this work, the fact that the optimised fit settings lead to significantly improved HCHO results compared to the standard settings should be stressed again. In particular, the agreement between retrievals and simulations under high NO_2 conditions was clearly improved. Further, the optimised fit settings yield highly consistent results regardless of the NO_2, HCHO and aerosol abundances, whereby also the elevation dependence of the absolute deviations from the reference values within a given sequence was reduced and thus the impact on the finally retrieved profile shape minimised.

Although consistent results together with small deviations from reference values (if available) like in this work should always be preferred, sometimes specific requirements on the results or specific measurement circumstances might lead to different decisions in the here-presented optimisation steps. For instance, if the focus for the selection of the NO_2 cross-sections included in the DOAS fit was purely put on the deviations from the true reference HCHO dSCDs, the combination additionally including the cross-section at 220 K, the linear Puķīte term and the Raman cross-section would be the best choice in case of high NO_2, while the combination additionally including the cross-section at 220 K, the square Puķīte term and the Raman cross-section would be the best choice if both NO_2 scenarios were considered together. It should be stressed that these two combinations perform relatively bad in terms of consistency between both NO_2 scenarios (compare Fig. 6.9) and that compensations of strong over- and underestimations for the individual cross-sections were observed. In case consistency should still be an important factor but low deviations the major criterion, the combination additionally including the cross-section at 220 K and only the linear Puķīte term would be a possible choice.

Further, the combination of O_3 cross-sections included in the standard settings would be the optimal choice if the variation between the different sun geometries was the only criterion in this optimisation step and the improvement in the agreement between retrievals and simulations which is achieved was ignored (compare Fig. 6.17).

Also other fitting intervals might be used depending on how much weight is put on the different aspects. For instance, in case of a spectral range covered by the instrument which is limited to wavelengths above ca. 330 nm or if a fixed FRS has to be used for the DOAS analysis, the commonly used 336 to 359 nm interval is a reasonable choice (compare considerations in Sect. 6.3). As discussed already above, a fitting interval from 319 to 360 nm would also be possible if a fit RMS which is as low as possible was not an important aspect. Lastly, a rather large range of fitting intervals with lower boundaries in the range of approximately 315 to 323 nm and upper boundaries between roughly 340 to 357 nm could be good choices if the focus was again solely put on low deviations from the reference dSCDs. As outlined in Sect. 6.3.2, such intervals are problematic for the real measured data which, however, might be different for other instruments. Therefore, sensitivity studies performed for specific instruments might be needed here. For more detailed discussions regarding the fitting interval, refer to Sect. 6.3.

Also specific atmospheric conditions might justify specific fit settings if only one measurement site with constant atmospheric conditions is considered. Nearly all findings in the here-presented sensitivity study were almost independent of the HCHO scenario. Hence, no specific recommendation in this regard can be given here and all optimisations apply to high and low HCHO conditions. The same is true for measurement sites regularly experiencing high aerosol abundances. As shown in Sect. 6.4.4, for the high AOD scenario notably better agreements between retrievals and simulations are found using the optimised fit settings compared to the low AOD case. Hence, these can be recommended regardless of the aerosol abundances.

For the usage of the Puķīte terms for NO_2, however, a few considerations can be made. In principle, these terms are not needed under low NO_2 conditions since no improvements of the HCHO results are achieved under such atmospheric conditions when these terms are included in the fit based on the results obtained for the synthetic spectra (compare Sect. 6.2.2.1). In fact, the finally chosen combination of NO_2 cross-sections leads to a slightly increased mean RMS_{abs} for the low NO_2 scenarios which is, however, not significant if the fluctuations of the results between the different sun geometries are considered. Therefore, if only the results of one measurement site at which very low NO_2 abundances prevail at all times are of interest, the additional NO_2 cross-sections can be excluded from the DOAS retrieval. In this work, however, the HCHO results of several stations are investigated and compared. Since most of them are regularly dominated by rather high NO_2 pollution and consistency is a major aim of this work, Puķīte terms for NO_2 are used for all stations. Based on the findings obtained in Sect. 6.2.2.2 using measurements from the ATTO site where low NO_2 conditions prevail throughout the year (compare Sect. 4.4 and Chapt. 13), this is not critical.

Since the results for the Puķīte terms of O_3 depend only slightly on the NO_2 scenario, no specific recommendation can be given in this regard. However, the usage of Puķīte terms for O_3 might be reassessed in case no NO_2 Puķīte terms are used.

Lastly, there was some indication that long fitting intervals (above ca. 40 nm) with upper boundaries at long wavelengths (above ca. 359 nm) tend to yield larger fit RMS-values. This effect was strongest at low elevations and is thus probably related to the inclusion of more and stronger NO_2 absorption bands. Also the elevation dependence of the absolute deviations within a given sequence was marginally larger for these intervals. No final conclusion on these potential effects of NO_2 absorptions on the results using such fitting intervals was provided in this work. If the usage of such intervals is desired under high NO_2 conditions, further studies are suggested here.

Finally, the HCHO results obtained with the recommended fit settings based on the here-presented sensitivity study were compared with the ones obtained using fit settings from IUP-Bremen and BIRA. Here, large parts of the observed differences were attributed to the different approaches of accounting for the Ring effect. These comparisons also stressed that the inclusion of the linear and square Puķīte terms for O_3 and NO_2 as recommended in this work is well justified.

Moreover, it became evident that the optimised fit settings yield highly consistent results for many atmospheric conditions.

7

IMPROVEMENT OF THE DOAS FIT SETTINGS OF FURTHER SPECIES

Besides the comprehensive investigations on the DOAS fit settings of formaldehyde, which were presented in the previous chapter, additional but less comprehensive studies on the fit settings of O_4 (Sect. 7.1) and glyoxal (Sect. 7.2) were performed. The main findings of these studies are summarised in this chapter.

7.1 WATER VAPOUR ABSORPTIONS IN THE O_4 FITTING WINDOW

The residuals of the first O_4 fits for the ATTO 80 m instrument showed strong systematic structures at around 363 nm. An example of such a residual is depicted in panel (a) of Fig. 7.1 for a spectrum recorded by the ATTO 80 m instrument at 3° elevation on 31 May 2020. Lampel et al. (2017a) reported on the evidence of water vapour absorptions around 363 nm in spectra of MAX-DOAS measurements and showed that the inclusion of the *POKAZATEL* line list (Polyansky et al., 2018) in the DOAS fits of UV absorbers can improve the results significantly. Further, an increase of the retrieved O_4 dSCDs was observed in that study when the *POKAZATEL* H_2O cross-section was included. Therefore, also in this work the usage of an H_2O absorption cross-section in the O_4 fit was tested in order to minimise the observed residual structure around 363 nm. Besides the *POKAZATEL* line list which is called "Polyansky (2018)" in the following, also the more recent HITRAN (2020) water vapour absorption cross-section (Gordon et al., 2022) was tested.

Panels (b) and (c) of Fig. 7.1 show the fit residual and the fitted H_2O cross-section, respectively, for the case when the Polyansky (2018) cross-section is included in the DOAS fit of O_4 for the above mentioned spectrum. Panels (d) and (e) show the same results for the HITRAN (2020) cross-section. Compared to panel (a), in both cases the pronounced structure around 363 nm is removed by the inclusion of a water vapour cross-section. Furthermore, panels (c) and (e) suggest that the residual structure can be solely attributed to water vapour absorptions, as its structure is well represented by the fit. The water vapour dSCD retrieved in the visible part of the spectrum is 5.57×10^{23} molec cm^{-2}, while values of 1.23×10^{24} and 3.78×10^{23} molec cm^{-2} are obtained in the UV part for Polyansky (2018) and HITRAN (2020), respectively. In the UV, these dSCDs lead to optical depths of slightly above 1×10^{-3} and the values found in this work are consistent with the numbers given in Lampel et al. (2017a) for similar water vapour dSCDs retrieved in the visible spectral range. The maximum optical depths reported by Lampel et al. (2017a) reach up to 2×10^{-3} for UV H_2O absorptions which can be exceeded for spectra measured at the tropical station ATTO. There the water vapour dSCDs retrieved in the visible spectral range can be in the order

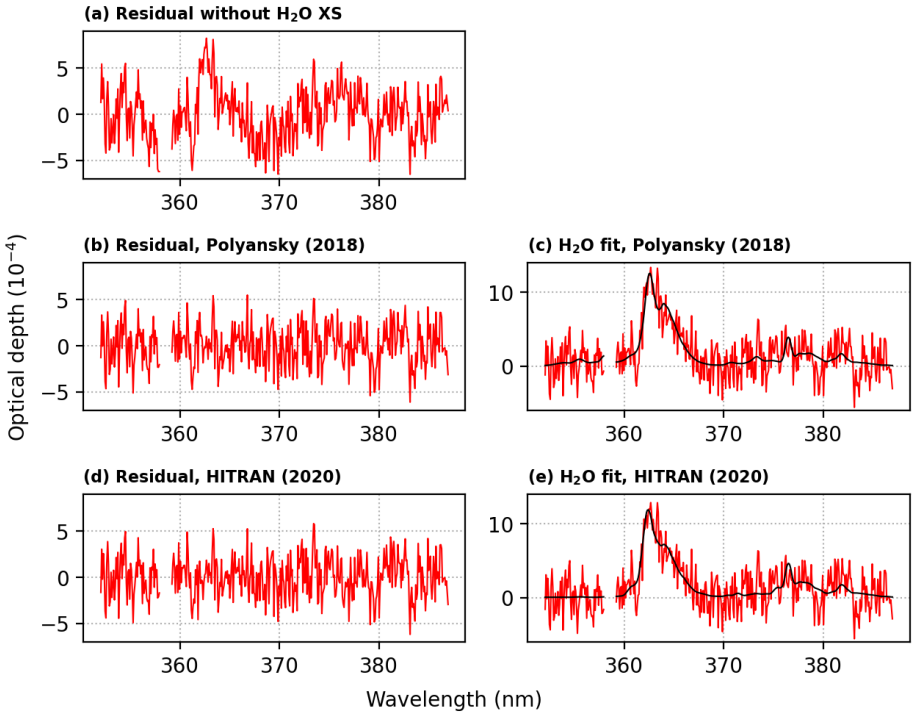


Figure 7.1: Residuals (left column) and fitted H_2O cross-section (right column) of an O_4 fit for a spectrum recorded by the ATTO 80 m instrument at 3° elevation on 31 May 2020. Panel (a) shows the residual for the case when no H_2O cross-section is included in the fit. Panels (b) and (c) show the residual and the fitted H_2O cross-section when the Polyansky (2018) cross-section is included, respectively. Lastly, panels (d) and (e) show the same for the case when the HITRAN (2020) cross-section is included in the fit. In panels (c) and (e) black lines indicate the fitted cross-section, while red lines show the fitted cross-section plus the residual.

of 1.5×10^{24} molec cm^{-2} and the corresponding opticals depths for the UV H_2O absorptions are about 2.3×10^{-3} for the Polyansky (2018) cross-section. Finally, also the reported increase in the O_4 dSCDs for moderate to high water vapour abundances of roughly 5 % can be confirmed and values of 4.6 % and 3.1 % are found for the two tested cross-sections, respectively, for the example spectrum.

Despite the clear and consistent findings, some questions remain. First, it has to be investigated how the inclusion of a water vapour absorption cross-section influences the absolute values of the retrieved O_4 dSCDs for a longer time series. Secondly, it is important to check whether the scatter of the retrieved O_4 dSCDs is enhanced or not. Next, the findings presented so far were only based on measurements at ATTO which experiences high absolute humidities throughout the entire year, but it has to be checked how this compares to data recorded at a more moderate location with lower water vapour abundances. Lastly, it has to be ensured that the observed fitted water vapour dSCDs in the UV represent real water vapour abundances in the atmosphere.

To answer these questions, data collected with the ATTO 80 m and Bayfordbury instruments were used and for both instruments data for one complete year

were analysed using three different scenarios. These scenarios are based on the standard O_4 fit settings summarised in Table 9.4, but fits were performed either without any H_2O cross-section or including the Polyansky (2018) or HITRAN (2020) cross-section. Note that for these fits I_0-corrections were applied for the NO_2 and O_3 cross-sections instead of intensity weighted convolutions like done for the final analyses. In the following, the effects on the retrieved O_4 dSCDs, the fit RMS and the scatter of the retrieved O_4 dSCDs are investigated.

Figure 7.2 shows time series of O_4 dSCDs (a), differences (b) and ratios (c) of the retrieved O_4 dSCDs and the fit RMS (d) for the different fit scenarios as described above for measurements performed at 3° elevation by the ATTO 80 m instrument in the year 2020.

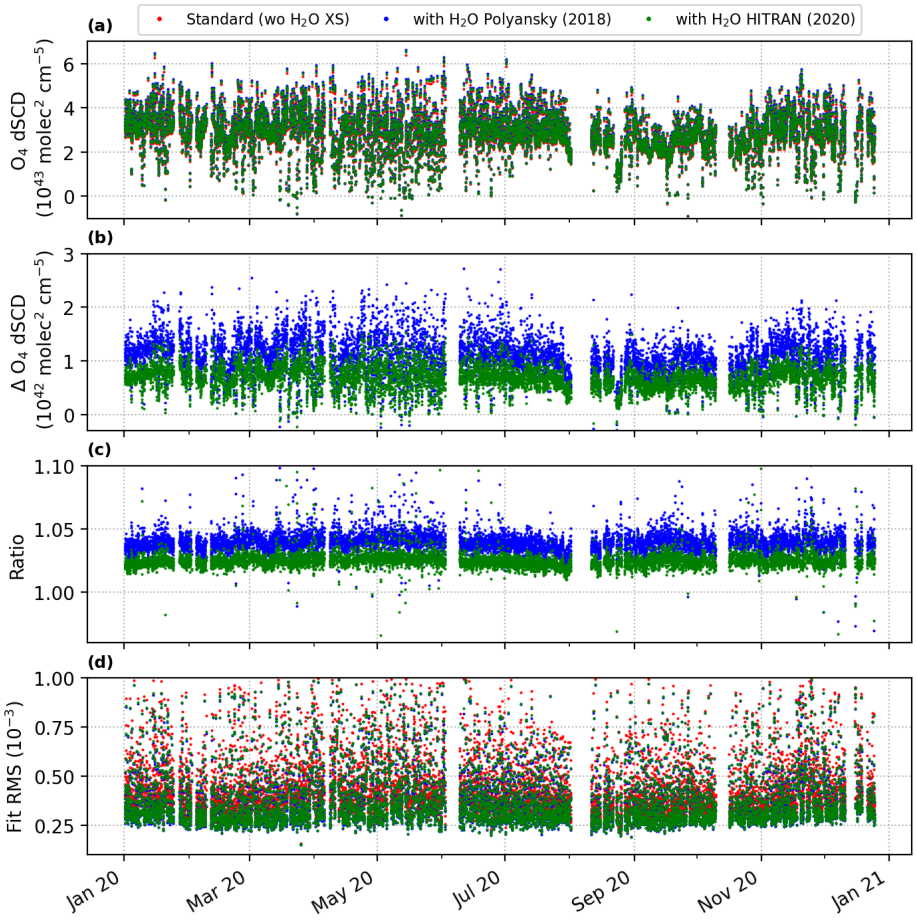


Figure 7.2: Time series of O_4 dSCDs (a), differences (b) and ratios (c) of the retrieved O_4 dSCDs and the fit RMS (d) for the different fit scenarios as described in the text for the ATTO 80 m instrument at 3° elevation.

The differences and ratios between the O_4 dSCDs retrieved with the two different water vapour absorption cross-sections included in the fit and the ones of the standard scenario without such an cross-section indicate an increase of the dSCDs when H_2O cross-sections are included. The absolute increases of the O_4 dSCDs at 3° are on average 1.09×10^{42} and 6.77×10^{41} molec cm^{-2} for the Polyansky (2018) and the HITRAN (2020) cross-section, respectively,

corresponding to mean relative effects of 4.2 and 2.5 %. The values found for the Polyansky (2018) cross-section are in good agreement with the above example and the studies by Lampel et al. (2017a). A closer look shows that the observed effects are weaker during the dry season (August to November). Also the fit RMS is clearly reduced when water vapour absorptions are accounted for. Here, the effects of both tested cross-sections are comparable.

The described effects on the O_4 dSCDs and the fit RMS become smaller with increasing elevation (not shown) which is an indication that the observed effects are related to real water vapour absorptions which become smaller at higher elevation angles.

To estimate the effect on the scatter of the retrieved dSCDs, the standard deviations of the three time series were calculated. Assuming that systematic variations of the dSCDs are captured identically by all three fit scenarios, changes in the standard deviations can be solely attributed to changes in the statistical fluctuations. For the time series at 3° elevation depicted in Fig. 7.2, values of 9.46×10^{42}, 9.79×10^{42} and 9.66×10^{42} molec cm^{-2} are obtained for the standard, the Polyansky (2018) and the HITRAN (2020) scenarios, respectively. The changes in the standard deviations are small, nevertheless, a slight increase of the scatter can be observed when H_2O absorption cross-sections are included in the fit.

Given the findings for ATTO, the usage of an H_2O absorption cross-section seems to be reasonable under humid conditions. To check whether this also applies to drier stations with lower absolute humidities, data collected at Bayfordbury were analysed for the year 2019 using the same three scenarios as described above. The resulting time series of the O_4 dSCDs, differences, ratios and fit RMS are depicted in the same way as for ATTO in Fig. 7.3.

At 3° elevation, slightly enhanced O_4 dSCDs are observed and the absolute mean differences are 6.55×10^{40} and 1.63×10^{40} molec cm^{-2} for Polyansky (2018) and HITRAN (2020), respectively. For both tested cross-sections this translates to relative changes of practically zero. Nevertheless, slightly larger values are visible in the summer months indicating a small effect for higher absolute humidities and thus stronger water vapour absorptions. The fit RMS is essentially not affected and only marginal improvements are observed.

Like for ATTO, the observed effects become smaller with increasing elevation (not shown) and are quasi zero for the highest elevations which again indicates a clear relation to real H_2O absorptions.

For Bayfordbury, standard deviations of the O_4 dSCDs of 13.77×10^{42}, 13.81×10^{42} and 13.78×10^{42} molec cm^{-2} are calculated for the standard, the Polyansky (2018) and the HITRAN (2020) scenarios, respectively. As already seen for the ATTO 80 m instrument, this indicates a slight increase of the scatter of the retrieved O_4 dSCDs.

The results obtained here together with the findings of Lampel et al. (2017a) suggest that it is reasonable to account for water vapour absorptions in the DOAS fits of O_4 in the UV spectral range. Actually, this also applies to other UV absorbers, e.g. HONO (compare Table 9.5), which are fitted in wavelength ranges that include the here investigated 363 nm water vapour absorption band. While Lampel et al. (2017a) recommended the Polyansky (2018) cross-section, here also

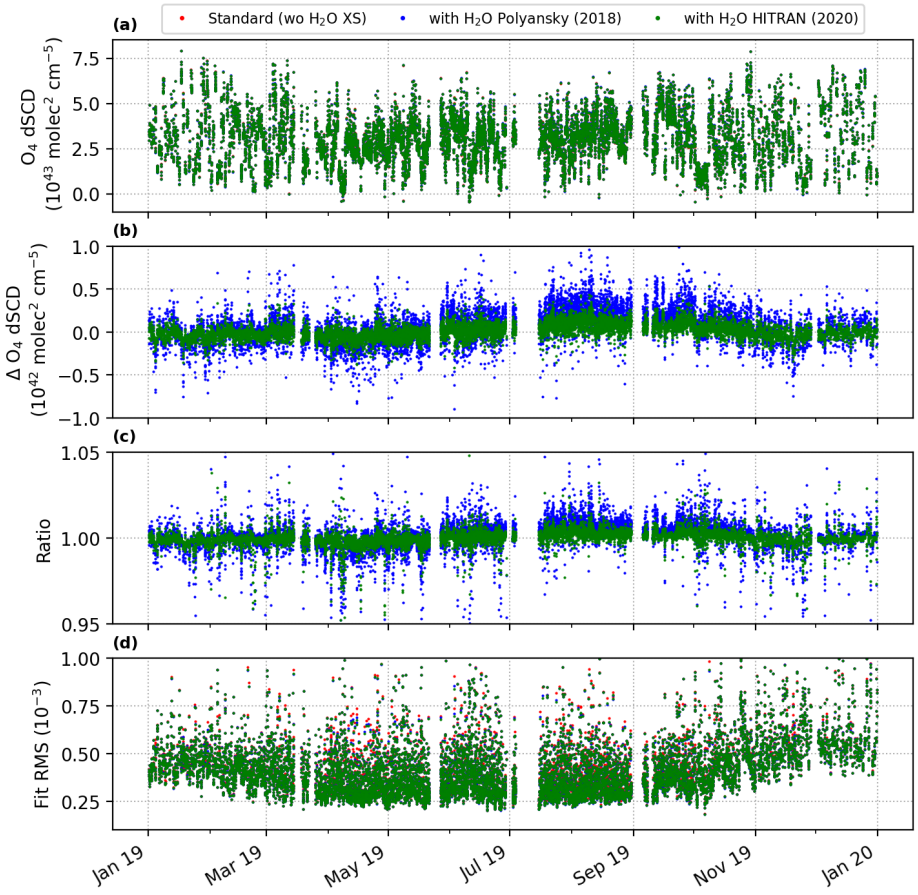


Figure 7.3: Time series of O_4 dSCDs (a), differences (b) and ratios (c) of the retrieved O_4 dSCDs and the fit RMS (d) for the different fit scenarios as described in the text for the Bayfordbury instrument at 3° elevation.

the more recent HITRAN (2020) cross-section was tested. The findings in panels (b) and (c) of Fig. 7.2 and 7.3 propose that the results obtained for the HITRAN (2020) scenario are more stable and show less scatter. Also the calculated standard deviations of the O_4 dSCDs are lower than the ones using the Polyansky (2018) cross-section. To have a closer look on which of the two cross-sections is better suited and to check once more if the observed effects can be solely attributed to real H_2O absorptions, the H_2O dSCDs retrieved in the UV are compared to the ones retrieved in the visible spectral range.

For that, spectra from the ATTO 80 m instrument for the year 2020 were analysed using the two O_4 fit scenarios from above which include the two tested water vapour cross-sections and one additional scenario that fits water vapour in the visible spectral range (400 to 460 nm). For the latter, the HITRAN (2012, Rothman et al., 2013) cross-section rescaled according to Lampel et al. (2015) was used. The obtained water vapour dSCDs are then correlated amongst each other. The corresponding correlation plots shown in Fig. 7.4 include data from all elevation angles. However, only results for SZAs below 80°, fit RMS-values below 1×10^{-3} and elevations starting at 2° are considered. Results of simple

linear regressions are shown in red, while 1-to-1 lines are shown as black dashed lines.

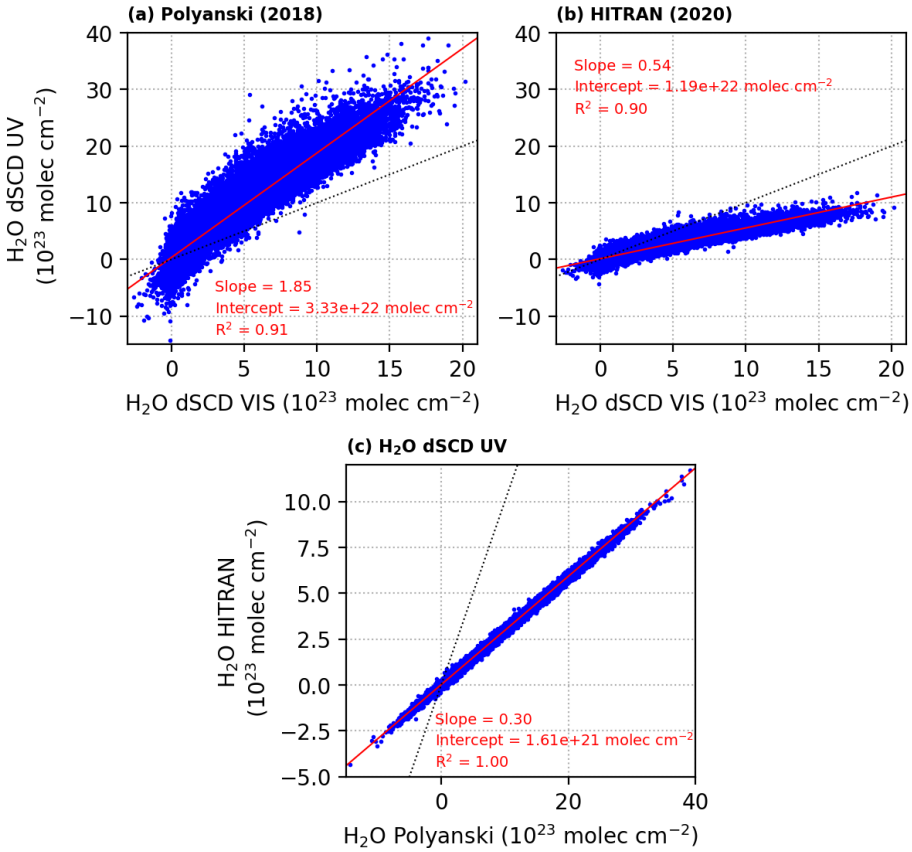


Figure 7.4: Water vapour dSCDs retrieved with the Polyansky (2018) (a) and HITRAN (2020) (b) cross-sections in the UV plotted against the corresponding dSCDs retrieved in the visible spectral range for the ATTO 80 m instrument in the year 2020. Panel (c) correlates the H_2O dSCDs in the UV using the Polyansky (2018) and HITRAN (2020) cross-sections. Results of simple linear regressions are shown in red, while 1-to-1 lines are shown as black dashed lines.

For both tested water vapour cross-sections, good correlations between the dSCDs retrieved in the UV and visible spectral ranges are found. The R^2-values are 0.91 and 0.90 for Polyansky (2018) and HITRAN (2020), respectively. The slopes of the linear regressions are 1.85 and 0.54 for the two cross-sections and differ by a factor of ca. 3.5 which describes differences between the absorption strengths of both cross-sections. The correlation between the two UV retrievals (c) indicates a high consistency and the slope basically reflects the ratio between the absorption strengths of both tested cross-sections. These analyses prove that the fitted H_2O cross-sections in the UV can be attributed to physical H_2O absorptions.

Comparing panels (a) and (b) suggests that the results obtained with the Polyansky (2018) cross-section show a slightly larger scatter which was also seen in the previous investigations on the O_4 dSCDs. Further, the H_2O dSCDs retrieved with HITRAN (2020) show physically more reasonable results since the obtained dSCDs are smaller than the H_2O dSCDs fitted in the visible spectral

range. This is to be expected since the AMFs are typically smaller in the UV (compare Sect. 6.4.4). Although the differences between the AMFs in the visible and UV probably do not explain the complete differences in the retrieved results, the HITRAN (2020) cross-section seems to be better suited than the Polyansky (2018) cross-section.

Altogether, the benefits of including a water vapour absorption cross-section in the fits of species which are retrieved in wavelength ranges that include the 363 nm water vapour absorption band exceed the slightly enhanced scatter of the obtained dSCDs of the target species. This is also the case for stations experiencing only moderate water vapour abundances like Bayfordbury. Thus, in the here-presented studies the HITRAN (2020) absorption cross-section is included in the O_4 and HONO fits, as indicated in Tables 9.4 and 9.5 for all stations. In the visible spectral range (e.g. NO_2 fits), still the HITRAN (2012) cross-section rescaled according to Lampel et al. (2015) is used.

7.2 IMPROVEMENTS OF THE GLYOXAL FIT

Several studies were conducted to improve the DOAS fit results for glyoxal. Brief summaries of these studies are presented in the following sections.

7.2.1 *Treatment of water vapour residual structures*

At typical spectral resolutions of MAX-DOAS instruments, glyoxal shows its strongest absorption features around 455 nm and is therefore typically retrieved in fitting ranges between 400 and 460 nm. Additionally, NO_2 is typically retrieved in this spectral range. Water vapour, however, also shows strong absorption structures in this spectral range which are not well characterised by the existing spectroscopic data for H_2O (Borger et al., 2020). Lampel et al. (2015) showed that this can lead to residual structures in DOAS retrievals which have a significant impact on retrievals of weak absorbers such as glyoxal. A possible effect here are spectral interferences that cause artificial dSCDs that are not related to real physical absorptions of trace gases. Also, the residuals of the first glyoxal and NO_2 fits for the ATTO 80 m instrument showed strongly pronounced structures which could be clearly attributed to strong water vapour absorptions (compare Fig. B.4 in the appendix). As the abundances of glyoxal and NO_2 are expected to be close to their detection limits at ATTO, the fits have to be significantly improved. To achieve this, several efforts were made, which are briefly listed here:

1. Usage of the HITRAN (2012) H_2O cross-section for which an additional scaling of the absorption strengths of the different absorption bands relative to each other was performed as recommended in Lampel et al. (2015).

2. Shift of the H_2O cross-section by $+0.021$ nm as for high water vapour dSCDs a shift of the cross-section that was allowed on a test basis converges towards this value. Further, the main absorption peaks in this spectral range at 416 and 442 nm were fitted separately (Wagner, 2019).

3. Usage of a symmetric Super Gaussian slit function (SLF) instead of a simple Gaussian SLF for the calibration and the convolutions of the cross-sections as described in Sect. 8.1, which in theory should represent the instrument's SLF better (Beirle et al., 2017).

As a consequence of these modifications in the treatment of the water vapour absorption cross-section, the retrieved glyoxal and NO_2 dSCDs got smaller and a slight improvement of the observed residual structures was achieved. An example for the fit results before the described modifications were applied is presented in Fig. B.4 in the appendix. The largest improvement was gained by using a Super Gaussian SLF instead of a simple Gaussian SLF. Nevertheless, despite all efforts, strong residual structures, especially between 440 and 450 nm, can still be seen as presented in Fig. 7.5. The structure around 416 nm could be clearly reduced and is nearly not visible any more indicating that it is caused by strong H_2O absorptions. However, the high frequency structures between 440 and 450 nm are still very prominent (compare Fig. 7.5a). The RMS for this spectrum could be improved from 5.50×10^{-4} to 4.42×10^{-4}. As can be seen in panels (b) and (c), the remaining structures probably have a systematic influence on the fits of NO_2 and glyoxal. The water vapour cross-section shown in panel (d) illustrates that these residual structures clearly overlap with its absorption structures.

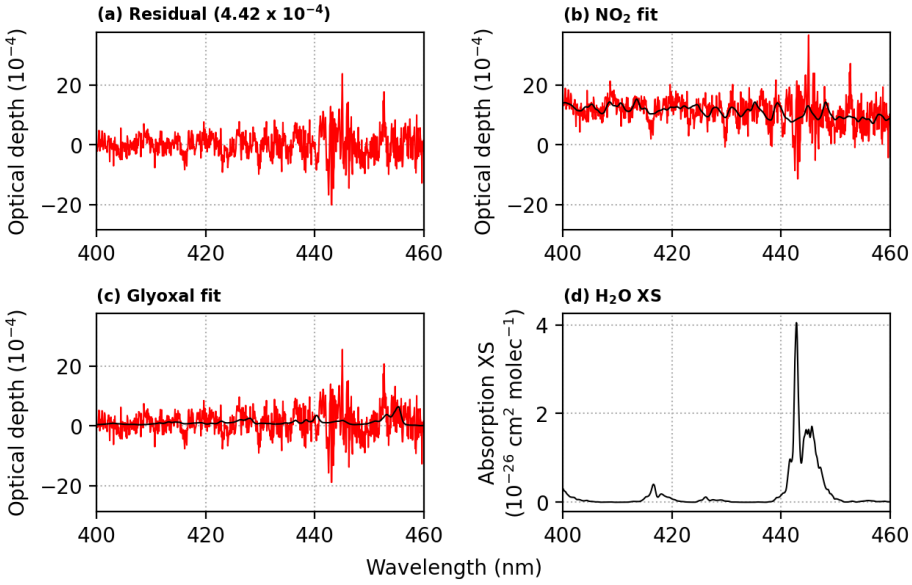


Figure 7.5: Panel (a) shows the residual of a spectral fit performed in the range from 400 to 460 nm for a spectrum recorded by the ATTO 80 m instrument at 2° elevation on 22 December 2017 at 15:30 LT. Panels (b) and (c) present the corresponding fits of NO_2 and glyoxal, respectively. Panel (d) shows the water vapour absorption cross-section.

In addition to the improvements listed above, also the use of Puḳīte terms (Sect. 3.1.5.2) for H_2O and the influence of an intensity weighted convolution of the H_2O cross-section were tested. Furthermore, it was also examined whether saturation effects due to very strong water vapour absorptions can explain the residual structures. For the latter, synthetic cross-sections were generated for high and

low water vapour absorptions using SCDs of 15×10^{23} and 1×10^{22} molec cm^{-2}, respectively. These synthetic cross-sections were then used in the DOAS fit instead of the normal water vapour cross-sections. However, neither of these efforts led to any improvements of the residual.

The glyoxal dSCD in the presented example is 1.2×10^{15} molec cm^{-2} and is thus probably close or below the detection limit which is rather high in view of the observed structures. For spectra including significant glyoxal absorptions (larger than ca. 2×10^{15} molec cm^{-2}), the fits look overall reasonable, however, further improvements would lower the detection limit. Moreover, the observed structures are even worse for the ATTO 298 m instrument as shown in Fig. 7.6, where the water vapour dSCD is comparable to the spectrum presented for the 80 m instrument. In addition, the residuals of the fits for the ATTO 298 m instrument show a strong structure around 430 nm which will be further studied in Sect. 7.2.4.

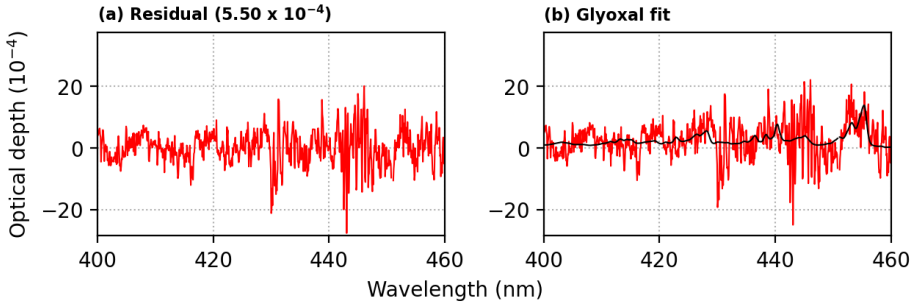


Figure 7.6: Panel (a) shows the residual of a spectral fit performed in the range from 400 to 460 nm for a spectrum recorded by the ATTO 298 m instrument at 2° elevation on 22 July 2019 at 12:49 LT. Panel (b) presents the corresponding glyoxal fit.

The largest improvement regarding the high frequency structures was achieved by a change from a Gaussian to a Super Gaussian SLF. Further, the observed structures have a FWHM of about 0.2 to 0.3 nm which is around half of the FWHM of the instrument's SLF. These two findings suggest that the observed high frequency structures are related to an insufficient characterisation of the SLF used for the convolution process. Unfortunately, no measured slit functions using spectral line lamps are available for the instruments used in this work. In order to test the previous hypothesis nevertheless, the HITRAN (2012, neither shifted nor split) water vapour cross-section was convolved with a normal Gaussian and a Super Gaussian, respectively, and the spectral difference was calculated[1]. Figure 7.7a shows the obtained difference as function of wavelength where strong high frequency structures are visible around the absorption bands at 416 and 442 nm.

These structures are especially prominent around 440 to 450 nm and roughly one order of magnitude smaller than the absorption cross-section itself in this spectral range. Panel (b) shows a zoom into the region around the observed high frequency structures at 440 to 450 nm and compares them to the fit residual of

1 The same was done for the solar spectrum, but no relevant structures were found.

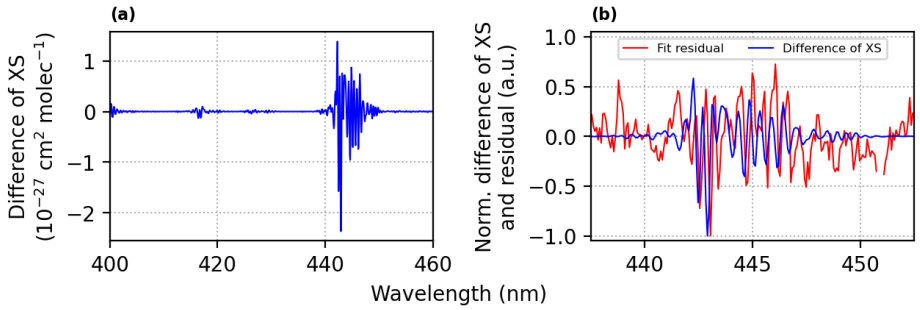


Figure 7.7: Panel (a) shows the difference of the HITRAN (2012) H_2O cross-section convolved with Gaussian and Super Gaussian SLFs, respectively, in the range from 400 to 460 nm. Panel (b) shows the same but zoomed into the region of the high frequency structures. Additionally, the fit residual from above is depicted in red as comparison. In panel (b), both curves are normalised to their corresponding maximum amplitudes.

the example spectrum from Fig. 7.6. The observed structures in the fit residual are similar to the structures which occur in the difference between the cross-sections convolved with the two different SLFs, while no perfect agreement of the structures is observed. Nevertheless, these observations show that an insufficient characterisation of the instrumental SLF and deviations from its correct shape can indeed lead to high frequency structures caused by strong water vapour absorptions. In theory, the difference between H_2O cross-sections convolved with the here-used Super Gaussian and the correct SLF should reproduce the residual structures perfectly. However, as pointed out before, the latter is not available for the used instrument. The same test as outlined here was conducted using a measured SLF from a different instrument, but no further improvements of the fit residual nor additional insights were achieved.

Although no measured SLF is available for the instrument used within this study, an attempt was made to determine spectral structures that arise from deviations of the assumed SLF from the shape of the correct SLF. These structures were then included as pseudo absorbers (PSAs) in the DOAS fit. For that, the already mentioned measured SLF from a different instrument was parametrised using a sum of a Gaussian and a Super Gaussian of the form

$$S = A_G \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) + A_{SG} \cdot \exp\left(-\left|\frac{x}{w}\right|^k\right). \tag{7.1}$$

A_G and A_{SG} are the amplitudes of the Gaussian and Super Gaussian terms, respectively. The standard deviation of the Gaussian term is described by σ, while w and k determine the width and shape of the parametrised slit function S. This function was then fitted to the measured SLF[2]. To make this fit more stable, initial values of all parameters were derived. For that, first a Gaussian was fitted to the tails of the measured SLF (distance from centre $\geq \pm 0.65$ nm). Afterwards, a Super Gaussian was fitted to the curve remaining after subtracting the previously fitted Gaussian from the measured SLF. Using these parameters as initial values,

2 Note that the measured SLF was upsampled to 0.01 nm resolution and normalised before the fit.

the function from Eq. 7.1 was finally fitted to the measured SLF. Lastly, the so yielded S describing the measured SLF was normalised such that the sum over the whole curve equals one which is necessary to use the function as convolution kernel. Note that the sum over each of the individual terms in Eq. 7.1 represents its share of the total function. Figure 7.8 presents the measured SLF (blue) as well as the fitted (green) and normalised (red) parametrisation.

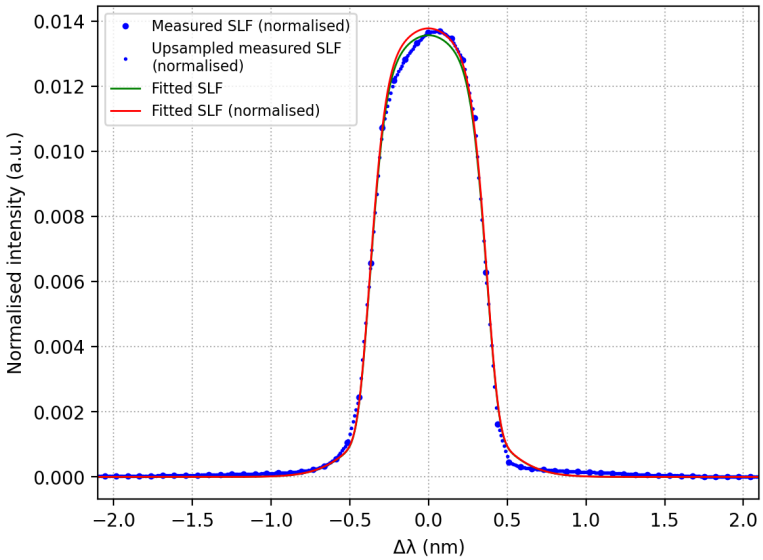


Figure 7.8: Measured slit function (large blue dots) and corresponding upsampled representation (small blue dots). The fitted curve using the parametrisation from Eq. 7.1 is shown in green, while the corresponding normalised fit curve is presented in red.

Despite a generally good fit result, the top and the tails are not perfectly captured by the fitted parametrisation. Further, the slight asymmetry is obviously not represented by purely symmetric functions. In principle, asymmetric Gaussian and Super Gaussian functions could have been used. However, tests showed that also these cannot well reproduce the slight asymmetry.

To investigate which spectral structures can be caused by deviations from the shape of the physically correct SLF, partial derivatives with respect to all five parameters of Eq. 7.1, namely A_G, σ, A_{SG}, w and k, were calculated numerically using

$$\partial_p S = \frac{S_{p^*} - S_0}{\Delta p}, \tag{7.2}$$

where S_0 is the fitted parametrised SLF and S_{p^*} describes the perturbed parametrisation. Further, p^* are the perturbed parameters according to $p^* = p_0 + \Delta p$ with p_0 and Δp being the unperturbed parameter and the perturbation, respectively. Thereby, a change rate of 5 % was applied to each parameter p. It is important to note that the normalisation of the sum over S_{p^*} to one has to be maintained. For perturbations in σ, w and k, this was achieved by a scaling of the amplitude of the respective term, while for A_G and A_{SG} a scaling of the amplitude of the other term had to be performed. The latter basically describes a change in the

weighting of both terms. Thus, only one of both amplitude perturbations has to be performed. The resulting partial derivatives are depicted in the four panels of Fig. 7.9.

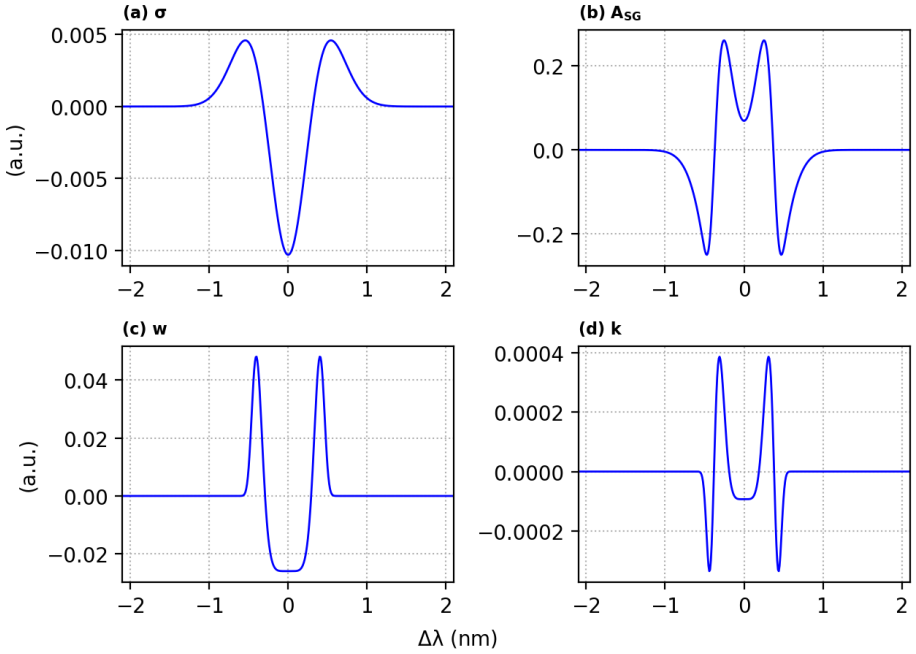


Figure 7.9: Partial derivatives obtained by perturbing each of the parameters of the parametrised SLF by 5 %. Note the different scales of the y-axes.

The spectral structures caused by the perturbations of the parameters of the SLF are obtained by convolving the water vapour cross-section with the different partial derivatives. Figure 7.10 presents the resulting spectra which show pronounced structures between 440 and 450 nm similar to the ones observed in the residual spectrum above showing that they can in principle be explained by an insufficient characterisation of the instrumental SLF. These spectra were then included as PSAs in the DOAS fit to possibly account for such structures.

Figure 7.11 compares the fit residuals without (a) and with (b) PSAs for the variations of the slit function parameters included in the DOAS fit for the example spectrum from above. However, only a marginal improvement is observed for the high frequency structures and also the fit RMS is only reduced by ca. 2 %. This can be explained by the fact that the measured SLF which was used to derive the parameters and their variations originated from a different instrument and was apparently too different from the actual SLF of the here-used instrument. As a consequence, the derivatives of a "wrong" parametrisation could not exactly reproduce the residual structures for this specific instrument. Nevertheless, the reason for the observed residual structures is most likely understood and a possible tool to deal with them was presented. For future instrument setups, it is therefore highly recommended to accurately characterise the instrumental slit functions, for instance by measurements using spectral line lamps.

As an application-oriented solution to remove the structures from the DOAS retrieval, the high frequency structures are now extracted from the fit residuals

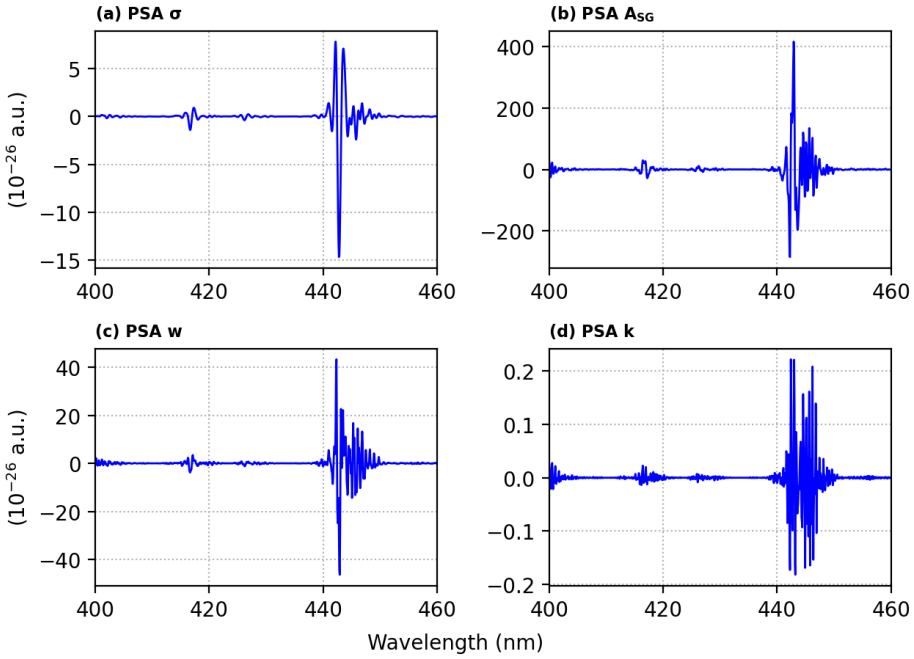


Figure 7.10: Spectral structures in the wavelength range between 400 and 460 nm caused by variations of the different slit function parameters.

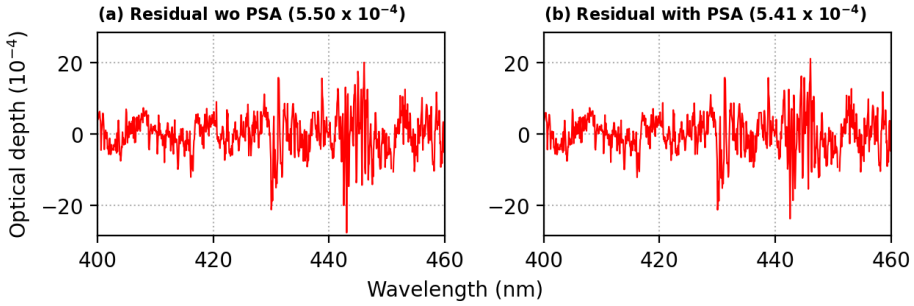


Figure 7.11: Panels (a) and (b) show the fit residuals of the same spectrum as above (see Fig. 7.6) for the scenarios without and with the determined PSAs for the variations of the slit function parameters included in the DOAS fit, respectively.

and included in the DOAS fit as pseudo absorber. For that, five residual structures of spectra recorded at 3° elevation on different days with relatively weak glyoxal and strong water vapour absorptions were extracted and averaged. These five residual spectra are shown in Fig. 7.12a, together with the mean structure which is finally included as PSA in the DOAS fits performed in the spectral range from 400 to 460 nm. Panel (b) depicts the fit residual of the fit for the used example spectrum when the PSA for the residual structure is included in the fit. Obviously, the high frequency structures are removed and only the strong structure at around 430 nm remains. Also the fit RMS is largely improved by approximately 16 %.

It should be stressed that including a residual structure in the DOAS fit can have certain pitfalls, such as artificially increased or decreased glyoxal dSCDs.

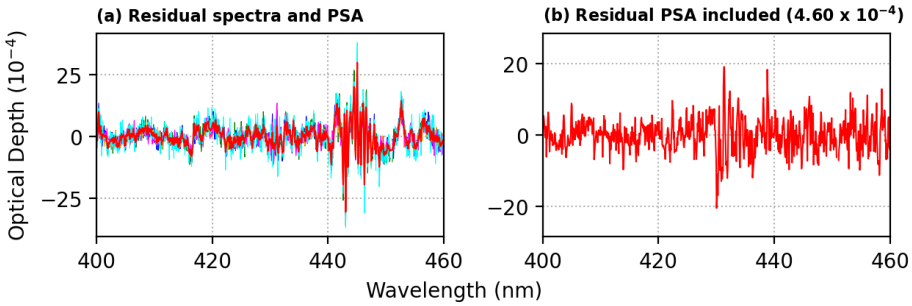


Figure 7.12: The different coloured curves in panel (a) represent the five residual spectra which were used to derive the PSA to account for the high frequency structures in the DOAS fit. The resulting PSA is shown in red. Panel (b) presents the residual of the fit for the example spectrum when the PSA is included.

However, various tests showed that this can be avoided by a careful selection of the used residuals. The main findings of these additional tests are summarised hereafter:

1. The structures of the PSA do not (strongly) correlate with the glyoxal absorption cross-section.
2. Similar NO_2 and glyoxal dSCDs are obtained with and without including the PSA in the fit. This is valid for days with low and high glyoxal abundances.
3. Clear improvements of the fit RMS and residual structures are achieved.
4. The fit coefficients of the PSA correlate well with the H_2O dSCDs.
5. The influence of the fitted PSA becomes weaker with increasing elevation as the H_2O absorptions get weaker.

These findings justify to include the mean residual structure as pseudo absorber in the DOAS fits of glyoxal and NO_2 in order to minimise the residual structures caused by strong H_2O absorptions in combination with an imperfect characterisation of the instrumental SLF. Such pseudo absorbers accounting for high frequency structures in the fit residuals were derived and included in the DOAS fits of NO_2 and glyoxal for all instruments used in this work (compare Tables 9.2 and 9.3), except for the Mainz instrument for which no such structures were observed.

7.2.2 *NO_2 Pukite terms under low NO_2 conditions*

For polluted stations like Bayfordbury, NO_2 Puḳīte terms (Sect. 3.1.5.2) have to be included in the DOAS fits of glyoxal since strong NO_2 absorptions lead to residual structures which influence the fit in an unfavourable way. An example of a glyoxal fit for a spectrum recorded at Bayfordbury is presented in Fig. 7.13. The NO_2 dSCD retrieved for this spectrum is 8.11×10^{16} molec cm^{-2} and thus

relatively high. Without NO_2 Puḳīte terms included in the fit, the residual (a) shows strongly pronounced spectral structures that can be clearly attributed to absorption features of NO_2 (compare cross-section in panel (e)). As becomes evident in panels (c) and (d), the usage of linear and square Puḳīte terms for NO_2 in the DOAS fit removes these structures and a clear glyoxal signal can be retrieved. Also the fit RMS is reduced by ca. 20 % in that way.

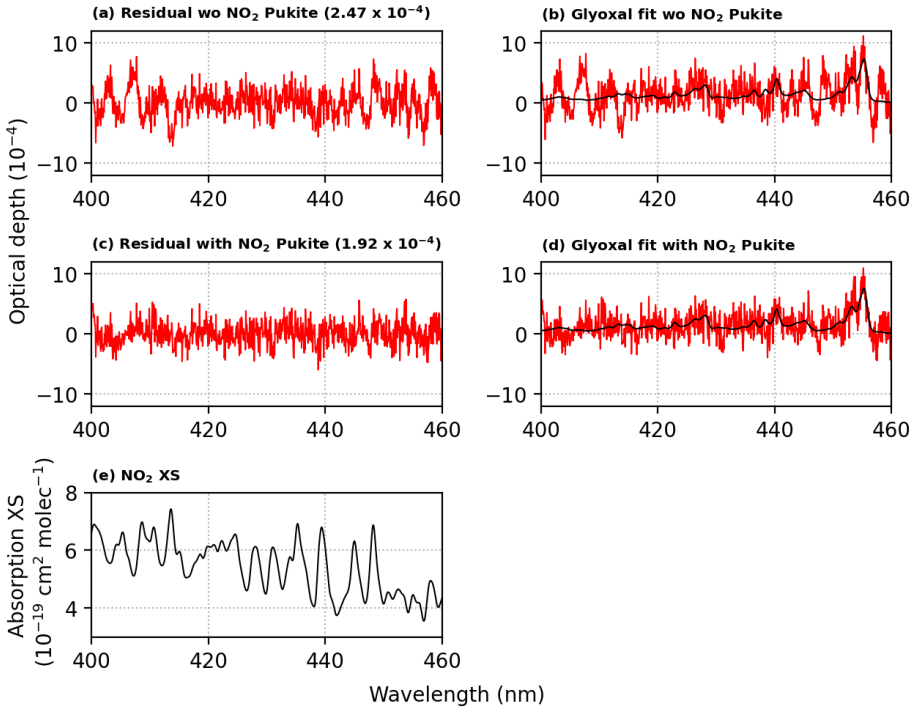


Figure 7.13: Residuals and glyoxal fits for a spectrum recorded by the Bayfordbury instrument at 3° elevation on 04 August 2019 at 11:57 LT. Panels (a) and (c) show the residuals for the cases without and with NO_2 Puḳīte terms included in the fit, respectively. Panels (b) and (d) show the corresponding glyoxal fits, where black lines indicate the fitted cross-sections, while red lines show the fit plus residual. Lastly, panel (e) depicts the NO_2 cross-section in the spectral range from 400 to 460 nm.

For the analysis of the ATTO measurements, Puḳīte terms for NO_2 are in principle not needed as only very small NO_2 abundances occur at this pristine location. However, since consistent fit settings for all stations are desirable, the effects of these terms on the DOAS fits of glyoxal under low NO_2 conditions were briefly investigated.

In order to do so, data from the ATTO 298 m instrument of the first measurement period in 2019 (compare Sect. 13.1) were analysed using two different fit scenarios based on the standard settings from Table 9.2. Note that this investigation on the Puḳīte terms for NO_2 in the glyoxal fit was performed at an early stage of this work and thus before the optimisation of the fitting interval (compare Sect. 7.2.3). Hence, a different fitting interval was used compared to the standard fit settings, namely 400 to 460 nm. The first scenario did not include NO_2 Puḳīte terms, while the second one included the respective linear and square Puḳīte

terms. The results for spectra recorded at 3° elevation are presented in Fig. 7.14 as time series of glyoxal dSCDs (a), differences (b) and ratios (c) of the retrieved glyoxal dSCDs between both scenarios and the fit RMS (d).

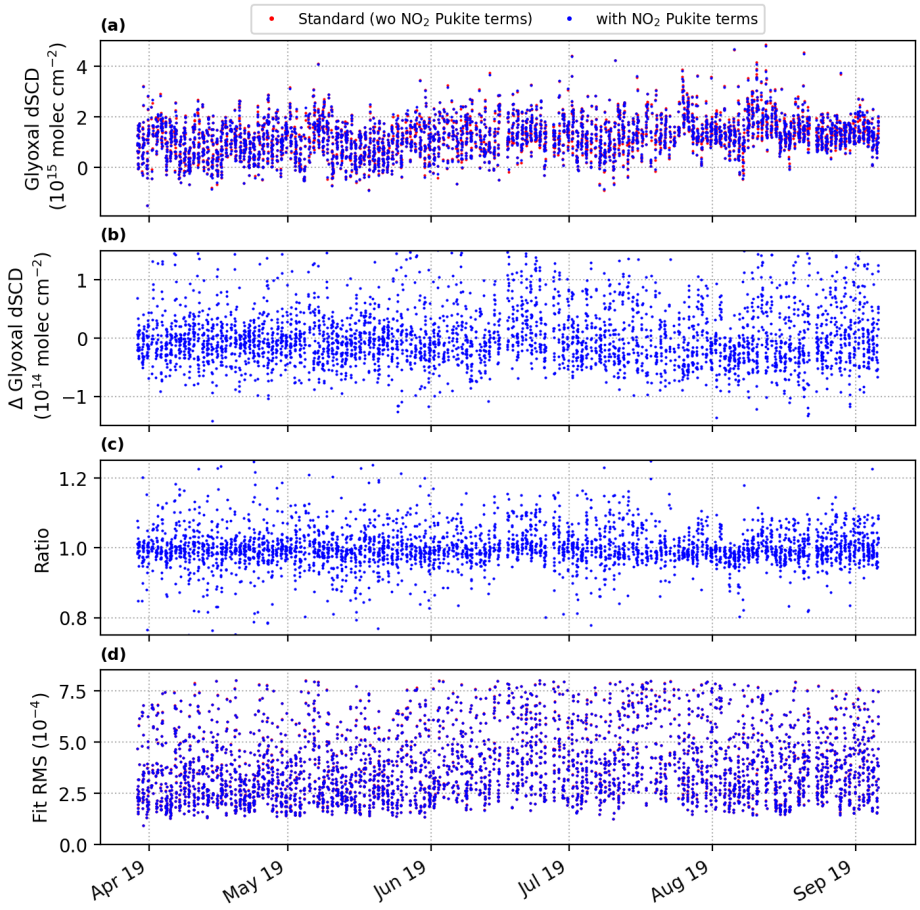


Figure 7.14: Time series of glyoxal dSCDs (a), differences (b) and ratios (c) of the retrieved glyoxal dSCDs and the fit RMS (d) for the two different fit scenarios as described in the text for the ATTO 298 m instrument at 3° elevation.

The retrieved glyoxal dSCDs are hardly affected since the differences and ratios between both scenarios mainly scatter around 0 and 1, respectively. Mean values of 0.01×10^{14} molec cm^{-2} and 1.01 are obtained for the two quantities. Further, the fit RMS is almost not influenced with observed changes of less than 1 %. Standard deviations of 7.30×10^{14} molec cm^{-2} are obtained for the retrieved glyoxal dSCDs for both scenarios indicating that also the scatter of the retrieved glyoxal dSCDs stays unchanged. All these findings are also valid for the not presented elevations, where also no significant effects were observed. Nevertheless, short individual periods with slightly larger effects occur. For instance, systematically larger glyoxal dSCDs are retrieved in the end of June if Puḳīte terms are included in the fit. However, this effect is still rather small and thus negligible. All described findings remain valid even if days with particularly low or high glyoxal abundances are considered. Lastly, also observed diurnal patterns in the glyoxal dSCDs stay the same.

Based on the brief case study presented here, it was concluded that the inclusion of Puḳīte terms for NO_2 in the glyoxal fit under low NO_2 conditions does not negatively affect the fit results. Since consistency is a major aim of this work and NO_2 Puḳīte terms are needed at the polluted stations, linear and square terms are also included in the DOAS fits for glyoxal at ATTO and thus at all stations (compare Table 9.2). Since the effect of Puḳīte terms is mainly important for the results of weak or less abundant absorbers, no such terms are included in the standard settings of the NO_2 fits as presented in Table 9.3.

7.2.3 *Choice of the fitting interval for glyoxal and NO_2*

The fitting interval from 400 to 460 nm, which was used so far for the glyoxal and NO_2 fits, contains several water vapour absorption bands. As explained above, these can cause problems in the DOAS fit in the case of strong water vapour absorptions. Although these are largely solved by including a PSA for the residual structure in the fit, the used fitting range might have negative side-effects:

1. The interval starting at 400 nm cuts off half an H_2O absorption band (compare red vertical line in Fig. 7.15) which is not optimal.
2. Studies showed that some of the absorption bands have wrong relative absorption strengths amongst each other. As a consequence, the two main bands in this spectral range are fitted separately.
3. The rather long spectral interval includes many strong Fraunhofer lines which could potentially lead to problems in the DOAS fit due to the detector non-linearity (compare Sect. 4.2.1.2).

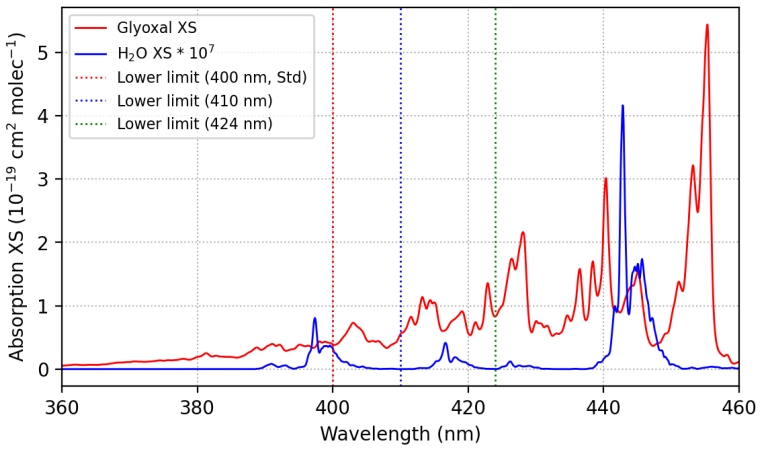


Figure 7.15: Absorption cross-sections of glyoxal (red) and water vapour (blue) between 360 and 460 nm. The water vapour cross-section was scaled with 1×10^7. The lower boundaries of the here-tested fit intervals are marked by coloured vertical lines.

To study these effects, two shorter fit intervals were tested. The first one starts at 410 nm and excludes the half water vapour band at 400 nm from the fit. In this

case, two major water vapour absorption bands at 416 and 442 nm are included in the fit which are fitted separately as outlined earlier. The second interval starts at 424 nm and includes, besides the main band at 442 nm, only one minor additional absorption band around 426 nm and thus removes the 416 nm band from the fit. For this interval, only one H_2O cross-section was included in the fit. The upper boundary at 460 nm was unchanged for both tested intervals. As both intervals are shorter than the standard one, less Fraunhofer lines are included in either case.

Again data from the ATTO 298 m instrument of the first measurement period in 2019 (compare Sect. 13.1) were used for these tests since the described effects are strongest for this instrument. It should be noted that the Pukīte terms of NO_2 were included in these fits. The results for spectra recorded at 3° elevation are presented in Fig. 7.16 while the results for the 30° spectra are shown in Fig. B.5 in the appendix.

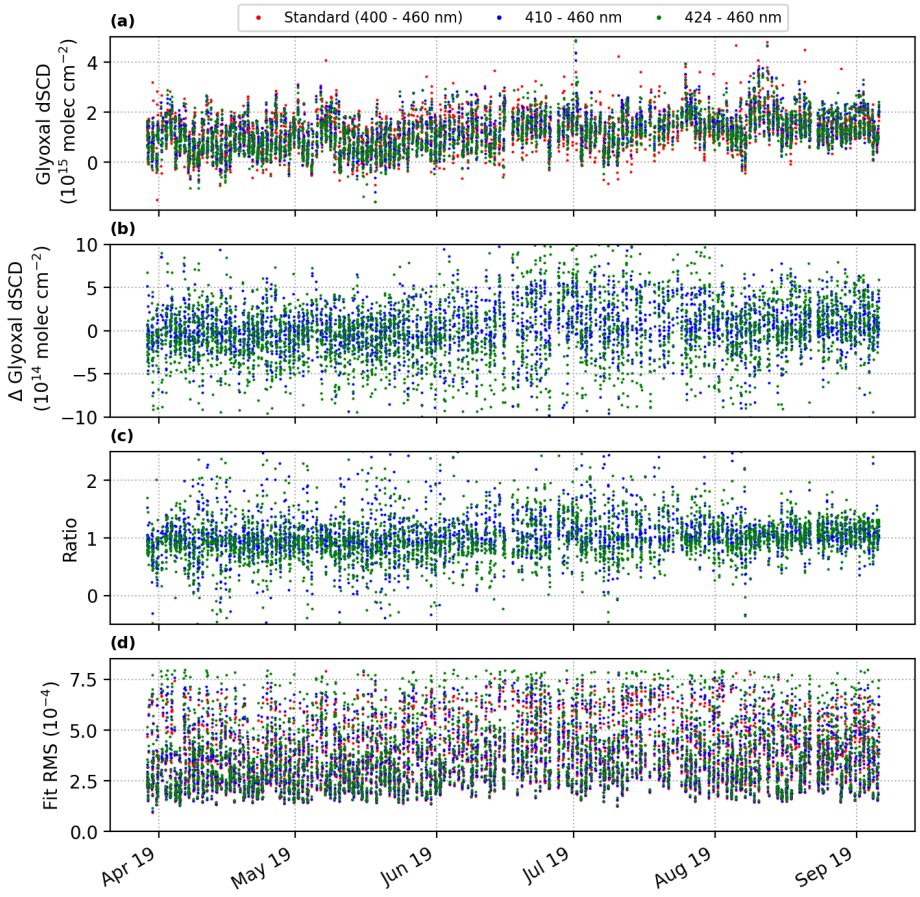


Figure 7.16: Time series of glyoxal dSCDs (a), differences (b) and ratios (c) of the retrieved glyoxal dSCDs and the fit RMS (d) for the different fit intervals described in the text for the ATTO 298 m instrument at 3° elevation.

Both tested fit intervals clearly improve the scatter of the retrieved glyoxal dSCDs compared to the standard scenario. While standard deviations of 6.88×10^{14} and 7.13×10^{14} molec cm^{-2} are obtained for the intervals starting at 410 and 424 nm, respectively, the standard deviation is 7.38×10^{14} molec cm^{-2} for

the standard interval. This is also valid for higher elevation angles, where even larger improvements are observed. Interestingly, the lowest scatter is found for the 410 nm interval. For this interval, the improvement is roughly 7 % at 3° elevation and ca. 37 % at 30° elevation. This decreased scatter is mainly due to less particularly low and high glyoxal dSCDs which are most likely outliers of the standard scenario as it becomes evident by the red dots in panel (a) of Fig. 7.16. The latter is especially pronounced at higher elevations (compare panel (a) in Fig. B.5 in the appendix).

At 3° elevation, the differences between the glyoxal dSCD retrieved with the shorter intervals and the standard interval mainly scatter around 0 and mean values of 0.08×10^{14} and -0.37×10^{14} molec cm^{-2} are calculated for the intervals starting at 410 and 424 nm, respectively. Besides this, a slight tendency towards larger dSCDs with increasing glyoxal dSCDs can be observed for the shorter fitting ranges. However, based on the presented time period, this is only a minor effect. The observed scatter in the absolute differences is relatively large for the shortest fitting interval (green dots in Fig. 7.16b). The ratios at 3° elevation mainly scatter around 1 (Fig. 7.16c). Since the mean values were found to be very sensitive to outliers caused by relatively large differences at rather low glyoxal dSCDs, median ratios are used and values of 1.00 and 0.96 are obtained for the intervals starting at 410 and 424 nm, respectively. Similar patterns are found for the other low elevation angles. However, at elevations above 10°, systematically lower glyoxal dSCDs are retrieved using the two shorter fit intervals. At 30°, median ratios of 0.69 and 0.61 are found, while the corresponding absolute differences are still close to 0 with values of -0.16×10^{14} and -0.45×10^{14} molec cm^{-2} for the 410 and 424 nm intervals, respectively (compare panels (b) and (c) in Fig. B.5 in the appendix).

The two shorter fit intervals yield larger fit RMS-values. While the observed increase is on average less than 3 % at 3° for the 410 nm interval, the fit RMS is significantly enhanced for the interval starting at 424 nm, where the relative increase is 10 % on average. For the enhancement of the fit RMS, no dependence on the elevation angle is observed.

Brief investigations of the diurnal variations of the retrieved glyoxal dSCDs showed that overall smoother diurnal patterns are obtained with the shorter fitting ranges, with the interval starting at 410 nm showing the smoothest diurnal variations. This is especially the case for spectra that show pronounced residual structures around 430 nm (compare Fig. 7.6) which are further studied in the next section.

Overall, the fit interval from 410 to 460 nm shows the smallest scatter of the retrieved glyoxal dSCDs, leads to the smoothest daily patterns and enhances the fit RMS only slightly compared to the previously applied standard interval. Further, it should be mentioned that the interval starting at 424 nm excludes several glyoxal absorption features from the fit, while not many structures exist between 400 and 410 nm (compare Fig. 7.15). For both tested shorter fit intervals, spectra with clear glyoxal signals are almost not affected by the changed fit interval. Based on these findings, the fit interval from 410 to 460 nm was selected

as new standard for the glyoxal and NO_2 fits performed in this work for all stations (compare Tables 9.2 and 9.3).

7.2.4 *Spectral gaps for the treatment of a residual structure around 430 nm*

As already mentioned in the previous sections, the fits of glyoxal and NO_2 sometimes show a pronounced residual structure at around 430 nm. This structure seems to be characteristic for the used AVANTES spectrometers and occurs for many TubeMAX-DOAS instruments, while the actual strength and spectral position of the structure varies between the instruments. Figure 7.17 depicts several examples of this structure in panels (a) to (e), all obtained for the ATTO 298 m instrument at 3° elevation during the first measurement phase in 2019. One possible factor causing this structure is a pronounced Fraunhofer line in the solar spectrum at ca. 430 nm as shown in panel (f). Nevertheless, even after various investigations, the cause of this structure is still not understood. However, one consistent observation in these investigations was that it mostly appears under very bright (high measured intensities) and clear sky conditions.

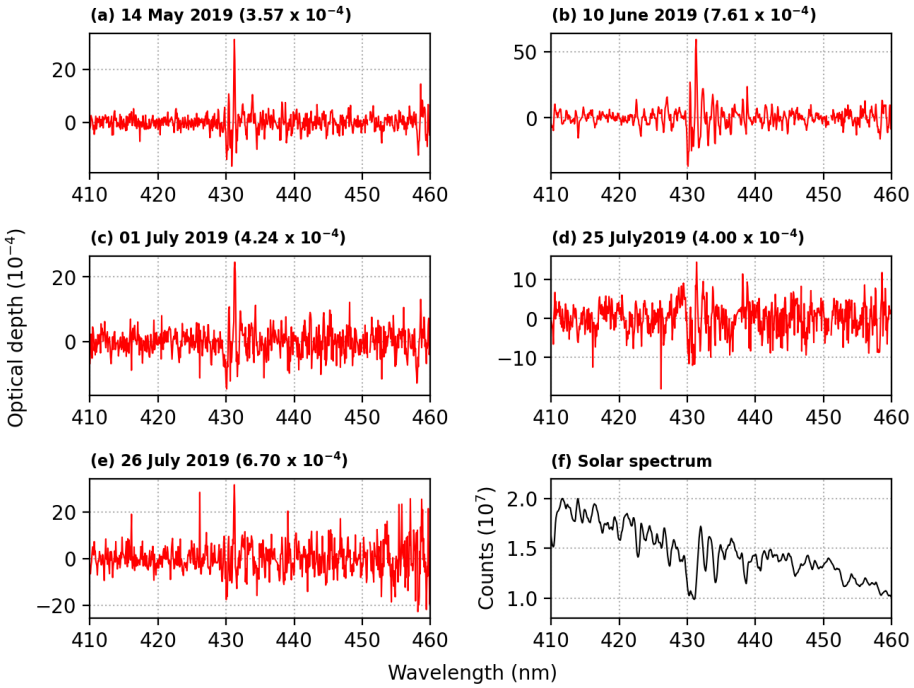


Figure 7.17: Panels (a) to (e) show examples for the observed residual structure at around 430 nm for the ATTO 298 m instrument at 3° elevation. The solar spectrum at instrument resolution in the affected spectral range is depicted in panel (f).

In addition, tests showed that this structure leads to spectral interferences in the DOAS retrieval having adverse effects on the glyoxal and NO_2 fits under specific measurement conditions. Thus, it is desirable to remove this structure from the fit which appears to be very challenging. Firstly, because the cause of this structure is not clear. Secondly, the exact shape and spectral position varies even for one instrument as becomes evident by comparing panels (a) to (e) of

Fig. 7.17. As a consequence of the latter, the simple removal by including a PSA of this structure in the fit, like done for the water vapour residual structures, is not possible. A last possible solution could be the inclusion of spectral gaps in the DOAS fits in the affected spectral range. Since the inclusion of gaps can have additional implications on the spectral fit, the gaps have to be chosen carefully and should be as small as possible. Based on manual selection, the following three gaps were tested: 429.6 to 432 nm (full), 429.8 to 432 nm (wide) and 430 to 431.6 nm (small). An overview of the applied gaps is presented in Fig. 7.18, together with the glyoxal cross-section in the spectral range of interest. The PSA for the water vapour residual structure was already included in the fit which was performed in the optimised fit interval from 410 to 460 nm for the same data of the ATTO 298 m instrument as above to test the influence of the gaps on the fit results.

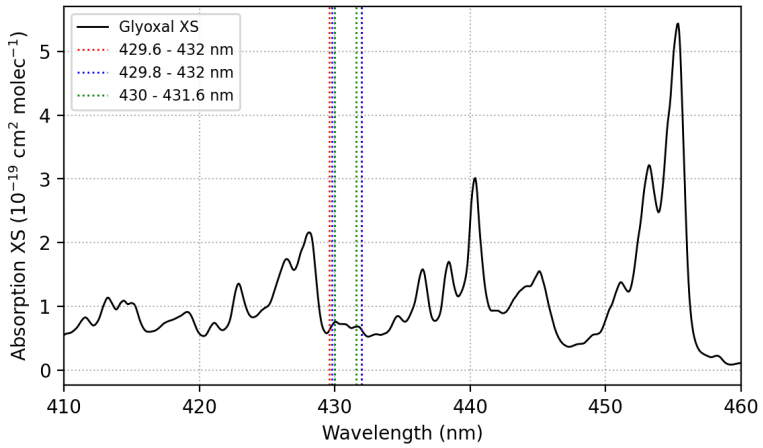


Figure 7.18: Absorption cross-section of glyoxal and tested gaps in the spectral range between 410 and 460 nm.

Again, the effects of these gaps on the retrieved glyoxal dSCDs, their scatter and the fit RMS were investigated. The results obtained with the different described fit scenarios for spectra recorded at 3° elevation are presented in Fig. 7.19.

All three gap scenarios slightly increase the scatter of the retrieved glyoxal dSCDs. While the standard deviation is 6.88×10^{14} molec cm^{-2} for the scenario without gap, standard deviations of 6.98×10^{14}, 6.98×10^{14} and 6.92×10^{14} molec cm^{-2} are obtained for the full, wide and small gap scenarios, respectively, corresponding to increases of less than 2 %. This is also valid for higher elevation angles, where similar enhancements are observed. However, for the smallest gap, at higher elevations the effect is slightly stronger pronounced and in the order of 5 %.

The differences between the glyoxal dSCDs retrieved with the standard scenario and the different gap scenarios mainly scatter around 0. At 3° elevation, mean differences of 0.11×10^{14}, 0.12×10^{14} and 0.67×10^{14} molec cm^{-2} are calculated for the three gap scenarios. Interestingly, the largest effect is found for the smallest included gap. The cyan dots in panel (b) show that this average tendency towards larger dSCDs for this scenario is caused by many individual measurements with

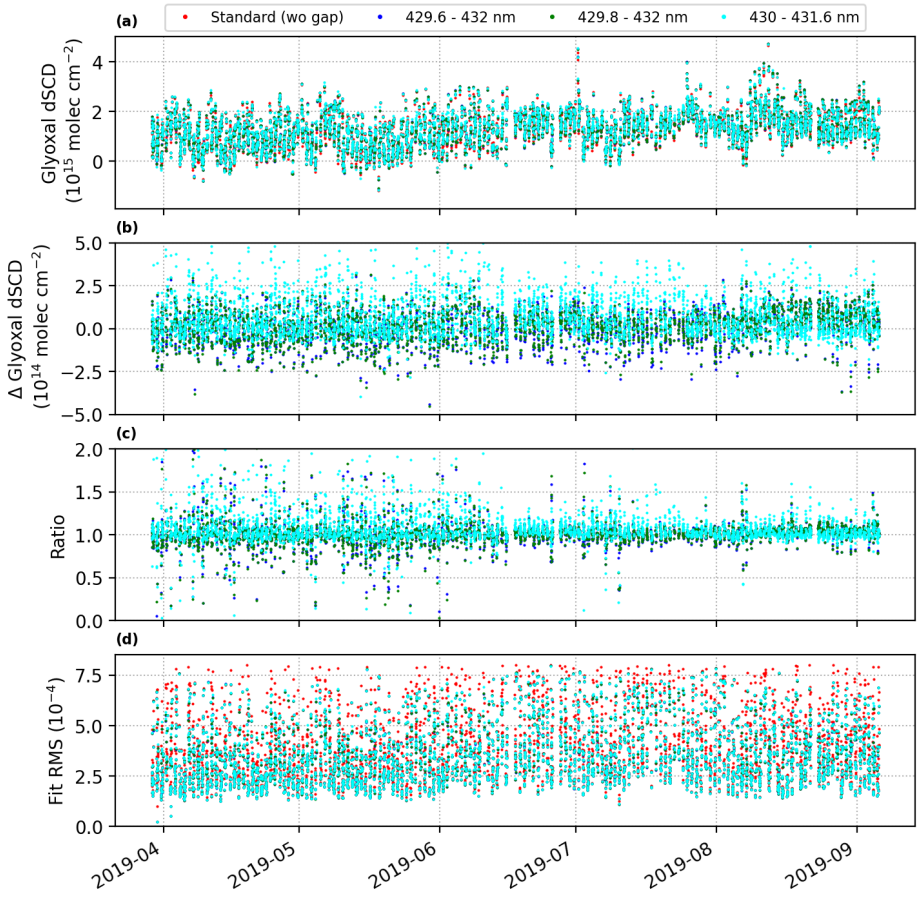


Figure 7.19: Time series of glyoxal dSCDs (a), differences (b) and ratios (c) of the retrieved glyoxal dSCDs and the fit RMS (d) obtained with different gaps included in the fit as described in the text for the ATTO 298 m instrument at 3° elevation.

increased dSCDs by above 2.5×10^{14} molec cm^{-2} in case of an included gap. As elaborated below, these measurements are associated with large fit RMS-values. Like already seen for the investigations regarding the fit interval, a slight tendency towards larger dSCDs with increasing glyoxal dSCDs can be observed for the two larger gaps. Again, this is only a minor effect. The ratios between the dSCDs retrieved with and without gaps mostly scatter around 1 (compare Fig. 7.19c). Median ratios of 1.01, 1.01 and 1.03 are calculated for the full, wide and small gap scenarios, respectively. No significant dependence on the elevation angle is observed for both differences and ratios.

Since fits with strong residual structures have large fit RMS-values, the fit RMS is largely improved when spectral gaps removing these structures are introduced to the fit (Fig. 7.19d). As a consequence, largest improvements are observed for particularly high fit RMS-values. On average, the fit RMS is improved by 10 % using either of the gaps, while marginally smaller improvements are observed for the smallest gap. For the fit RMS, no dependence on the elevation angle is observed.

To gain more detailed insights into the relation between the changes of the dSCDs and the fit RMS, the glyoxal dSCDs obtained with the different gap

scenarios are plotted against the dSCDs of the standard scenario in Fig. 7.20, where the fit RMS of the standard scenario is colour-coded.

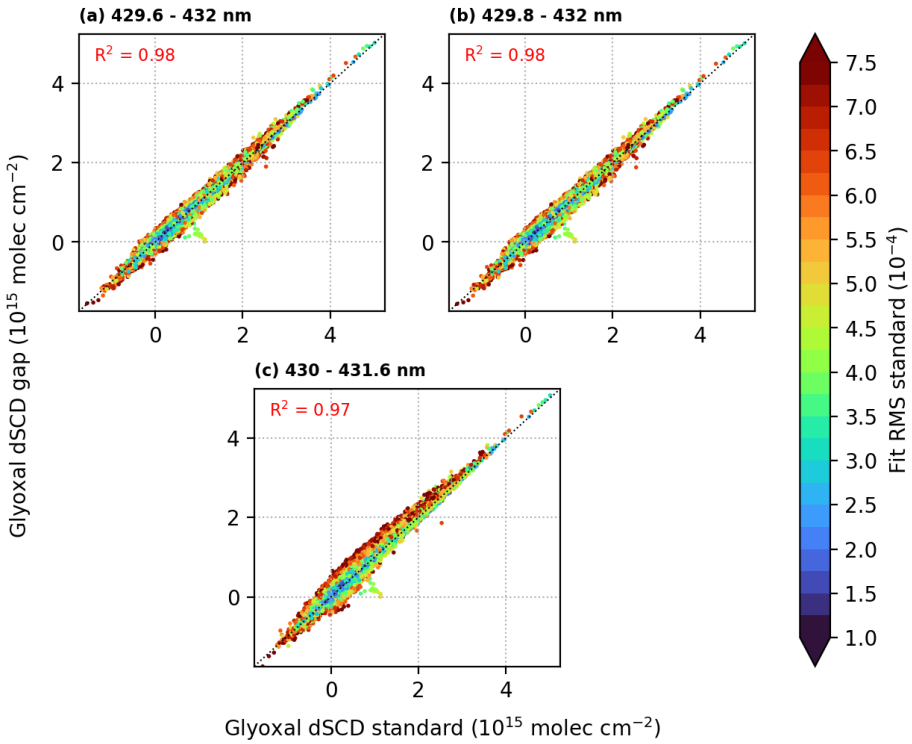


Figure 7.20: Correlation between glyoxal dSCDs retrieved with different gap scenarios and the standard scenario. The fit RMS of the standard scenario is colour-coded, while 1-to-1 lines are marked by black dashed lines.

Consistent to the previous findings, for all three tested gaps, good correlations are observed and the dSCDs are overall not largely affected by the gaps. Especially, at low fit RMS-values the individual data pairs are close to the 1-to-1 lines. However, a closer look reveals that for the smallest gap, whose results are shown in panel (c), larger deviations are found for large fit RMS-values, where a tendency towards an increase of the dSCDs is observed. This is consistent to the findings above. These large fit RMS-values are associated with fit residuals that show strong residual structures. Interestingly, as becomes evident in panels (a) and (b), the complete removal of the residual structure by relatively large gaps does not strongly influence the retrieved dSCDs. Contrary to that, an incomplete removal of the residual structure affects the spectral fit of glyoxal stronger. This is most likely caused by spectral interferences of the remaining part of the residual structure with a small spectral feature of glyoxal close to 430 nm (compare glyoxal cross-section in Fig. 7.18).

Altogether, the use of a spectral gap does not seem to be necessary, as no significant changes in the glyoxal dSCDs occur if the already optimised fit settings are used. Only the fit RMS can be significantly improved by introducing a gap. Therefore, it was decided to not include a spectral gap at 430 nm in the standard fit settings of glyoxal and NO_2. Nevertheless, the retrievals of both species were run for two scenarios, namely using the standard settings summarised in Tables 9.2

and 9.3 as well as a test scenario where the gap from 429.8 to 432 nm is included. Brief investigations for all stations showed that besides enhanced scatter of the retrieved dSCDs no benefits are found for this test scenario and thus it is not further considered in this work. Finally, it should be mentioned that for some instruments further smaller spectral gaps had to be included in the fits of several species to deal with so-called hot pixels which are pixels that produce an above-average number of electrons and thus create strong spikes in the measured spectra. These gaps are described in detail in the respective measurement overview sections in the analysis part of this work.

8

FURTHER IMPROVEMENTS OF THE DOAS RETRIEVAL

Besides the optimisation of the DOAS fit settings presented in the two previous chapters, also the effect of the assumed slit function shape of the used MAX-DOAS instruments (Sect. 8.1) and potential influences of changing spectrometer temperatures (Sect. 8.2) on the DOAS fit results were investigated. The results of these two studies are briefly summarised in this chapter.

8.1 SPECTRAL CALIBRATION ASSUMING A SUPER GAUSSIAN SLIT FUNCTION

As outlined in Sect. 3.1.5.3, during the measurements using a spectrometer with a comparably low spectral resolution the highly structured spectra of scattered sunlight (including trace gas absorptions) are smoothed by the instrument's slit function. In cases for which no measurements of the instrumental SLF are available, it can be determined by comparing a HR solar spectrum with a measured spectrum and assuming a certain parametrisation for its shape. Since in this work a solar reference spectrum from Kurucz et al. (1984) is used, this procedure is called "Kurucz-Fit" in the following. Besides the spectral calibration, parameters describing the assumed shape of the SLF are output by this fit. Often a simple symmetric Gaussian of the form

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{8.1}$$

is assumed which is described by one single parameter, namely its standard deviation σ. However, Beirle et al. (2017) suggest the usage of a Super Gaussian function which can be expressed as

$$SG(x) = A(w,k) \cdot \exp\left(-\left|\frac{x}{w}\right|^k\right) \tag{8.2}$$

with two parameters, w and k, describing its width and shape, respectively. Further, $A(w,k)$ represents a factor needed to normalise the integral over SG to 1 which is required for the usage as SLF. The advantage of the Super Gaussian parametrisation is that it can describe a variety of different shapes and thus potentially better reproduce the correct shape of the instrumental SLF. Already Beirle et al. (2017) showed that also for the AVANTES spectrometers used in this work significant improvements of the Kurucz-Fit results can be obtained when using a Super Gaussian instead of a simple Gaussian. Since the parametrised SLF is also used to convolve the reference absorption cross-sections included in the DOAS fit, an imperfect characterisation of the instrumental SLF can markedly affect the DOAS fit results. Therefore, besides the influences on the Kurucz-Fit results, the effect on the retrieved trace gas dSCDs is examined in the following.

The Kurucz-Fits in this work are performed using the *QDOAS* software (compare Sect. 9.1) yielding the spectral calibrations of the respective spectrometers

and the (wavelength dependent) parameters describing the corresponding slit functions. Figure 8.1a compares the fit residuals of such Kurucz-Fits assuming Gaussian (blue) and Super Gaussian (red) SLFs for a zenith spectrum measured at Bayfordbury which was also used as universal reference for the analysis of the first measurement period (compare Table 4.1).

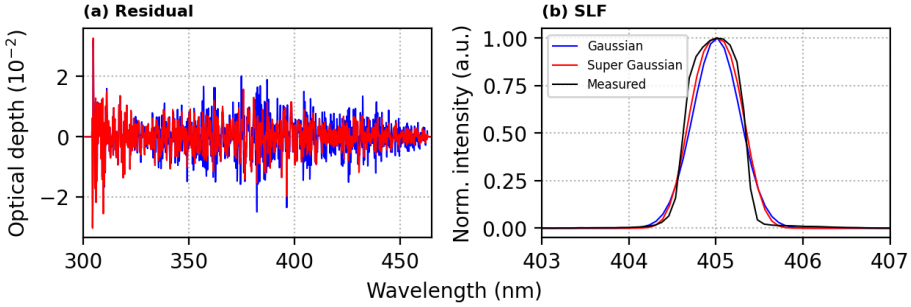


Figure 8.1: Panel (a) compares the residuals of Kurucz-Fits assuming Gaussian (blue) and Super Gaussian (red) SLFs for a zenith spectrum measured at Bayfordbury on 23 July 2019 at 12:12 LT. Panel (b) compares a measured SLF at 405 nm with a delta peak convolved with the Gaussian and Super Gaussian SLFs obtained from the Kurucz-Fits for the Bayfordbury instrument.

Obviously, the residual is improved if a Super Gaussian parametrisation is used. This is also reflected by the reduction of the fit RMS of the Kurucz-Fit by ca. 20 % which is 5.50×10^{-3} and 4.48×10^{-3} for the Gaussian and Super Gaussian SLF, respectively. The slit function parameters at 405 nm obtained from these fits were then used to convolve a delta peak. Since also for the Bayfordbury instrument no measured SLF is available, in Fig. 8.1b the obtained results are compared to the measured SLF from a different instrument which was already used in Sect. 7.2.1 and is supposed to be generally representative for the used AVANTES spectrometers. Consistent to the findings in Sect. 7.2.1, also the Super Gaussian slit function cannot perfectly describe the correct shape of the SLF. Nevertheless, some improvement compared to the simple Gaussian function is achieved as the shape of the Super Gaussian is slightly closer to the measured shape.

To study the influence of the choice of the parametrisation for the SLF on the resulting trace gas dSCDs, both parametrisations of the SLF were used to convolve the standard set of reference absorption cross-sections and DOAS fits for the standard trace gases were performed for both scenarios. Since these studies were carried out at an early stage of this work, only data until early March 2020 are included and the fit settings slightly deviate from the final settings summarised in Tables 9.1 to 9.5. This mainly affects the included absorption cross-sections, whereas the overall findings are expected to be not influenced by these differences.

The comparison of the dSCDs and fit RMS-values obtained for both scenarios shows that for trace gases retrieved in the UV spectral range, such as HCHO and O_4, the results are almost not affected by the choice of the SLF parametrisation. For these species, slopes and R^2-values of practically 1 are obtained for the

corresponding correlation plots[1]. For trace gases retrieved in the visible spectral range (e.g. NO_2 and glyoxal), however, the fit RMS is improved by choosing a Super Gaussian slit function in many cases. As an example, Fig. 8.2 depicts the comparison of NO_2 dSCDs (upper part) and the corresponding fit RMS (lower part) obtained assuming Gaussian (blue) and Super Gaussian (green) SLFs during the Kurucz-Fit for spectra measured at Bayfordbury at 3° elevation.

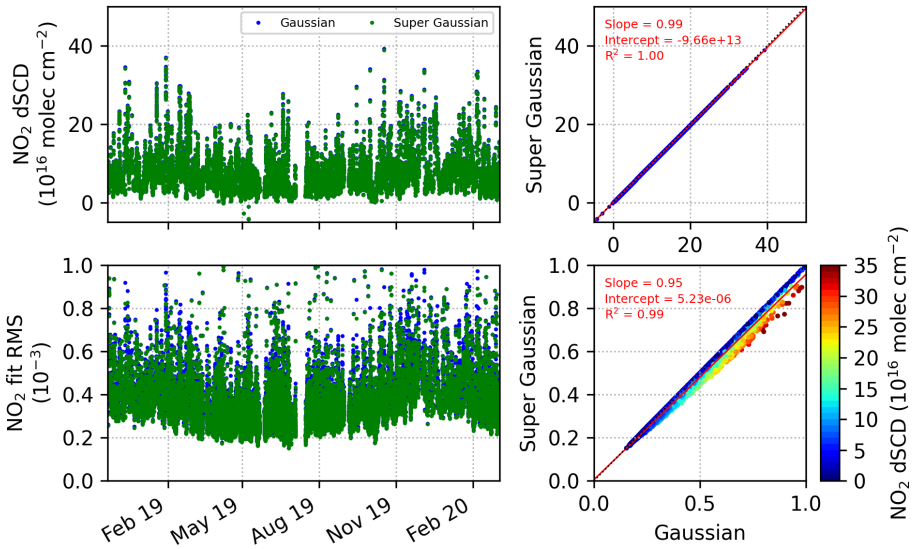


Figure 8.2: Comparison of NO_2 dSCDs (upper part) and corresponding fit RMS (lower part) obtained assuming Gaussian (blue) and Super Gaussian (green) SLFs for spectra measured at Bayfordbury at 3° elevation. Left panels show time series of the results for both scenarios, while the right panels show the corresponding correlation plots. The results of simple linear regressions are shown in red. NO_2 dSCDs are colour-coded in the lower right panel.

Largest improvements of the fit RMS are found for high NO_2 dSCDs, while the retrieved dSCDs are almost not affected by the choice of the SLF parametrisation. Nevertheless, marginally smaller values are obtained for the highest dSCDs. Similar findings hold for the DOAS fits of water vapour and, to a slightly smaller extent, glyoxal. Another observation is that the effects on dSCDs and fit RMS become weaker with increasing elevation[2]. Together with the finding that largest improvements are observed for high NO_2 dSCDs, this indicates an effect that is related to strong NO_2 absorptions causing residual structures in the fit residuals due to an imperfect characterisation of the SLF similar to the studies presented for water vapour in Sect. 7.2.1. This effect can obviously be minimised by assuming a Super Gaussian SLF. However, as already shown for water vapour in Sect. 7.2.1 and as it becomes evident from Fig. 8.1b, also the usage of a Super Gaussian parametrisation for the instrumental SLF does not perfectly represent the correct shape of the SLF. Therefore, it should be stressed again that for future measurements it is highly recommended to accurately characterise the instrumental slit functions by spectral line lamp measurements.

1 Not shown here in sake of shortness.
2 Not shown here in sake of shortness.

For the residual structures caused by strong NO_2 absorptions[3], further improvements of the fit RMS are achieved when Puḳīte terms for NO_2 (compare Sect. 3.1.5.2) are included in the DOAS fits performed in the visible spectral range. For the glyoxal retrieval, the latter has been shown to be essential for polluted stations, such as Bayfordbury, in Sect. 7.2.2.

Given the results described in this section, for the calibrations and convolutions of reference absorption cross-sections used for the final DOAS analyses symmetric Super Gaussian parametrisations of the SLFs were used for all instruments.

8.2 INFLUENCE OF CHANGING SPECTROMETER TEMPERATURE ON SPECTRAL CALIBRATION AND CONVOLUTION

Although the spectrometer temperatures of the utilised MAX-DOAS instruments are stabilised using Peltier units, in some specific cases the temperatures could not be kept stable over the entire measurement period. For instance at Bayfordbury, a defective fan of the instrument's cooling unit required an increase of the spectrometer temperature from 15 to 28 °C with a transition period where temperatures varied between 20 and 30 °C. For details, refer to Sect. 10.1 and Table 4.1. Naturally, these three temperature regimes lead to three periods for which different sets of DC and offset spectra are needed for the correction of the measured spectra. The extent to which also the optical properties, i.e. spectral calibration and SLF, of the instrument are affected by these temperature changes is briefly examined in this section.

Since the spectral calibration of the instrument might change due to such temperature changes, spectral calibrations were performed for all three temperature regimes. To compare the pixel-to-wavelength assignments of these three calibrations, the differences of the assignments of the second and third period with respect to the first one were calculated and are depicted in panel (a) of Fig. 8.3.

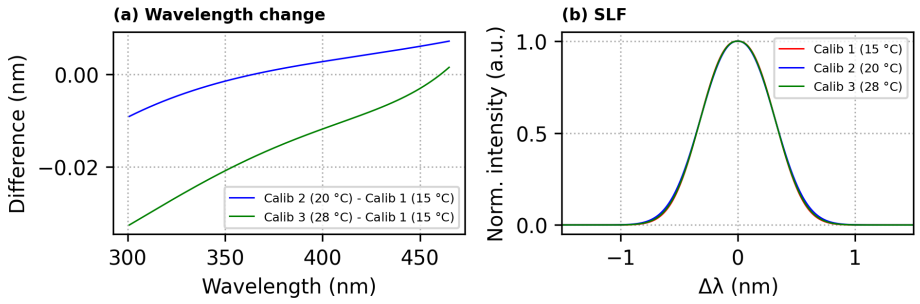


Figure 8.3: Panel (a) shows the differences in the pixel-to-wavelength assignments for the second (blue) and third (green) temperature regimes compared to the first one. The x-axis corresponds to the spectral calibration of the first period. Panel (b) compares the corresponding Super Gaussian SLFs at 341.5 nm (centre of the HCHO fitting window) for the three temperature regimes.

3 As outlined in Sect. 7.2.1, this is not the case for residual structures caused by strong H_2O absorptions.

While comparably small differences between the first and second calibration regimes are found, larger differences are obtained for the comparison between the first and third calibration periods as also the temperature difference between these regimes is larger. For the centre of the HCHO fitting range (341.5 nm), the differences are approximately -0.002 and -0.023 nm, while the differences are ca. 0.005 and -0.007 nm in the NO_2 fitting interval for the second and third period, respectively.

Figure 8.3b compares the (Super Gaussian) SLFs at 341.5 nm obtained from the Kurucz-Fits for the three temperature regimes. Apparently only marginal differences in the shapes of the slit functions for the different calibration regimes exist indicating stable optical characteristics of the used AVANTES spectrometers over a notable temperature range of ca. 15 °C.

Another interesting aspect is the extent to which the trace gas results are affected by the small changes in the spectral calibrations (compare Fig. 8.3a) caused by the changed spectrometer temperatures. To investigate this, spectral fits of HCHO and NO_2 were performed using one fixed spectral calibration, namely the one valid for the first temperature regime, for the whole measurement period. Note that similar to the previous section, these studies were conducted at an early stage of this work. Thus, only a limited data set for Bayfordbury was available and the fit settings slightly deviate from the final settings summarised in Tables 9.1 to 9.5. Again the general findings are expected to be not influenced by these small differences.

Figure 8.4a shows time series of the so-called "pre-shifts" of these HCHO (blue) and NO_2 (green) fits. The pre-shift is a spectral shift that is applied to the (interpolated) FRS (compare Sect. 9.1) of each individual spectral fit before the DOAS fit itself is performed in order to align the calibrations of the individual (interpolated) sequential FRS to that of the universal FRS.

During the period for which the used spectral calibration is valid (until end of April 2020), the pre-shifts are rather small and constant in both fitting windows. From late April to late June 2020, varying spectrometer temperatures were used. In this period large variability in the pre-shifts is found for both trace gas fits indicating effects of slightly changed wavelength calibrations caused by temperature variations. In late June 2020, the spectrometer temperature was fixed to 28 °C and the pre-shifts stay rather constant afterwards. Nevertheless, the pre-shifts in this period differ from the ones observed before the temperature was changed, whereby the UV spectral range (HCHO fit) is stronger affected and pre-shifts of around 0.021 and 0.006 nm are found from July to October 2020 in the HCHO and NO_2 fitting ranges, respectively. Comparing these values to the differences in the pixel-to-wavelength assignments between first and third calibration regime in the respective fitting ranges (compare Fig. 8.3a) shows that the spectral pre-shifts compensate reasonably well for slightly changed spectral calibrations in the DOAS fit. The reason for the decrease of the pre-shifts for both fitting ranges after October 2020 is not clear at the moment. Further, the spectral shift that is applied to the analysed spectrum is not affected by changed spectrometer temperatures as no distinct pattern is visible after April 2020 (compare Fig. 8.4b).

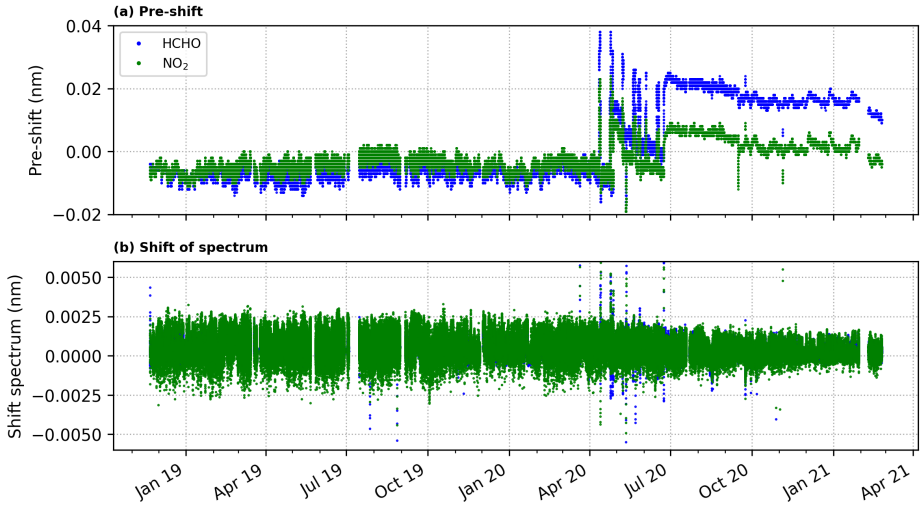


Figure 8.4: Panel (a) shows time series of the spectral pre-shifts of HCHO (blue) and NO_2 (green) fits performed for the Bayfordbury instrument using one fixed spectral calibration and SLF for the whole measurement period. Panel (b) depicts time series of the corresponding shifts applied to the measured spectra during the fits. For both plots, only data with elevations above 2°, fit RMS-values smaller than 1×10^{-3} and SZAs below 80° are considered.

Lastly, the HCHO results obtained using the calibration, which is valid for the first temperature regime, for the entire data set are compared to the results obtained for the case when the (correct) calibrations of the individual temperature regimes are used. For that, Fig. 8.5 shows time series of the corresponding HCHO dSCDs (a), the differences (b) and ratios (c) of these dSCDs and the corresponding fit RMS (d) at 3° elevation.

Apparently the effect of slightly changed spectral calibrations caused by changed spectrometer temperatures on the HCHO results is rather small as red and blue dots in panel (a) mostly overlap. Nevertheless, faintly larger dSCDs are obtained with the wrong calibration as the mean absolute difference for the dSCDs at 3° elevation is 0.22×10^{15} molec cm^{-2} which corresponds to less than 1 % of the mean formaldehyde dSCD (2.58×10^{16} molec cm^{-2}) in the considered measurement period. The mean ratio is 1.02 also indicating only a small effect. For the fit RMS, no clear pattern is found and the values can be slightly larger or lower for the case when a fixed calibration is used. Since the absolute difference is practically independent of the elevation angle the observed effect can be considered to be systematic but small. A last finding is that for the case when the (correct) individual calibrations are used, about 8 % more valid, i.e. fulfilling basic filter criteria (fit RMS smaller than 1×10^{-3}), DOAS fits are obtained.

In conclusion, it can be stated that changes of the spectrometer temperature in the order of ca. 10 to 15 °C lead to only small changes in the optical properties of the used AVANTES spectrometers. While the instrumental SLF is only marginally affected, small changes in the spectral calibration emerge from these temperature changes which are basically compensated by the spectral pre-shift of the DOAS analyses leading to only small changes in the final trace gas results. Consequently,

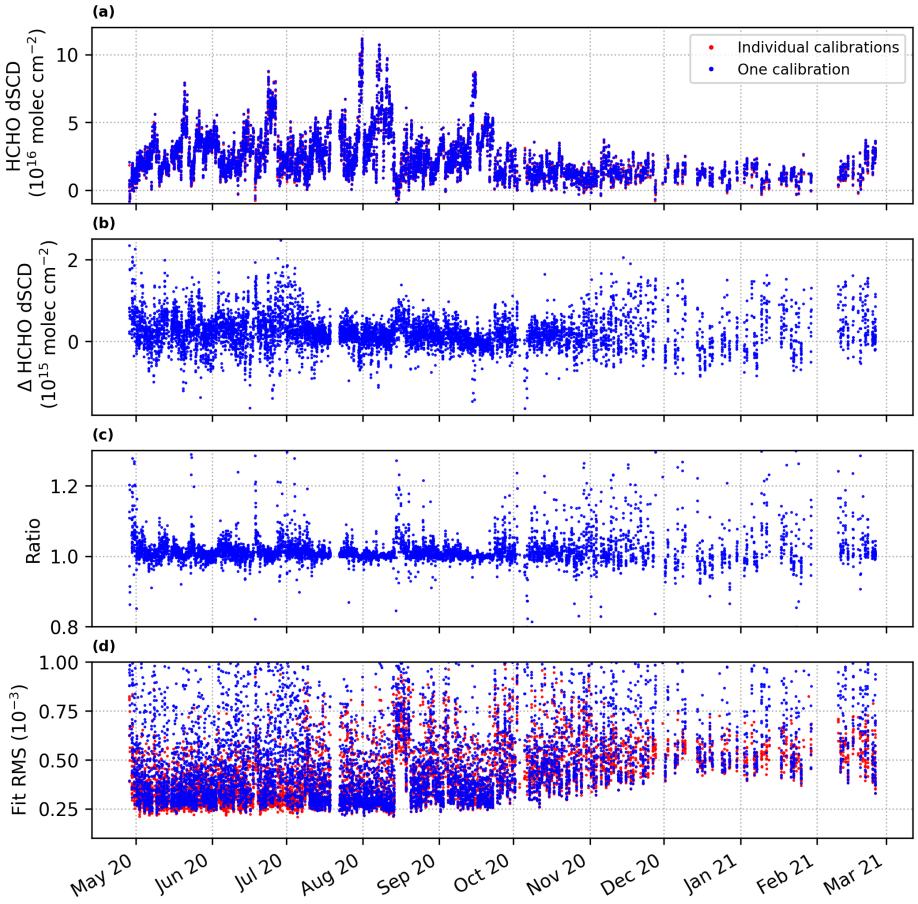


Figure 8.5: Time series of HCHO dSCDs (a), differences (b) and ratios (c) of the retrieved HCHO dSCDs and the corresponding fit RMS (d) obtained using individual adjusted calibrations (red) and one fixed calibration (blue) as described in the text for the Bayfordbury instrument at 3° elevation.

apparently it is no large issue to use one fixed calibration and instrumental SLF for the used AVANTES spectrometers for cases in which the spectrometer temperature is changed. However, despite these findings, three different spectral calibrations and SLFs were used for the DOAS retrievals for Bayfordbury in this work (compare Table 4.1) to assure high quality results for the entire measurement period.

Part IV

MEASUREMENTS OF TROPOSPHERIC FORMALDEHYDE

9 DATA ANALYSIS

This chapter outlines the analysis procedure of the spectra collected at the different measurement sites. First, the final DOAS fit settings which were optimised in the previous part are summarised and listed in Sect. 9.1. Further, the profile retrieval using MAPA and the applied settings are described in Sect. 9.2. This includes the presentation of a tool to identify inaccuracies in the elevation calibration by the MAPA results themselves and the description of how near-surface concentrations were obtained from the MAPA results. Section 9.3 summarises the filter criteria which were applied to the MAPA data before the results are visualised and discussed. Finally, a comparison of the air temperatures and near-surface winds of the ERA5 data set with the corresponding measurements of local weather stations is shown in Sect. 9.4.

9.1 SPECTRAL ANALYSIS

As outlined in Chapt. 3, the first step to obtain vertical column densities and profiles of trace gas concentrations (or aerosol extinction) from spectra of scattered sunlight is the spectral analysis using the DOAS principle. For this, version 3.4.4 of the *QDOAS* software package (Danckaert et al., 2017) was used.

First, the spectra are spectrally calibrated based on Fraunhofer lines of a solar reference spectrum (Kurucz et al., 1984) also using *QDOAS*. As outlined earlier, for the spectral calibration and the subsequent convolution of the absorption cross-sections a Super Gaussian shaped slit function is assumed (compare Sect. 7.2 and 8.1). The parameters describing this slit function are an output of the calibration process itself.

A large part of this work is the optimisation of the used DOAS fit settings, especially of formaldehyde and glyoxal, which was presented in the last chapters. Tables 9.1 to 9.5 summarise the final standard fit settings of all species relevant for this work. Some further specifications are given below.

As Fraunhofer reference spectrum the 90° spectra before and after the corresponding sequence are interpolated to the measurement time of the spectrum to be analysed. The absorption cross-sections of O_3 and NO_2 are convolved applying intensity weighted convolutions (compare Sect. 3.1.5.3), while standard convolutions are applied to the other cross-sections. For the Ring effect, two pseudo absorbers are derived from the universal FRS (compare Sect. 3.1.5.1). For the spectral retrievals of glyoxal and NO_2, the water vapour cross-section was split up into two parts to separate the absorption structures of the major bands. The results of the stronger one at around 442 nm from the NO_2 retrieval are further used as results for the H_2O abundance. Lastly, these fits also include a PSA accounting for residual structures which are attributed to water vapour absorptions obtained as described in Sect. 7.2.1.

Table 9.1: Overview of the standard fit settings for HCHO.

PARAMETER	SETTINGS
Fitting interval	324 – 359 nm
Trace gas	
HCHO	Meller and Moortgat (2000), 298 K
BrO	Wilmouth et al. (1999), 228 K
NO_2	Vandaele et al. (1998), 294 K, 220 K, linear and square Puḳīte terms, intensity weighted convolution
O_3	Serdyuchenko et al. (2014), 223 K, 243 K, linear and square Puḳīte terms, intensity weighted convolution
O_4	Thalman and Volkamer (2013), 293 K
Ring effect	Two pseudo absorbers derived from universal FRS
FRS	Sequential, interpolated to measurement
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

Table 9.2: Overview of the standard fit settings for glyoxal.

PARAMETER	SETTINGS
Fitting interval	410 – 460 nm
Trace gas	
NO_2	Vandaele et al. (1998), 294 K, linear and square Puḳīte terms, intensity weighted convolution
Glyoxal	Volkamer et al. (2005b), 296 K
O_3	Serdyuchenko et al. (2014), 223 K, intensity weighted convolution
H_2O	Rothman et al. (2013), 293 K, rescaled according to Lampel et al. (2015), shifted by 0.021 nm, split up
O_4	Thalman and Volkamer (2013), 293 K
Ring effect	Two pseudo absorbers derived from universal FRS
H_2O structure	PSA created from residual spectra
FRS	Sequential, interpolated to measurement
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

Table 9.3: Overview of the standard fit settings for NO_2.

PARAMETER	SETTINGS
Fitting interval	410 – 460 nm
Trace gas	
NO_2	Vandaele et al. (1998), 294 K, intensity weighted convolution
Glyoxal	Volkamer et al. (2005b), 296 K
O_3	Serdyuchenko et al. (2014), 223 K, intensity weighted convolution
H_2O	Rothman et al. (2013), 293 K, rescaled according to Lampel et al. (2015), shifted by 0.021 nm, split up
O_4	Thalman and Volkamer (2013), 293 K
Ring effect	Two pseudo absorbers derived from universal FRS
H_2O structure	PSA created from residual spectra
FRS	Sequential, interpolated to measurement
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

Table 9.4: Overview of the standard fit settings for O_4.

PARAMETER	SETTINGS
Fitting interval	352 – 387 nm and 338 – 370 nm
Trace gas	
NO_2	Vandaele et al. (1998), 294 K, intensity weighted convolution
O_3	Serdyuchenko et al. (2014), 223 K, intensity weighted convolution
O_4	Thalman and Volkamer (2013), 293 K
HCHO	Meller and Moortgat (2000), 298 K
BrO	Wilmouth et al. (1999), 228 K
H_2O	Gordon et al. (2022), 296 K
Ring effect	Two pseudo absorbers derived from universal FRS
FRS	Sequential, interpolated to measurement
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

Table 9.5: Overview of the standard fit settings for HONO.

PARAMETER	SETTINGS
Fitting interval	335 – 373 nm
Trace gas	
HONO	Stutz et al. (2000), 298 K
NO_2	Vandaele et al. (1998), 294 K, 220 K, linear and square Puķīte terms, intensity weighted convolution
O_3	Serdyuchenko et al. (2014), 223 K and 243 K, intensity weighted convolution
BrO	Wilmouth et al. (1999), 228 K
O_4	Thalman and Volkamer (2013), 293 K
HCHO	Meller and Moortgat (2000), 298 K
H_2O	Gordon et al. (2022), 296 K
Ring effect	Two pseudo absorbers derived from universal FRS
FRS	Sequential, interpolated to measurement
Polynomial	5^{th} order
Intensity offset	Constant and 1^{st} order term
Shift and stretch	Spectra are shifted and stretched (1^{st} order) against FRS

The fit settings for HONO are based on Wang et al. (2020), but were slightly modified. In particular, intensity weighted convolutions were applied for O_3 and NO_2 (including the respective Puķīte terms) instead of I_0-corrected convolutions (compare Sect. 3.1.5.3). Furthermore, different absorption cross-sections were used for O_3, BrO and water vapour in order to use a consistent set of cross-sections for the retrievals of all species. Note that the here-used HITRAN (2020) water vapour cross-section was not yet available when Wang et al. (2020) performed their study.

For some instruments, slight modifications of the standard settings, e.g. the introduction of spectral gaps caused by hot pixels, were unavoidable. If this was the case, details are outlined in the introduction section of the result chapter of the respective station. An aspect worth mentioning already here is that at ATTO limitations of the instrumental setup on the tower hinder measurements at 90° elevation. Therefore, spectra recorded at 88° and 87° are used as FRS for the 80 m and 298 m instrument, respectively.

Lastly, Fig. 9.1 presents typical fit results for the different species relevant for this work obtained with the standard fit settings summarised above. While for HCHO, glyoxal, NO_2 and O_4 fit examples for spectra recorded by the Bayfordbury instrument in July 2019 (for details refer to the figure caption) are shown, the HONO fit example is based on a spectrum measured by the Mohali instrument on 5 November 2020 since no HONO was detected at Bayfordbury (compare Chapt. 10). Obviously, the shorter integration time at Mohali (compare Table A.1 in the appendix) leads to a higher fit RMS caused by a stronger influence of

noise. The examples were chosen in a way that they represent cases with neither particularly high nor low trace gas abundances.

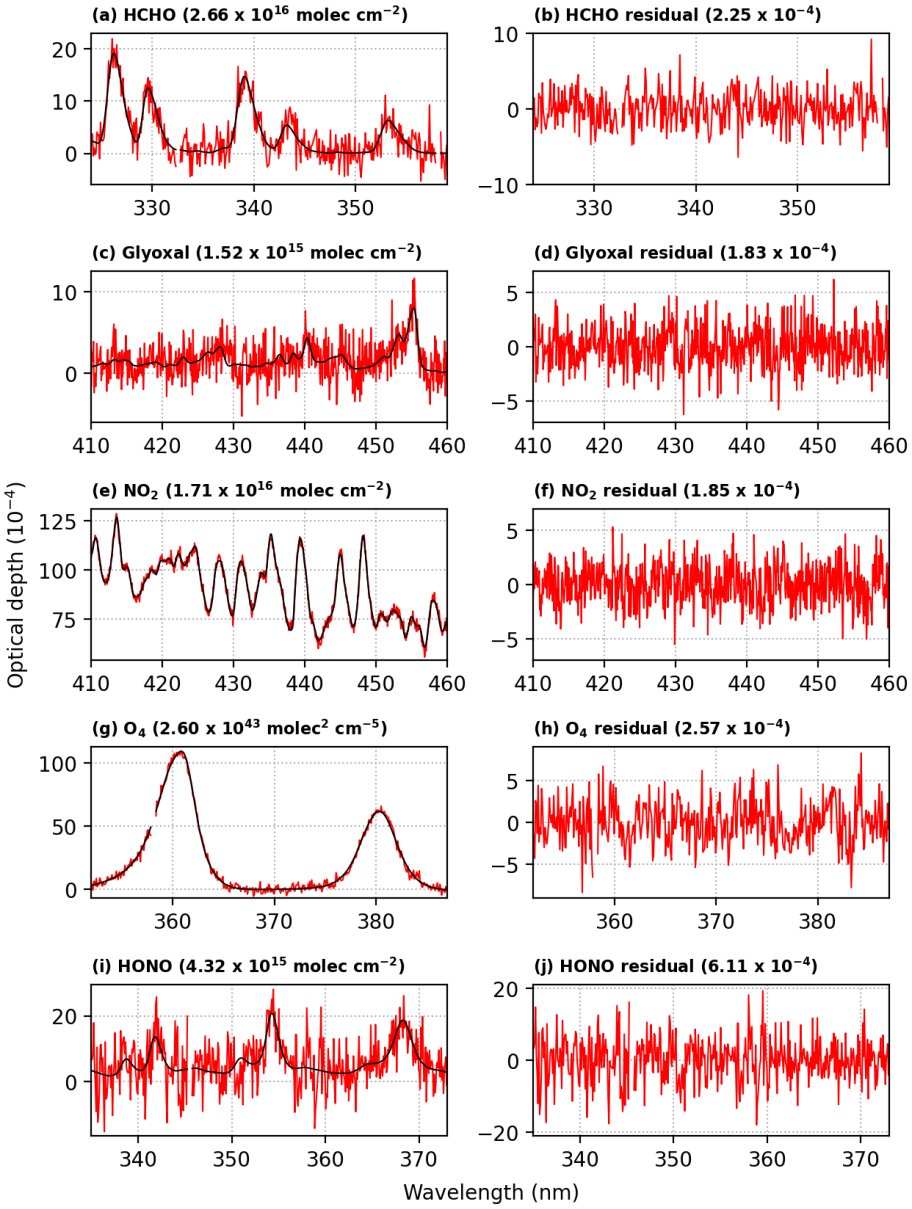


Figure 9.1: Fit examples (left column) and corresponding fit residuals (right column) for HCHO, glyoxal, NO_2, O_4 and HONO fits using the standard fit settings summarised in Tables 9.1 to 9.5. In the left column, black lines indicate the fitted cross-sections, while red lines show the fitted cross-sections plus the residuals. The numbers above the individual panels indicate the resulting dSCDs (left column) and fit RMS-values (right column). The examples are shown for spectra recorded at 3° elevation by the Bayfordbury instrument for HCHO (20 July 2019, 14:29 LT, SZA: 41.6°), glyoxal (24 July 2019, 11:15 LT, SZA: 33.5°), NO_2 (same as HCHO) as well as O_4 (16 July 2019, 11:05 LT, SZA: 32.6°) and the Mohali instrument for HONO (5 November 2020, 08:15 LT, SZA: 72.4°).

9.2 MAPA RETRIEVAL

The output of the spectral analyses are differential slant column densities at multiple elevation angles. MAPA version 1.01 was then used to derive vertical profiles of aerosol extinction and trace gas concentrations, including total columns, from the trace gas dSCDs. The basic concepts of the MAPA algorithm are outlined in Sect. 3.3, while here the used standard MAPA settings are summarised. In addition, the treatment of the retrieval results is described.

9.2.1 *Retrieval settings*

Default settings as described in Beirle et al. (2019) were used for the MAPA algorithm itself. The calculation of the a-priori O_4 VCDs needed for the aerosol retrieval was done by MAPA during the preprocessing of the input data utilising the provided ERA5 meteorological profiles as described in Beirle et al. (2022).

Further, LUTs interpolated to the effective wavelengths of the fit intervals of the respective trace gases were used for the species retrieved in the UV, namely aerosol, HCHO and HONO. Thereby, the effective wavelength of a given absorber is defined as the area (absorption) weighted mean wavelength within its fitting range. This approach is well justified since the (broad-band) wavelength dependence of the Rayleigh scattering probability is strongest at short wavelengths as it is proportional to λ^{-4}. As a consequence, the AMF and thus the sensitivity varies within a given fitting interval especially in the UV spectral range (compare Sect. 3.2.3). Therefore, the standard approach of using the AMFs at the wavelength of the main absorption feature of an absorber (e.g. 360 nm for aerosol/O_4) within its fitting interval is possibly not the perfect choice since the effective absorption probably occurs at a different wavelength having different AMFs. This is also supported by brief comparisons of both approaches based on one year of Bayfordbury data (2019) which showed that slightly better results in terms of more valid MAPA retrievals (compare Sect. 3.3) and better agreement with AERONET data (for aerosol) can be achieved when AMFs at the effective wavelengths are used for the profile inversions.

These brief studies on the MAPA settings also showed that more valid MAPA retrievals for aerosol and subsequently for HCHO are obtained when the aerosol results based on the O_4 dSCDs retrieved in the 338 to 370 nm fitting interval are considered. However, stronger absorption features of O_4 exist in the 352 to 387 nm spectral range. Consequently, the aerosol results based on O_4 dSCDs retrieved in the latter fitting range are probably more robust and more representative for the aerosol properties at longer wavelengths (visible spectral range). A comparison of the results of the spectral retrievals in both spectral ranges themselves showed highly consistent results. Therefore, two MAPA retrievals were performed for aerosol using the O_4 results from both fitting intervals (compare Table 9.4). In accordance with the previous paragraph, LUTs at 357 nm and 367 nm were used for the retrievals based on the O_4 dSCDs from the 338 to 370 nm and 352 to 387 nm fitting interval, respectively. The corresponding results were used for the trace gas retrievals for absorbers in the UV (HONO and HCHO) and visible

Table 9.6: Overview of the wavelengths of the LUTs used for the profile retrievals and aerosol scenarios used for the MAPA retrievals of the different trace gases.

SPECIES	PARAMETER	
	Aerosol	Wavelength
Aerosol 357 nm	–	357 nm
Aerosol 367 nm	–	367 nm
HCHO	357 nm	336 nm
Glyoxal	367 nm	430 nm
NO_2	367 nm	430 nm
H_2O	367 nm	430 nm
HONO	357 nm	357 nm

(NO_2, H_2O and glyoxal), respectively. Further, the aerosol results at 367 nm are exploited for the studies on the aerosol quantities themselves. Table 9.6 provides an overview of the wavelengths of the LUTs which were used for the different retrievals and the utilised aerosol scenarios for the different trace gases.

To scale the aerosol results to the wavelength used for the trace gas retrieval, the so-called Ångström exponent describing the wavelength dependence of the AOD is used as described in Sect. 5.2 (Ångström, 1929). In theory, this exponent depends on the specific optical properties of the aerosol and thus varies between different atmospheric scenes. Hence, MAPA offers the possibility to provide Ångström exponents, e.g. obtained from AERONET measurements, or to use the default value of 1. The studies on the MAPA settings using Ångström exponents inferred from AERONET measurements (compare Sect. 5.2) and the default value of 1 showed that the results could not be significantly improved (i.e. not more valid data). Also the resulting trace gas VCDs were not systematically changed. Therefore, the default Ångström exponent of 1 was applied for all MAPA retrievals in this work.

Several publications on MAX-DOAS measurements report on the need to scale the measured O_4 dSCDs to achieve agreement between measurements and the forward model simulations of the profile retrieval (for details refer to Wagner et al., 2019 and Beirle et al., 2019). Therefore, MAPA offers the option to define a fixed O_4 scaling factor or to determine the best matching, variable, scaling factor during the retrieval process itself (compare Sect. 3.3 and Beirle et al., 2019). Also the brief studies on the MAPA retrieval settings performed here showed the need of an O_4 scaling factor (SF) to notably improve the agreement of the retrieved AODs with the ones measured by AERONET and to markedly increase the amount of valid retrievals. Thereby, best and consistent results were obtained with a fixed SF of 0.8 and the variable SF, while a scaling by 0.9 was not sufficient. Hence, the MAPA retrievals were performed applying two different scaling factors, namely a variable scaling factor and a fixed one of 0.8. For the atmospheric investigations presented in this work, mainly the results for variable SFs are shown in the main

part. For details, refer to Sect. 9.3 which also provides an overview of the variable O_4 scaling factors obtained in this work.

Additionally, for all instruments the need of an elevation bias was found, details are discussed in the next section (Sect. 9.2.2). Lastly, the default MAPA output grid was used, which has a resolution of 20 m up to a height of 200 m and 200 m above this altitude up to a height of 4 km. Above that height, a few more layers exist with an uneven spacing.

9.2.2 *Elevation biases in MAPA retrievals*

For each MAX-DOAS instrument, the calibration of the elevation angles was optimised during the installation using different combinations of approaches presented in Donner et al. (2020). As suggested in that publication, daily horizon scans were performed at Bayfordbury and Mohali, while daily so-called obstacle scans were performed at ATTO, to monitor the temporal evolution of the elevation angle calibrations (unfortunately, no such monitoring is available for the Mainz measurements). For most stations, regular patterns and for some stations additionally abrupt changes of the elevation calibrations could be identified in that way. As MAPA offers the possibility to apply an elevation angle offset to correct for inaccuracies in the reported elevation angles, the influence of elevation biases on the MAPA retrievals was investigated. Further, a possible new tool to infer inaccuracies in the elevation calibrations from the MAPA results themselves is introduced in this section.

Since the elevation calibration at Bayfordbury was well characterised during the installation phase of the instrument in November 2018, this station is used for the following investigations. An expected horizon elevation of 0.52° was determined for this instrument. However, varying horizon elevations showing both smooth and rather abrupt changes were retrieved from the daily horizon scans (compare Fig. B.6 in the appendix). Overall, too high horizon elevations were retrieved, with values reaching up to ca. 2.25°, suggesting that a correction of the reported elevation angles by applying an elevation bias in the MAPA retrieval is necessary.

In MAPA the reported elevation angles α_{reported} are corrected via

$$\alpha_{\text{corr}} = \alpha_{\text{reported}} - \alpha_{\text{bias}}, \tag{9.1}$$

where α_{corr} and α_{bias} are the corrected elevations and the applied elevation bias, respectively. Note that in this definition too high reported angles require a positive bias. Using the possibility to correct the elevation angles in the MAPA retrieval and the well monitored changes of the elevation calibration at Bayfordbury, several MAPA runs were performed applying different elevation biases, including a scenario were the bias was intentionally applied in the wrong direction. The purpose of the latter was to have a proof of concept run.

These studies were only performed for inversions of aerosol and HCHO profiles for limited time periods, in which the horizon scans yielded quasi constant horizon elevations. Without going into details, which would go beyond the scope of this work, no clear results were found in these investigations with respect to the influence of an elevation bias on the number of valid MAPA retrievals

as well as to the agreement of the obtained AODs with AERONET data. As a consequence, it was decided to run the MAPA analysis for two scenarios. For the first run, the assumption was that the calibration of the elevation angles remains stable, thus, no bias was applied. The second run was performed assuming that the elevation calibration is changing as suggested by the horizon scans and a bias of 1° was applied which was expected to best represent the varying elevation bias on average (compare Fig. B.6 in the appendix). Both runs were carried out for a variable O_4 scaling factor and a fixed one of 0.8.

The influence of the applied elevation bias on the retrieved AODs is presented in Fig. 9.2. Panels (a) and (b) show the differences between the AODs retrieved with and without an elevation bias applying variable and fixed O_4 SFs, respectively. In case of a variable SF, the applied elevation bias leads to lower retrieved AODs, while larger AODs are obtained for a fixed SF. Compared to the respective mean AODs, the observed mean differences correspond to around −10 % and 20 %. Further, larger scatter can be observed for the fixed SF compared to a variable SF, while less retrievals are indicated as valid by MAPA for both cases if a bias is applied.

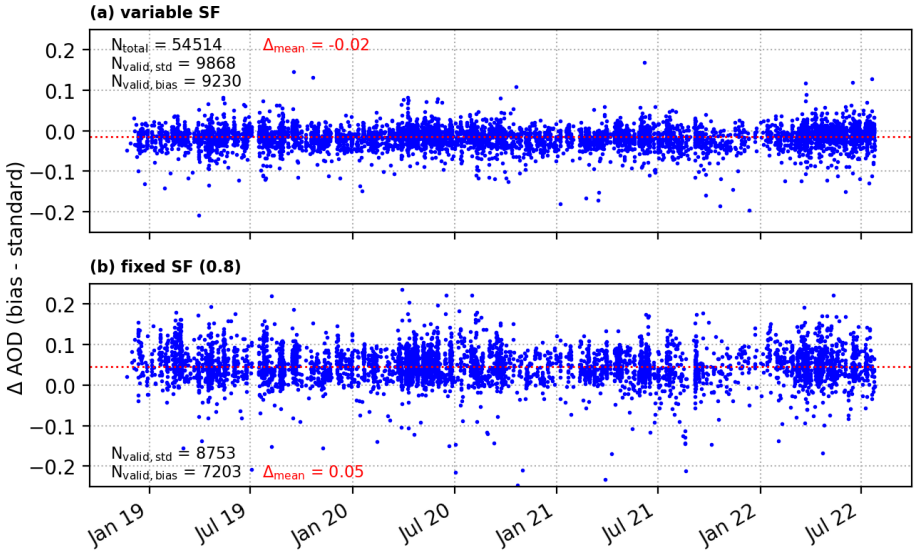


Figure 9.2: Differences between the AODs retrieved with and without an elevation bias applying a variable (a) and a fixed O_4 SF (b). Mean differences are shown in red.

For the run applying a variable scaling factor, also the effect on the resulting SFs can be investigated. Histograms of the O_4 SFs obtained without and with applying an elevation bias in the MAPA analysis are shown in Fig. 9.3. Applying an elevation bias leads to a 0.06 larger mean SF and a more symmetrical frequency distribution. The higher number of valid retrievals for the standard scenario in the case of a fixed SF of 0.8 can be explained by the obtained SFs since for the standard scenario the variable SFs appear to be closer to 0.8. For the scenario with a variable scaling factor, this observation is not fully understood.

In case of a variable O_4 scaling factor, parts of possibly existing inaccuracies of the elevation angle calibration seem to be compensated since smaller differences

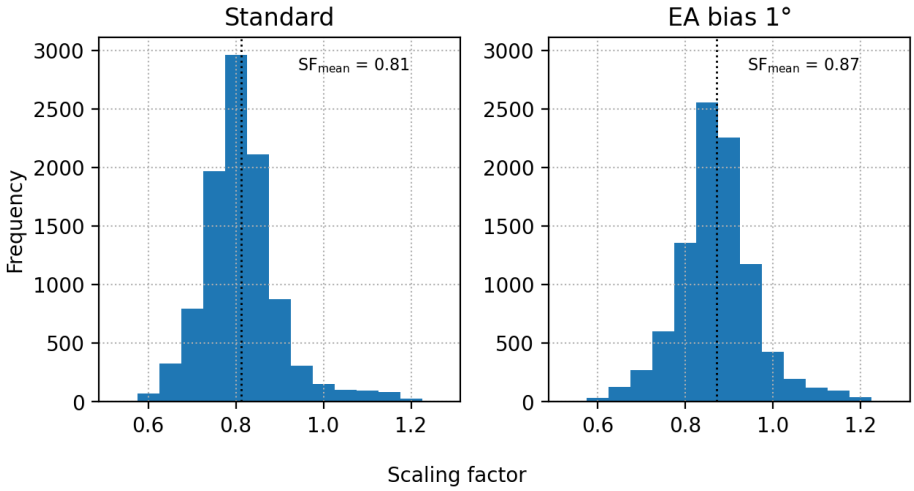


Figure 9.3: Histograms of the O_4 SFs obtained without (left) and with (right) applying an elevation bias of 1° in the MAPA analysis. Additionally, the respective mean values are indicated.

between the scenarios without and with elevation bias are found. Further, the clear influence of the elevation bias on the resulting variable SFs supports this hypothesis. Moreover, it should be mentioned that the described effects of the applied biases on the retrieved AODs and variable SFs are consistent with MAPA tests performed by Beirle (2020). These tests yielded a decrease of the SFs of 0.05 per degree downward tilt which is in good agreement with the value of 0.06 found here. However, the effects on the AODs were found to have the same magnitude for fixed and variable scaling factors. Nevertheless, at least the sign of the differences was the same as found in this work.

Lastly, the retrieved AODs at 367 nm and water vapour VCDs are compared to AERONET measurements. The corresponding correlation plots for the standard case (left) and the MAPA run with elevation bias (right) are depicted in Fig. 9.4, where results of simple linear regressions are additionally shown in red.

A general underestimation of the AODs by MAPA as compared to AERONET becomes evident, whereby low AODs are generally better reproduced especially for the standard scenario. An excellent agreement is found for water vapour. For aerosol, both scenarios yield comparable correlations, however, a minor worsening can be observed when an elevation bias is applied. For water vapour, the already very good agreement is even marginally improved when an elevation bias is applied. The results for the fixed scaling factor show overall similar patterns, with slightly worse agreement for the AODs. It should be mentioned that also an additional run without applying any scaling of the O_4 VCDs was performed, however, only little data is valid in this case and the agreement to AERONET is bad.

For trace gases, Beirle (2020) found larger VCDs in the case of a downward tilt of the instrument as compared to a correct elevation calibration for both variable and fixed SFs. Here, it should be stressed that the scenario with elevation bias in this work effectively corresponds to a correction of the elevations. Consistent to

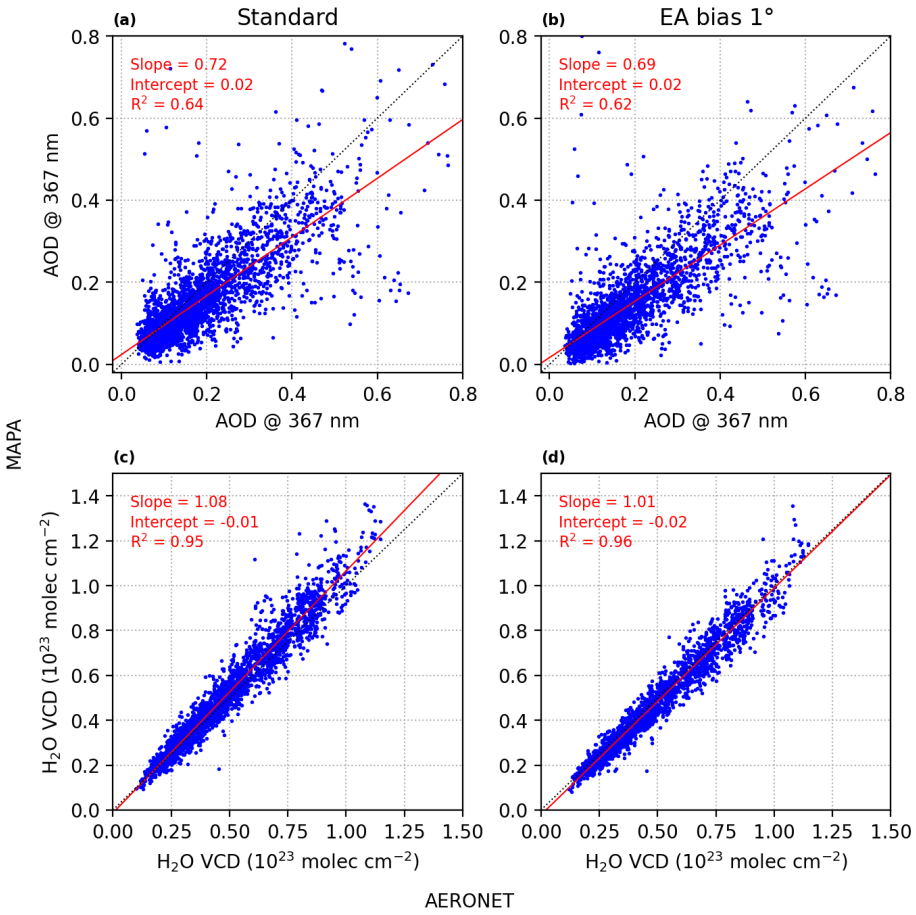


Figure 9.4: Correlations between MAPA and AERONET AODs at 367 nm without and with applying an elevation bias for the case of a variable SF are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

this, also in this work the MAPA runs where an elevation bias is applied yield slightly lower trace gas VCDs in the order of a few percent. Like for aerosol, less valid MAPA retrievals are obtained when an elevation bias is applied. An interesting aspect here is that, amongst others, also some of the clear outliers are flagged as *warning* when a bias is applied.

The studies by Beirle (2020) also suggest that the profile shape is strongly affected by a wrong elevation calibration. Specifically, downward tilts are associated with elevated trace gas and aerosol layers, where a large fraction of the total column is concentrated in a confined layer which is not located at the ground. While such profile shapes might occur for some species, especially for trace gases emitted at the ground, e.g. NO_2, such profile shapes are not expected. However, such elevated profiles occur very frequently for all species, when no elevation bias is applied. Although such a detailed study like presented here was only conducted for the measurements at Bayfordbury, it is expected that the main findings are also valid for the other stations. Interestingly, the frequency of elevated profiles is clearly reduced, when an elevation bias is applied. Although elevated

layers still occur, their height can be reduced and their vertical extent increased resulting in fewer unphysical confined layers. Surprisingly, this effect emerges very prominently for the vertical profiles of the water vapour concentrations. The diurnal evolution of the water vapour concentration profiles on 23 April 2020 is presented in Fig. 9.5, together with the corresponding VCDs and near-surface concentrations for both scenarios, without (a) and with (b) applying an elevation bias.

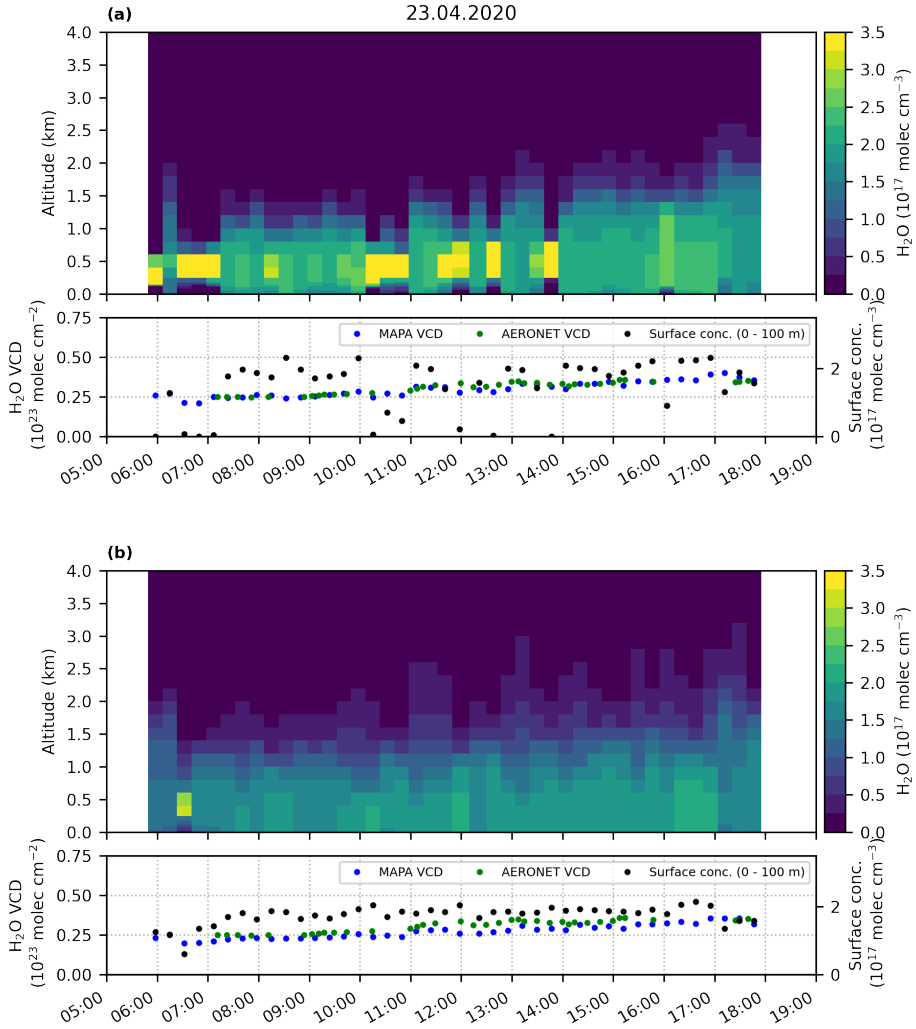


Figure 9.5: Diurnal evolution of water vapour concentration profiles on 23 April 2020. Parts (a) and (b) show the results for the scenarios without and with elevation bias, respectively. Colour-coded are the concentrations in the different grid boxes. The lower panels indicate the corresponding MAPA and AERONET H_2O VCDs as well as the MAPA near-surface concentrations. Results are shown for a variable O_4 scaling factor.

Obviously many profiles showing elevated water vapour layers with concentrations abruptly approaching zero towards the surface are obtained. Applying an elevation bias basically removes all of them yielding profiles that reach overall higher and do not show steep concentration gradients towards the ground. As

mentioned above, the VCDs are slightly decreased when an elevation bias is applied. For the presented day, the number of valid profiles is identical for both scenarios and a slight worsening of the agreement to the AERONET VCDs is observed. However, the correlation plots in panels (c) and (d) of Fig. 9.4 suggest that when the complete data set is considered the agreement between the MAPA results and the AERONET data is marginally improved if an elevation bias is applied. In addition, the improvement achieved for the profile shape can be considered to be more important than the slight worsening of the VCDs on individual days. Figure 9.5 depicts results for a variable SF, however, similar patterns are observed for the fixed SF of 0.8 (not shown).

The observed strong effect of an elevation bias on the profile shape offers the possibility to use the retrieved water vapour profiles to assess whether an elevation bias, i.e. a correction, is needed or not. The basic assumption here is that water vapour usually does not show elevated layers. As an example, panel (a) of Fig. 9.6 depicts monthly mean water vapour concentration profiles for April and July 2020 extracted from the ERA5 data set (compare Sect. 5.1). A rapid decrease of the H_2O concentration with increasing altitude is clearly visible for both months and the main difference are the absolute values of the concentrations which are higher in summer. For both months, no elevated layers occur.

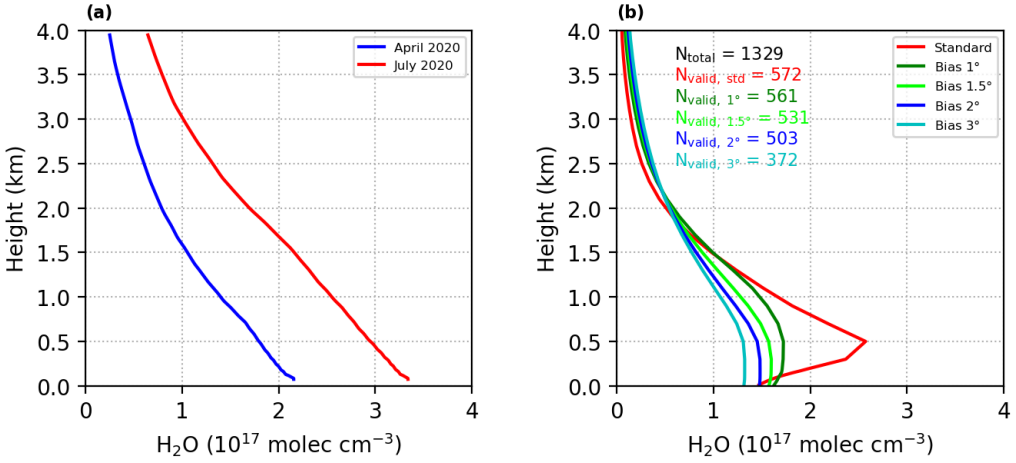


Figure 9.6: Panel (a) shows mean water vapour concentration profiles for April and July 2020 for Bayfordbury extracted from the ERA5 data set. Panel (b) presents mean water vapour concentration profiles for April 2020 retrieved with MAPA applying a variable SF and different elevation biases. Additionally, the number of valid MAPA retrievals is indicated.

Running the MAPA retrieval for water vapour and checking the shapes of the resulting profiles, thus, allows to assess whether an elevation angle bias is needed or not. If necessary, runs with different biases can be conducted and the bias yielding profiles that match the expected shape with no elevated layers best can be chosen. Additionally to the already mentioned MAPA retrieval applying an elevation bias of 1°, several MAPA runs using biases of 1.5°, 2° and 3° were performed for April 2020. The obtained mean profiles are presented in panel (b)

of Fig. 9.6. The largest improvement is observed when a bias of 1° is used, while the additionally tested biases do not lead to significant further improvements. On the contrary, the number of valid profiles rapidly decreases the higher the applied bias gets. Further, it should be mentioned that the smallest changes in the obtained profile shapes between the two applied SFs are found for a bias of 1° (not shown). As a consequence, for Bayfordbury the elevation angles were corrected by 1° for the MAPA retrieval.

As a final aspect, it should be discussed that the profile parametrisation used in MAPA cannot reproduce purely exponential profiles (compare Sect. 3.3). This becomes also evident in Fig. 9.6b, where despite the application of an elevation bias still no perfect agreement with the expected profile shape shown in panel (a) can be observed. As shown for formaldehyde (a) and NO_2 (b) in Fig. 9.7, for other trace gases a quite steep decrease of the concentrations towards the surface remains visible even after correcting the elevation angles in the MAPA retrieval. This remaining systematic structure is characteristic for profiles obtained with MAPA v1.01 and is caused by a combination of the parametrisation used in the algorithm and the calculation of the output profiles. Here, the mean of the Monte Carlo ensemble (compare Sect. 3.3) was used. Single profiles with shape parameters larger than 1, corresponding to elevated box profiles, which are included in the ensemble lead to the observed steep concentration gradient towards the ground for the "mean output" profiles. This effect also has to be considered when calculating concentrations that shall be representative for near-surface values as outlined in the next section (Sect. 9.2.3).

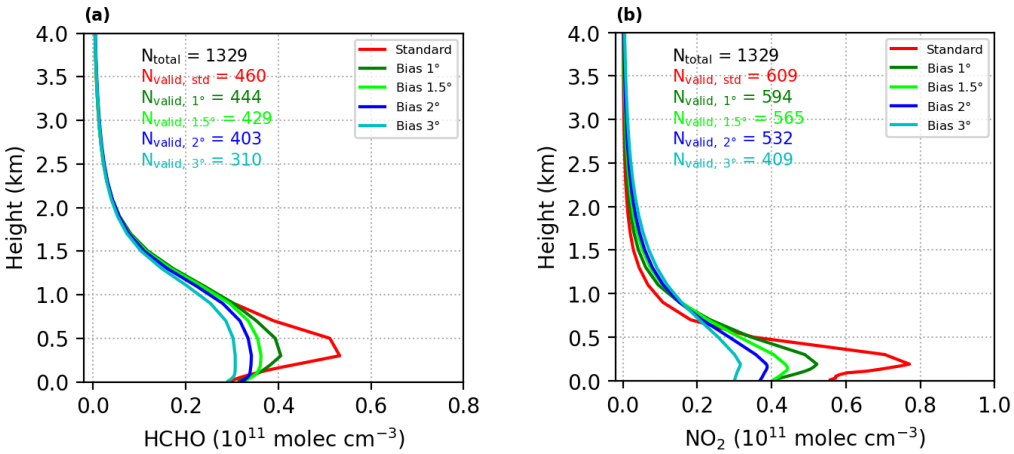


Figure 9.7: Mean formaldehyde (a) and NO_2 (b) concentration profiles for April 2020 retrieved with MAPA applying different elevation biases and a variable SF. Additionally, the number of valid MAPA retrievals is indicated.

In combination with daily horizon or obstacle scans (if available), the best matching elevation bias was determined for each of the five instruments using the procedure outlined above. For Mohali and Mainz, the effects of the applied elevation biases on the water vapour profile shapes are not so clear and complex behaviours are observed. Most likely, this can be explained by high aerosol loads

concentrated very close to the ground which frequently occur at Mohali and Mainz.

Finally, the resulting biases that were used for the MAPA analyses for the different instruments are summarised in Table 9.7, together with the lowest elevation angles that were included in these MAPA analyses. The latter is needed since for all stations the lowermost (positive) elevations had to be excluded from the retrievals since they were blocked by obstacles. Note that the uncorrected elevation angles are given in the following table. Negative elevation angles, i.e. measurements for which the instrument points below the horizon, included in the elevation sequences of the different instruments (compare Table A.1 in the appendix) were generally excluded from the MAPA analyses for two reasons. Firstly, the LUTs needed for the MAPA retrieval do not include negative elevations. Secondly, such measurements typically do not yield high quality results. For most stations, all other available elevations starting at the one listed in Table 9.7 were used for the MAPA retrievals, whereas for the two ATTO instruments also elevations of 45° and above were excluded. A major reason for this decision is consistency since such elevations were not included in the measurement sequences at the other stations (compare Table A.1 in the appendix). Moreover, tests showed that the inclusion of measurements at these high elevations does not improve the MAPA results. A last aspect contributing here is that the LUTs do not include elevations above 45°.

Table 9.7: Overview of the applied elevation biases and lowest elevation angles used in the MAPA analyses for the different instruments.

STATION	ELEVATION BIAS	LOWEST ELEVATION
Bayfordbury	1°	3°
Mainz	1°	2°
Mohali	1°	2°
ATTO 80 m	1°	2°
ATTO 298 m	0.5°	2°

9.2.3 *Comparison of near-surface concentrations and total columns*

Besides vertical profiles of trace gas concentrations and aerosol extinction, the MAPA output provides integrated columns in terms of trace gas vertical column densities (VCDs) and aerosol optical depths (AODs). For the atmospheric investigations carried out in the following chapters (Chapt. 10 to 13), two quantities are used. While the total amount will be characterised by VCDs and AODs, near-surface values will be described by mean concentrations and extinctions over the few lowermost MAPA layers[1]. In this section, both quantities are compared to each other to investigate how independent from each other they are. This gives an indication about the information content of the retrieved MAX-DOAS profiles.

1 Note that MAPA yields vertical profiles above instrument altitude.

Several aspects have to be taken into account when deciding up to which level the concentrations and extinctions should be averaged. On one side, the averaging height should not be too high, since otherwise processes like mixing or chemical processing might not be resolved by the comparison between both quantities. Further, the concentrations of some species already start to decline above a few hundred metres. On the other side, the profiles obtained with MAPA show systematic artefacts in the lowest layers of the profiles as outlined in the previous section and the surface values might be underestimated when the averaging height becomes to low. As a compromise, it was decided to average the concentrations over the lowest 400 m of the total column in order to obtain near-surface concentrations. This height range corresponds to the lowest 11 layers of the retrieved MAPA profiles. Since the vertical MAPA output grid changes its spacing at 200 m, a weighting with the individual layer thicknesses had to be applied. It should be noted that for all stations, except for ATTO, the terms "near-surface concentration" and "surface concentration" are used synonymously in the following.

To investigate the information content of the retrieved MAX-DOAS profiles, Fig. 9.8 shows correlations of HCHO near-surface concentrations and their corresponding VCDs measured at Bayfordbury. Only the results of valid MAPA retrievals applying a variable scaling factor are presented. The four panels show the results of individual measurements as well as hourly, daily and monthly means. Altogether, good correlation between the two quantities can be observed. Nevertheless, especially the individual measurements (single) show that the correlation is not perfect and that there is a certain diversification in the scatter plot. For instance, the increase of the near-surface concentration by a certain amount does not always lead to the exactly same increase of the corresponding VCD. This indicates that the concentration profiles of HCHO contain more information than just the total amount expressed by the VCDs. Thus, the investigation of both HCHO near-surface concentrations and VCDs as well as their interplay might reveal interesting additional insights. With increasing averaging period, the R^2-values increase yielding almost 1 for monthly means. Here, the additional information content, e.g. systematic changes of the profiles during the day or the meteorological variability, gets lost and mainly the seasonal patterns are left.

For the other species, the patterns are similar, although for absorbers located relatively close to the ground, e.g. NO_2, a clearer splitting of the point cloud can be observed (compare Fig. B.8 in the appendix). This shows that for these species the additional information content can be even larger. The correlations of the near-surface concentrations/aerosol extinctions and the corresponding VCDs/AODs of glyoxal, NO_2 and aerosol are shown in Fig. B.7 to B.9 in the appendix. No significant differences between the results of the two applied scaling factors are found.

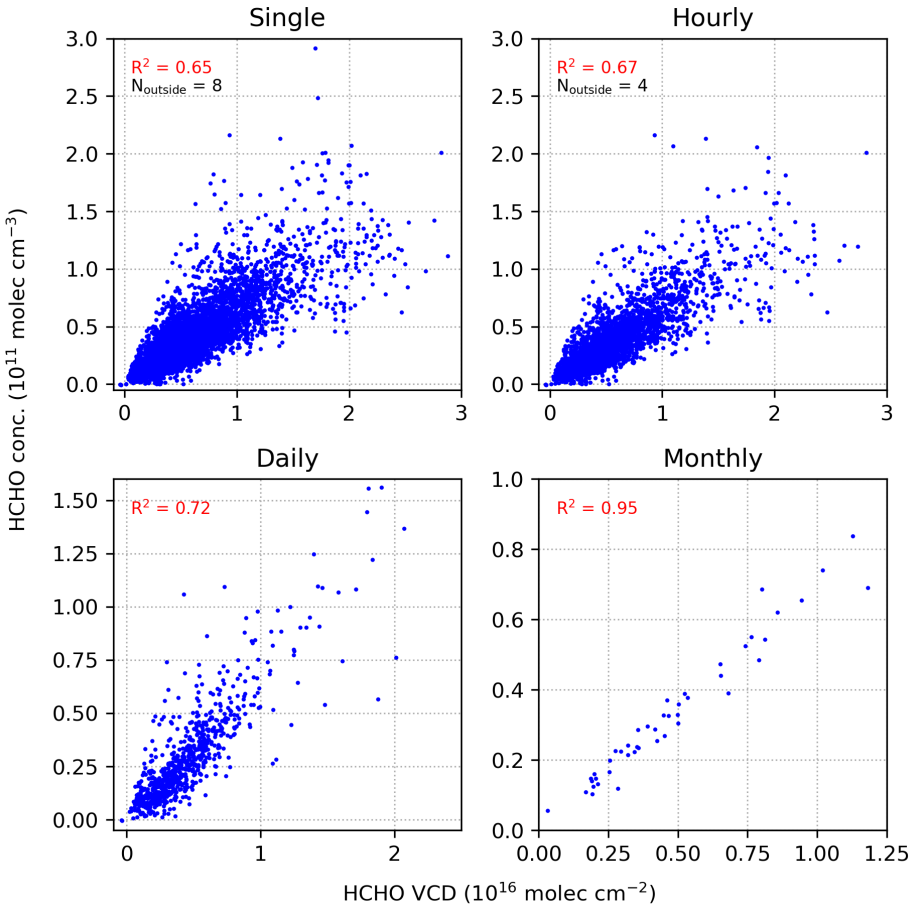


Figure 9.8: Correlation of HCHO near-surface concentrations and corresponding VCDs measured at Bayfordbury. Only valid results for a variable O_4 SF are shown. The four panels present individual measurements as well as hourly, daily and monthly means.

9.3 DATA SELECTION

Before the MAPA results are visualised and discussed, some filters were applied to keep only high quality results. This section briefly outlines and justifies the applied filter criteria.

As described in Sect. 3.3, in the here-used standard configuration, MAPA yields the best matching parameter combination as well as an ensemble of parameter sets containing up to 100 members. Several statistical quantities can be calculated from this ensemble which are, amongst others, the mean, the weighted mean and the standard deviation. As already seen in Sect. 9.2.2, the usage of mean profiles leads to strong concentration gradients towards the ground in case profiles with shape parameters larger than 1 are included in the ensemble. Despite this effect and the slightly worse agreement to the measured dSCDs, the usage of mean profiles might be more robust since consecutive profile shapes can be significantly different if only the best matching parameter set is considered. Comparisons showed that the VCDs and AODs obtained using different statistical quantities

agree very well. Thus, it was decided to use the mean of the Monte Carlo ensemble for all investigations performed in this work.

Due to decreasing measured intensities, the fit RMS of the spectral DOAS analysis typically increases with increasing SZA, even if sequential reference spectra are used. Further, the measurement routine which adjusts the integration times used for the recording of the spectra gets automatically slower with decreasing intensities. As a consequence of the latter, the completion of one full measurement sequence takes longer at large SZAs and the atmospheric state possibly changes more within one sequence. Both effects together lead to a poorer quality of the finally obtained MAPA results, especially for species retrieved in the UV, like for example HCHO. Partly this is accounted for by the MAPA flagging (see below), however, experience shows that still some problematic results remain. Therefore, all valid results obtained at SZAs above 80° are additionally excluded from the studies presented in this work. Since HONO is photolysed quickly under the influence of solar radiation and thus substantial amounts typically are expected to occur only during morning hours, measurements up to an SZA of 90° are considered for this species.

MAPA provides a detailed flagging of the retrieval results and warnings as well as errors are raised based on the performance of the profile retrieval. A more in-depth description of the MAPA flagging was already given in Sect. 3.3, while details and the applied thresholds are described comprehensively in Beirle et al. (2019). The MAPA flags are generally rather strict with relatively conservative thresholds especially for warnings. During the analysis of the measurement data in this work, it became evident that the MAPA flags are particularly sensitive to cloudy scenes which are associated with complex radiative transfer conditions that are problematic for the MAPA algorithm. This is even worse when additionally rapid changes in cloud coverage occur which is very often the case at the tropical site ATTO (compare Sect. 4.4). Under these cloudy conditions, often the measured dSCDs within a sequence do not follow a smooth pattern but are more noisy leading to a large fraction of flagged retrievals. Also during the comparison with satellite data, it became clear that in addition to the satellite quality criteria the MAPA flags indicate measurements which are associated with quite large cloud fractions reported within the satellite data set. Statistics of the raised MAPA warnings showed that the most frequent warnings are the RMS and consistency flags.

While retrievals flagged as errors should be rejected in general, investigations using measurement data from different stations showed that profiles marked with warnings can in principle contain useful information. In particular, VCDs and AODs often seem to have reasonable values, however, these are often associated with unreasonable vertical concentration profiles. Thus, the corresponding near-surface concentrations yield problematic values. Further, the scatter of the data is largely increased if results with warnings are included in the investigations. As a consequence, only valid profile retrievals without any flag are considered for the atmospheric investigations presented in the following chapters. This approach leads to a large fraction of excluded measurements, but given the long time series at most stations this is not a major problem since there are still enough

data available to obtain good statistics. However, it should be mentioned that the MAPA flags potentially exclude specific meteorological conditions such as cloudy or rainy scenes. Hence, the results potentially have a clear-sky bias and are therefore possibly not fully representative for all meteorological conditions.

As outlined above, the MAPA retrievals in this work were performed applying two different O_4 scaling factors, namely a variable SF and a fixed one of 0.8, since significant improvement of the agreement between retrieved AODs and the ones measured by AERONET could be achieved in that way. Table 9.8 summarises the fractions of valid aerosol, HCHO and NO_2 retrievals obtained for the five instruments applying either SF. Note that in addition to the MAPA flags, the SZA filter was already applied here, i.e. only valid MAPA results at SZAs below 80° were taken into account to calculate the listed fractions. However, as the SZA filter is applied to all retrievals regardless of the applied SF, the difference in the fractions of valid results can be attributed solely to the number of valid MAPA results obtained with either SF.

Table 9.8: Overview of the fractions of valid aerosol, HCHO and NO_2 retrievals for the five instruments obtained with fixed (0.8) and variable (var) O_4 SFs. Additionally, the total number of available profiles is given for each instrument.

STATION	VALID DATA						
	Total	Aerosol		HCHO		NO_2	
		var	0.8	var	0.8	var	0.8
Bayfordbury	54 514	16.9 %	13.2 %	10.0 %	7.6 %	16.3 %	12.8 %
Mainz	148 954	20.4 %	15.4 %	5.2 %	4.0 %	19.0 %	14.5 %
Mohali	115 646	33.9 %	34.9 %	27.4 %	27.7 %	32.1 %	33.2 %
ATTO 80 m	40 083	13.5 %	10.3 %	10.8 %	7.5 %	3.9 %	2.9 %
ATTO 298 m	16 747	14.3 %	11.8 %	12.0 %	10.0 %	4.1 %	3.4 %

Two clear conclusions can be drawn. First, the fractions of valid retrievals are small and typically below 20 % with the exception of Mohali. Second, the variable SF obviously leads to significantly more valid retrievals for all species, except for Mohali, where marginally more valid retrievals are obtained with the fixed scaling factor. The reason for the latter is still not clear, however, the mean of the obtained variable SFs is largest at Mohali yielding a value of 0.90 (compare below). As described in Wagner et al. (2019, 2021), the reason why a scaling of the O_4 dSCDs in some profile retrievals is necessary at all is not yet understood and several publications on MAX-DOAS measurements even report that no scaling is needed (e.g. Spinei et al., 2015; Ortega et al., 2016). Wagner et al. (2019) summarise that the best matching O_4 SFs vary between 0.75 and 0.9 in many cases while the factors determining the actual values of these best matching SFs are not fully clear. Figure 9.9 shows (relative) histograms of the O_4 SFs determined during the MAPA retrievals using variable SFs for the five instruments considered in this work. Further, the corresponding mean values (vertical dashed lines), which are also summarised in Table 9.9 together with their standard deviations, are shown.

For these histograms, only valid retrievals based on the O_4 dSCDs retrieved in the 352 to 387 nm fitting interval[2] and data with SZAs below 80° were considered.

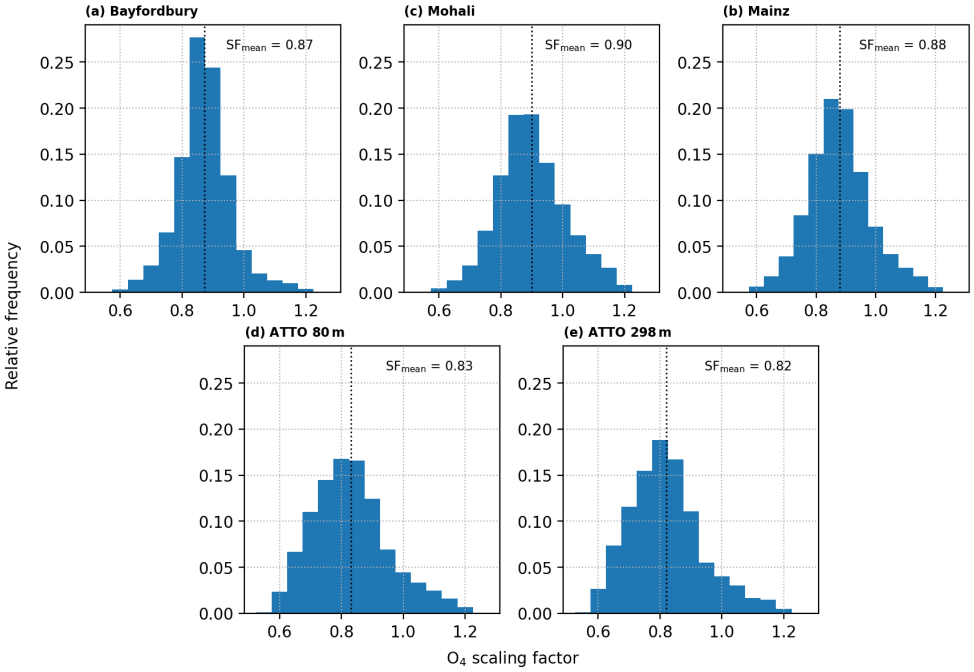


Figure 9.9: Histograms of variable O_4 SFs obtained for the five instruments using variable SFs in the MAPA retrievals. Only valid retrievals were considered. Additionally, the respective mean values are indicated by vertical dashed lines.

Table 9.9: Overview of the mean O_4 SFs obtained for each of the five instruments using variable SFs in the MAPA retrievals. Also the corresponding standard deviations are listed.

STATION	MEAN SF	STANDARD DEVIATION
Bayfordbury	0.87	0.09
Mohali	0.90	0.11
Mainz	0.88	0.11
ATTO 80 m	0.83	0.12
ATTO 298 m	0.82	0.12

Consistent with the mentioned recent literature values, the mean O_4 scaling factors obtained for the five instrument range from 0.82 to 0.90. While the distribution is rather narrow for the Bayfordbury measurements (standard deviation of 0.09), the distributions are broader for the two ATTO instruments (standard deviations of 0.12). The latter indicates larger variations of the best matching variable SFs which might be related to the frequent occurrence of rapidly changing cloud cover at ATTO strongly affecting the MAPA retrievals. The reasons

2 A comparison showed that similar histograms and mean SFs are obtained for the retrievals based on the O_4 dSCDs retrieved in the 338 to 370 nm fitting interval.

for the systematically different mean SFs for the different instruments remains unclear also in this work. Possible factors contributing here are different aerosol abundances and different aerosol properties. This is supported by the finding that at Mohali where high aerosol abundances occur frequently (compare Chapt. 12) the on average largest variable SFs are found, whereas for the two ATTO instruments consistently the smallest means are obtained. As outlined above, also the accuracies of the elevation calibrations probably play a role here. However, these aspects should be in the focus of further investigations regarding profile retrievals. Lastly, for some stations small systematic annual variations of the variable O_4 SFs were observed. Yet, no clear reason for this could be identified in this work based on comparisons of the annual patterns of the SFs with meteorological quantities (temperature, pressure and humidity).

To check which scaling factor yields better results, the retrieved AODs and water vapour VCDs were correlated with the corresponding AERONET measurements of these quantities (compare Sect. 5.2). The obtained correlation plots for Bayfordbury are shown in Fig. 9.10, where results for the fixed and variable SFs are shown in the left and right columns, respectively.

While at low AODs an overestimation is found for the fixed SF, an underestimation is visible for the variable scaling factor. The tendency towards lower AODs for the variable SF is found at all stations (compare upper panels in Fig. B.10 to B.13 in the appendix). Excellent agreement is found for the water vapour VCDs for both scaling factors. The performed linear regressions show that for both aerosol and water vapour similar results are obtained for both applied SFs. However, slightly better agreement is found in case of a variable SF for aerosol. It should be mentioned that similar investigations were also performed for the other stations yet not for all of them so clear conclusions like for Bayfordbury can be drawn. The corresponding correlation plots are presented in Fig. B.10 to B.13 in the appendix.

Lastly, it should be stressed that for the general results of the different species and most results of the atmospheric studies performed in this work the choice of the O_4 scaling factor is not critical. Given this and in light of the fact that significantly more valid retrievals are obtained for the variable scaling factor, which increases the relevance of possible conclusions, only results for the variable scaling factor are presented in the following. For the few cases where significant differences between both scenarios are observed, they will be carefully assessed, as for example slightly different profile shapes are sometimes retrieved (compare Sect. 10.5).

9.4 ERA5 VS. LOCAL METEOROLOGICAL DATA

At all four measurement sites, in principle local weather data collected by weather stations in the vicinity of the MAX-DOAS instruments exist. However, not all of these weather stations collected the meteorological data in a consistent way, e.g. the weather station at ATTO is located at the top of the tall tower at an altitude of 325 m. In addition, the weather data at some stations do not cover the complete measurement periods of the MAX-DOAS instruments. As consistency is a major

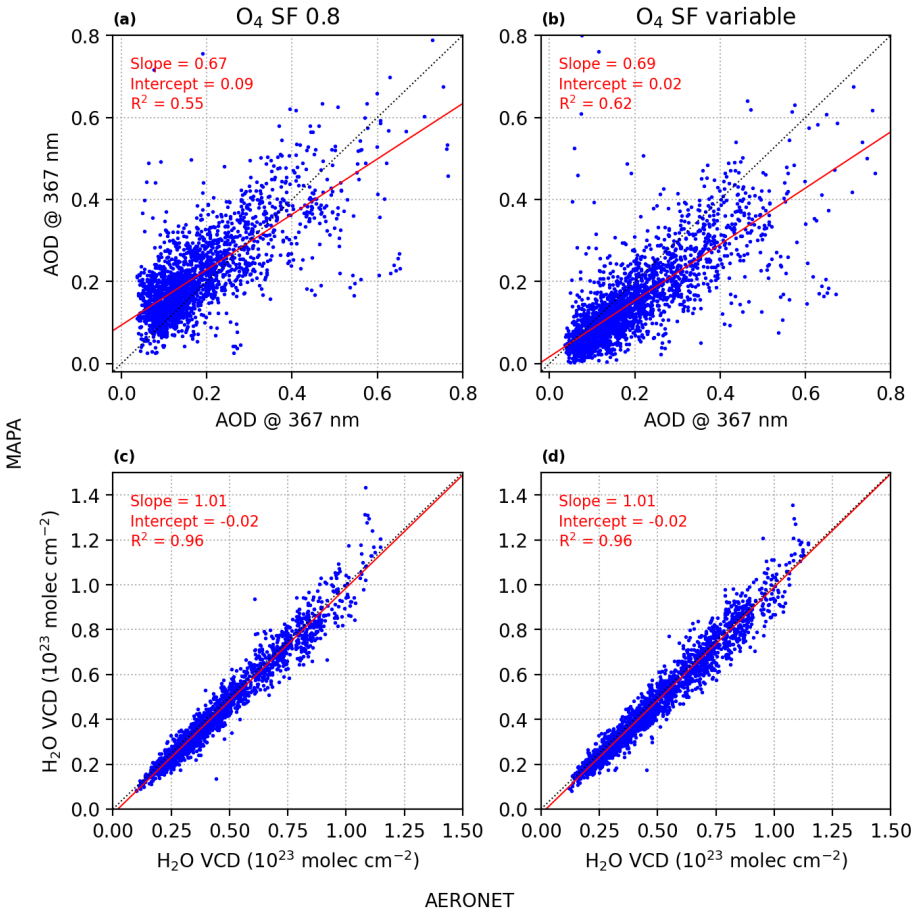


Figure 9.10: Correlations between MAPA and AERONET AODs at 367 nm applying fixed (0.8) and variable O_4 scaling factors for the Bayfordbury measurements are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

aspect of this work, it is desirable to use a consistent meteorological data set which covers the whole measurement period for each of the instruments. This can be achieved by using the ECMWF ERA5 reanalysis data set introduced in Sect. 5.1. This section compares the local temperature and wind data against the ERA5 data set. This is done in detail for one selected station, namely Bayfordbury.

Profiles of temperature and wind were already extracted from the ERA5 reanalysis data set for Bayfordbury in a preprocessing step of the MAPA analysis. Thus, for each measurement sequence temperature and wind data are available on an uneven time grid which are used in the following. Note that this approach leads to the limitation that values are only available during the day. For the comparison with local data, temperatures and wind vectors of the lowest model layer were used which represents values at a height of approximately 10 m. The temporal resolution of the local weather station is 1 min. To compare both data sets, they were brought on a one hour time grid by averaging them over one hour bins. The wind data of the local weather station have a detection limit of 0.44 m s^{-1} and

often wind speeds of 0 are measured. Therefore, only hours with at least 5 min of non-zero winds were considered in the comparison of the wind data, while all available data were used for the comparison of air temperatures.

The correlation between ERA5 and locally measured temperatures is shown in panel (a) of Fig. 9.11, where the results of a simple linear regression are additionally shown in red. In general, good agreement between both data sets can be observed yielding an R^2 of 0.93. However, the extreme values are not well captured by the ERA5 temperatures which can be explained by the temporal and spatial resolutions of the data set of six hours and $1° \times 1°$, respectively. Given these resolutions and the fact that a reanalysis model is involved, single extreme values of air temperatures close to the ground, which are largely influenced by the local circumstances, cannot be expected to be accurately represented by ERA5.

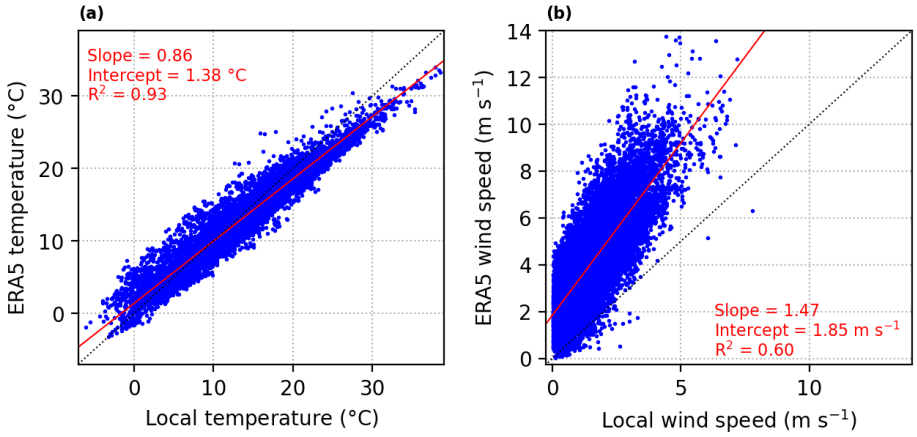


Figure 9.11: Correlations between ERA5 and locally measured temperatures (a) and wind speeds (b) at Bayfordbury. The results of simple linear regressions are displayed in red, while 1-to-1 lines are indicated by black dashed lines.

Panel (b) of Fig. 9.11 presents the correlation between ERA5 and locally measured wind speeds as well as the results of a simple linear regression. Despite reasonable correlation, significantly higher wind speeds occur for the ERA5 data set. At least two effects might explain this observation. First, the surroundings of the local weather station might lead to a shielding effect, as for example trees, the domes of astronomical telescopes and two buildings are within a few tens of metres. Second, the ground height in the model deviates by around 7 m from reality. However, the latter is probably a minor effect.

For the comparison of the wind directions of both data sets, the angles between the hourly wind vectors were calculated in a way that only values between $-180°$ and $180°$ are possible. Positive and negative angles correspond to clockwise and counter-clockwise deviations of the wind vectors, respectively. Figure 9.12a shows the frequency distribution of the angular differences in bins of 5°. A nearly Gaussian shaped distribution is obtained indicating that no large systematic differences between the wind directions of both data sets exist. The mean angular difference is calculated as 2.6° with a standard deviation of 30.2° yielding a 2σ interval of $-57.8°$ to 63°. The standard deviation shows that at least 68.2 % of the data have similar main wind directions, i.e. the winds originate from the same

±45°-sector around the reference wind direction (local wind data). The median of the unsigned differences is 13.6° (not shown) illustrating that 50 % of the data agree also in the secondary wind direction (±22.5°-sector around the reference wind direction).

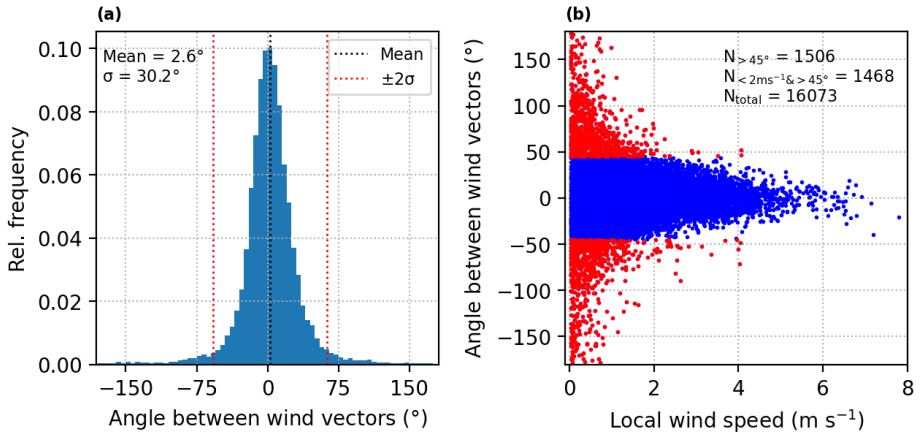


Figure 9.12: Panel (a) shows the frequency distribution of the angular differences between ERA5 and locally measured wind directions in bins of 5°. Additionally, the mean difference and the 2σ interval are indicated. Panel (b) presents the angular differences as function of locally measured wind speeds. Differences larger than 45° are marked in red.

The dependence of the obtained angular differences between ERA5 and locally measured wind directions on wind speed is investigated in Fig. 9.12b, where differences larger than 45° are shown in red. Additionally, the number of data points above this threshold is indicated as well as the number of data points that in addition have measured wind speeds lower than 2 m s^{-1}. Approximately 90 % of the data are found to have the same main wind directions, while nearly all differences larger than 45° occur at measured wind speeds below 2 m s^{-1}. At these low winds speeds the measurements are expected to show a larger scatter and to be more influenced by local conditions. Also the uncertainties in the ERA5 wind fields might be larger under such conditions. As a consequence, the observed differences of the wind directions appear to be understandable and uncritical.

Apart from the larger differences between ERA5 and measured wind speeds, both data sets agree reasonably well. Thus, for Bayfordbury in principle both data sets could be used since also the locally measured data set covers the whole MAX-DOAS measurement period.

A detailed comparison was only provided for Bayfordbury, but also for the other stations reasonable agreements between the ERA5 data set and the local weather data can be observed. While for temperatures in general good agreements are found for all stations, the wind data deviate more strongly at ATTO and Mohali. For ATTO, the agreements of both temperatures and wind data can be improved by utilising the model layer corresponding to ca. 325 m instead of the lowest model layer. For Mohali, however, also the choice of a different model layer cannot improve the agreement. The relatively large differences between model and measured wind vectors found for Mohali can most likely be explained by

the proximity of the Himalayan mountain range in connection with the relatively coarse spatial resolution of the model (compare Sect. 5.1).

Nevertheless, ERA5 offers the possibility to use a consistent data set that covers the measurement periods of all instruments without gaps which is not the case at all measurement stations for the local measurements. This is justified since overall good agreement between ERA5 and local measurements is obtained. Further, the ERA5 wind data are probably more representative for regional transport processes. Given all these aspects, it was decided to use the ERA5 meteorological data on a one hour time grid for the following investigations of the dependencies of the trace gas results on temperature and wind.

10

BAYFORDBURY, UK

In this chapter, the results of the measurements conducted at the Bayfordbury Observatory (Sect. 4.2), United Kingdom, are discussed, focussing on HCHO, glyoxal, NO_2 and aerosol results. First an overview of the measurements and the particularities of the analysis is provided in Sect. 10.1, followed by a discussion of annual and diurnal cycles of the species of interest in Sect. 10.2. Investigations on the dependencies of the measured trace gas and aerosol abundances on wind characteristics and air temperature are outlined in Sect. 10.3, while the relation amongst the different species themselves is analysed in Sect. 10.4. Finally, the seasonal behaviours of the vertical distributions of the main species are investigated in Sect. 10.5 and a short comparison of the MAX-DOAS and TROPOMI formaldehyde VCDs is performed in Sect. 10.6.

10.1 OVERVIEW AND PARTICULARITIES OF THE ANALYSIS

The MAX-DOAS measurements at Bayfordbury were started on 21 November 2018 and although the measurements are in principle still running (as of July 2023), only data up to 24 July 2022 were taken into account for this work. This time range includes lockdown measures due to the COVID-19 pandemic in 2020 and partly the extreme heat wave in Great Britain in summer 2022. The latter led to measured air temperatures as high as 38.9 °C (19 July 2022) at Bayfordbury which is the highest temperature measured during the entire measurement period.

Given an excellent support by local observatory staff, only minor data gaps caused by rare instrumental issues, e.g. spectrometer software crashes, occurred during this period. The biggest disturbance in Bayfordbury so far is the defective fan of the instruments cooling unit which broke in April 2020. As a consequence, the spectrometer temperature had to be increased from 15 to 28 °C with a transition period in May and June 2020 in which different temperatures were tested. This resulted in three periods for which different DC and offset spectra were needed for the correction of the measured spectra. Details of how the changing spectrometer temperature affects the optical properties of the instrument were elaborated in Sect. 8.2 and despite the rather small effects reported there, three different spectral calibrations and SLFs were used for the DOAS retrievals for Bayfordbury (compare Table 4.1). This procedure assures that, besides only marginally increased noise levels and fit RMS-values, no adverse effects on the retrieval process have been recognised for the affected time periods and the respective data can be used without limitations.

Besides three spectral gaps (332.4 – 332.7 nm, 345.8 – 346.1 nm and 357.9 – 358.3 nm) introduced to the DOAS fit due to hot pixels, for the Bayfordbury measurements no modifications of the standard fit settings listed in Sect. 9.1 had to be applied. Also the MAPA inversions were performed using the standard settings

outlined in Sect. 9.2.1. Additionally, the whole measurement period yielded high quality results so that all data can be used for the following atmospheric studies. The elevation bias applied for the MAPA retrievals is 1° which was determined using the tool presented in Sect. 9.2.2. Finally, also all standard selection criteria introduced in Sect. 9.3 could be applied for Bayfordbury.

10.2 ANNUAL AND DIURNAL CYCLES

As a first characterisation of the results, annual and diurnal cycles of formaldehyde, glyoxal, NO_2 and aerosol results are discussed, after the data were filtered according to the criteria defined in Sect. 9.3.

To get an overview on the general temporal characteristics on an annual scale, the data were plotted as time series. Thereby, individual measurements are displayed as well as daily and monthly means. For the calculation of the latter, two filters based on the sampling sizes within the time periods over which the means were calculated have been applied. For daily means, days with less than three valid data points were marked with light colours, while monthly means were only calculated for months with at least 10 valid individual measurements. In addition, months with less than 30 valid single measurements were labelled with open circles.

To gain more detailed insights into typical seasonal patterns of the different species, box-whisker-plots were generated for each month. Both was done for total columns as well as for near-surface concentrations and aerosol extinctions (compare Sect. 9.2.3).

Figure 10.1 presents the described annual plots for formaldehyde, where the two upper panels show the time series of VCDs and near-surface concentrations. Both quantities show overall similar temporal evolutions with pronounced enhancements for the individual measurements (light grey) in late summer of all three depicted years which becomes even more evident if daily (blue) and monthly (red) means are considered. Low formaldehyde abundances are measured during winter periods. Especially for the VCDs, a close connection to the temperature becomes apparent. The summer of 2021, which was a rather moderate one, shows obviously lower values than the other summers. Further, the highest measured individual formaldehyde VCDs, exceeding values of 3×10^{16} molec cm^{-2}, occur at the end of the considered measurement period which coincides with the extreme heat wave in Great Britain in summer 2022. For the near surface concentrations, both phenomena are less pronounced but still visible. Detailed investigations of the temperature dependencies will be performed in the next section (Sect. 10.3).

The two lower panels of Fig. 10.1 depict the average seasonal behaviours of the formaldehyde VCDs (left) and surface concentrations (right) in terms of monthly box-whisker-plots. While the lowest values occur in January (mean VCD: around 0.2×10^{16} molec cm^{-2}, mean concentration: around 0.15×10^{11} molec cm^{-3}), both quantities show a rather smooth constant increase which peaks in August with

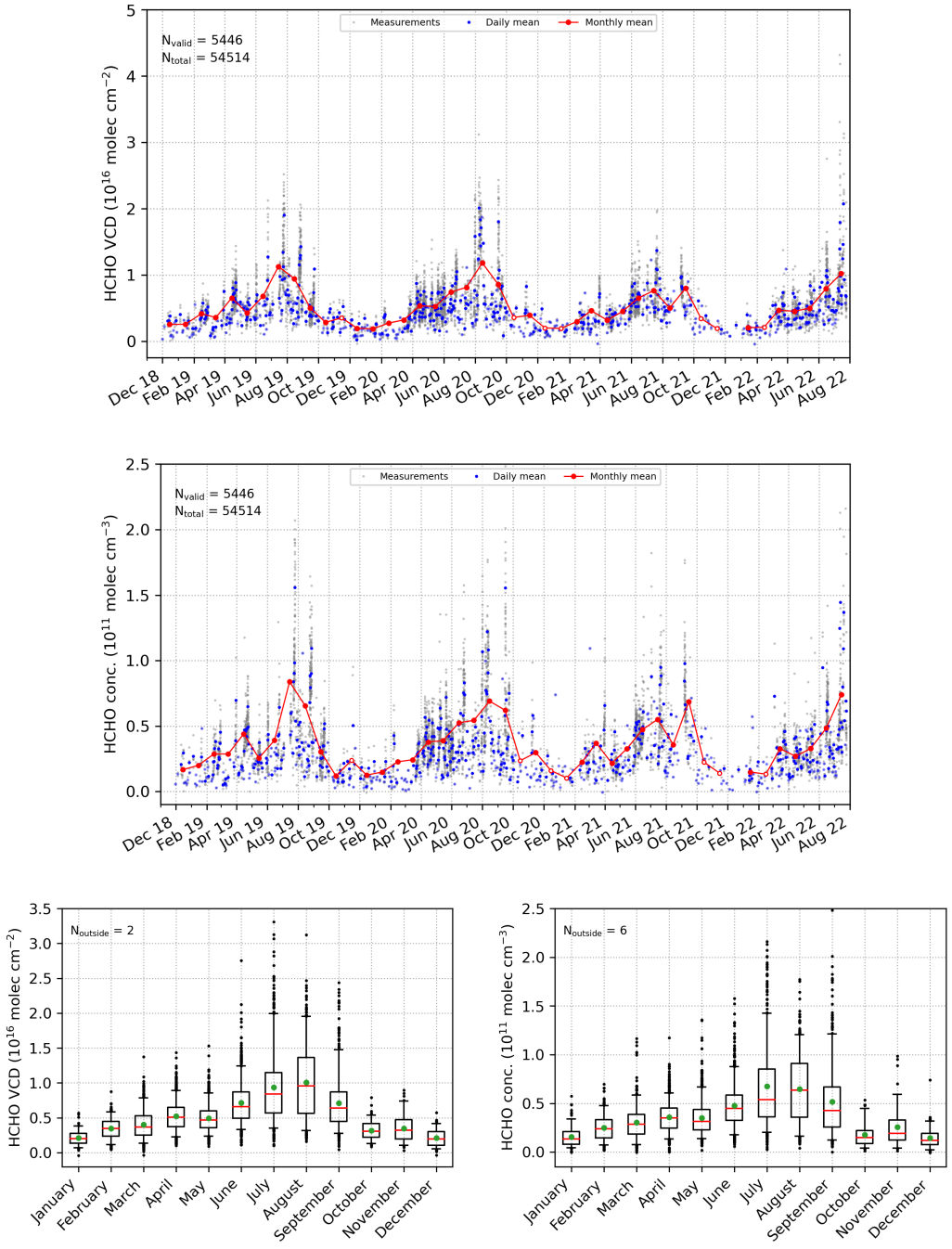


Figure 10.1: The two upper panels show time series of HCHO VCDs and surface concentrations for Bayfordbury. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of HCHO VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

mean vertical columns and surface concentrations[1] of 1×10^{16} molec cm^{-2} and 0.65×10^{11} molec cm^{-3}, respectively. Also the ranges of the measured formaldehyde VCDs and surface concentrations increase in summer (compare whiskers in the two lower panels of Fig. 10.1), probably because the conversion from precursors to HCHO strongly depends on meteorological parameters. After a slight decrease in September, an abrupt decline to values comparable to the minimum values in January is observed in October which is similarly pronounced for both quantities potentially indicating missing strong formaldehyde sources in autumn and winter with only background levels measured.

As described in Sect. 9.3, only MAPA results for a variable O_4 scaling factor (SF) are presented in the main part of this work. For Bayfordbury, no major differences in the annual characteristics of the formaldehyde results are observed for the SF of 0.8. While practically no differences are observed for the VCDs, slightly lower surface concentrations are obtained for the variable SF. This finding indicates differences in the profile shapes and how the total column is distributed vertically, however, details are discussed in Sect. 10.5.

Next, the diurnal cycle of formaldehyde is investigated. To do so, the data were separated according to the four meteorological seasons occurring at Bayfordbury (compare Sect. 4.2). The seasonal data were then binned into 1-hour bins and median values were calculated for each of these bins. In principle, also means could be used here, however, since means are more sensitive to outliers, rather noisy diurnal cycles were obtained and consequently median cycles were chosen to obtain more robust results.

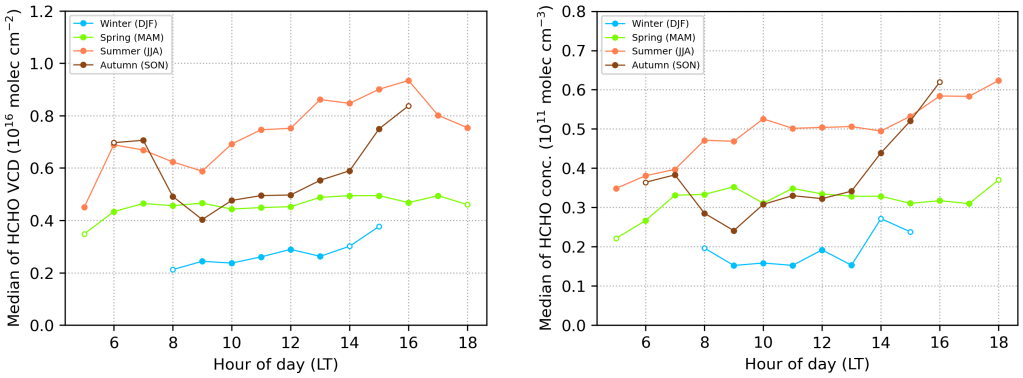


Figure 10.2: Diurnal cycles of HCHO VCDs (left) and surface concentrations (right) for Bayfordbury for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

Diurnal cycles of formaldehyde VCDs and near-surface concentrations obtained that way are shown in Fig. 10.2 for the four seasons. It should be noted that only values with more than 10 valid retrievals in the respective bins are displayed and hours with less than 5 % valid MAPA retrievals in the respective season are

1 The highest mean surface concentration is actually obtained in July. However, this value is comparable to the one in August and the medians (red lines), which are less sensitive to extreme values, suggest that the average formaldehyde abundance is higher in August.

marked by open circles. The latter labels hours with a below-average rate of valid MAPA retrievals and thus less robust results.

For both quantities, only weak diurnal cycles can be observed in winter and spring, whereas summer and autumn exhibit diurnal cycles with a notable increase of the formaldehyde abundance over the course of the day. Since this occurs for both VCDs and surface concentrations, this probably indicates sunlight-driven chemical formation of HCHO from precursor species. The enhanced values for both quantities during the early morning hours in autumn originate from the fact that median values at these times contain only data from September where sunrises are early enough to cover these hours with measurements and formaldehyde values are still enhanced compared to October and November. For summer, a similar effect can be observed for the VCDs in early morning and late evening. Here, measurements from June where days are longest and formaldehyde abundances are lower than in July and August lead to a low bias.

Following the same approach as for HCHO, Fig. 10.3 shows the annual characteristics of glyoxal. Similar to formaldehyde, glyoxal shows enhanced VCDs and surface concentrations in the late summer months (compare upper two panels). In late autumn and during winter, the glyoxal VCDs and surface concentrations are practically 0 which was not the case for HCHO, where still minimum values significantly above 0 could be observed. Like already seen in the case of formaldehyde, the summer of 2021 shows lower glyoxal abundances compared to the other years and the heat wave of 2022 leads to very high values. However, both features are less pronounced for glyoxal compared to HCHO.

The statistical representations of the annual cycles for glyoxal VCDs and near-surface concentrations (lower panels of Fig. 10.3) confirm that in winter values scattering faintly around 0 are retrieved. Starting in February both quantities show a rather smooth increase peaking in late summer. While HCHO reaches its maximum in August, the glyoxal values in July and August are basically the same. Mean values of around 1×10^{14} molec cm^{-2} and 1×10^{9} molec cm^{-3} are obtained in these months for VCDs and surface concentrations[2], respectively. As already seen for HCHO, the spreads of glyoxal VCDs and surface concentrations are largest in summer indicating a larger range of measured glyoxal abundances. In September, the glyoxal values are slightly decreasing for both quantities, followed by an abrupt decline to their minimum winter values of practically 0 in October. This abrupt decline is even stronger pronounced as compared to formaldehyde and indicates the absence of relevant glyoxal sources in late autumn and winter.

As for formaldehyde, almost no differences of the results obtained with the two applied O_4 scaling factors are observed for the VCDs, while smaller surface concentrations are retrieved using a variable SF.

The average diurnal variations of glyoxal VCDs (left) and near-surface concentrations (right) in the different seasons are presented in Fig. 10.4.

Like for formaldehyde, only a weak diurnal cycle can be observed in winter and values stay at practically 0 fitting the observed winter values of the annual cycles. The sudden increase of the surface concentrations at 3 p.m. is probably

2 Like for HCHO, the maximum of the surface concentrations is actually observed in July. However, the medians indicate similar values for both months.

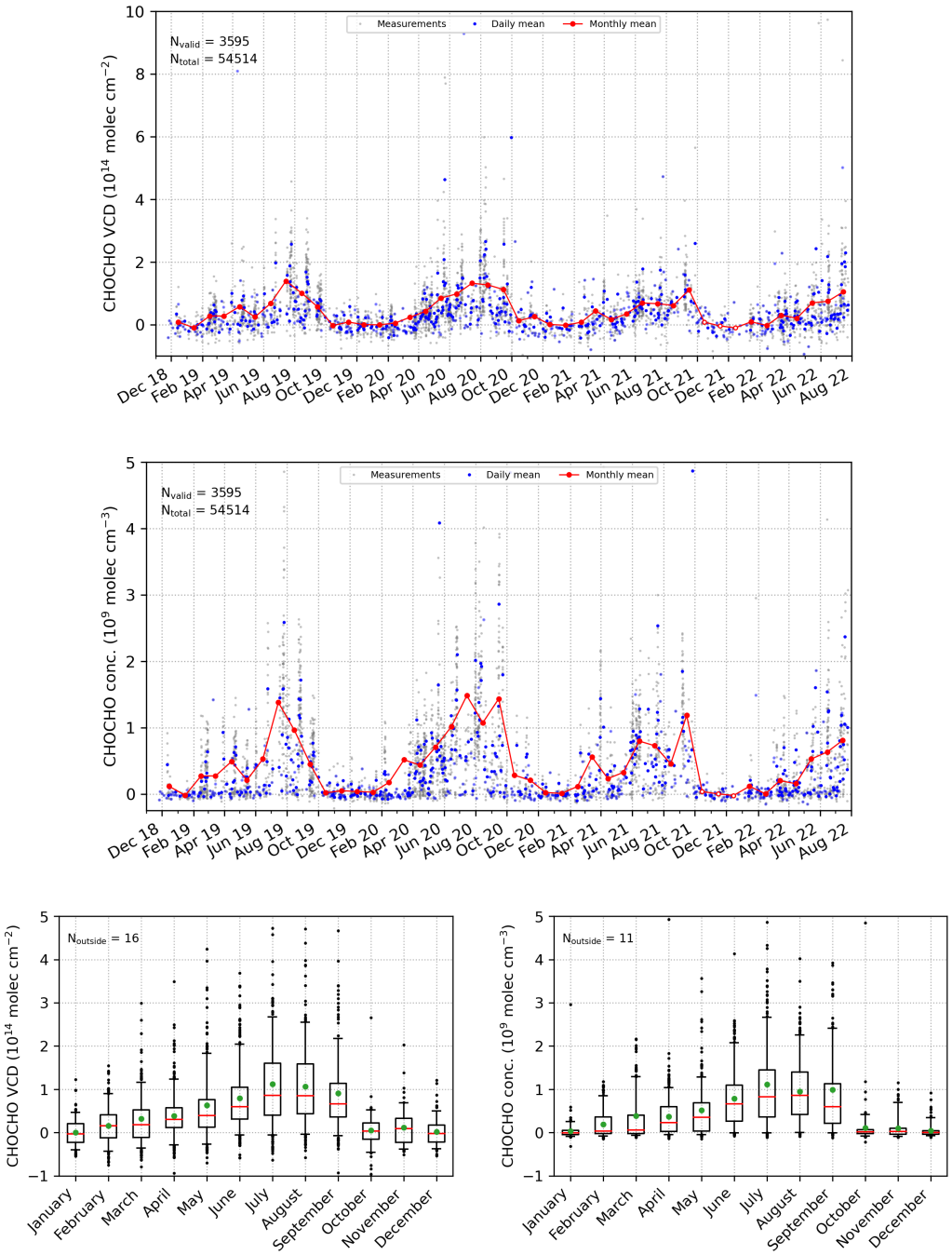


Figure 10.3: The two upper panels show time series of glyoxal VCDs and surface concentrations for Bayfordbury. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of glyoxal VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

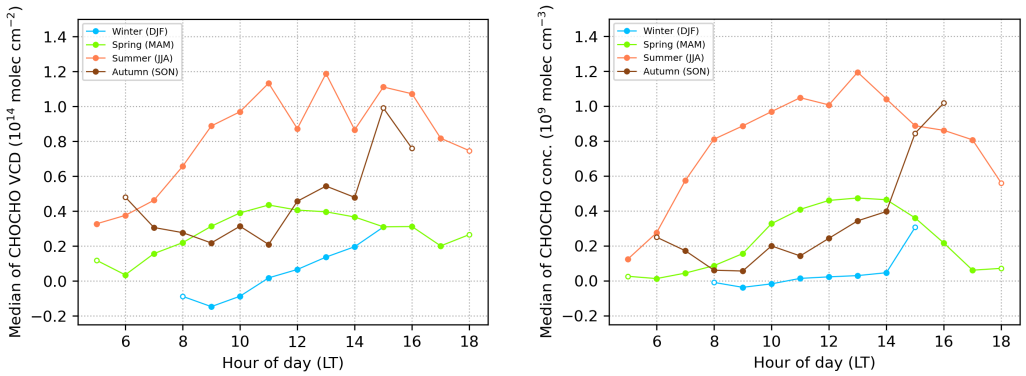


Figure 10.4: Diurnal cycles of glyoxal VCDs (left) and surface concentrations (right) for Bayfordbury for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

not significant as this data point is flagged for having less than 5 % valid MAPA retrievals. Contrary to HCHO, in spring a pronounced diurnal cycle can be noticed with an increase of the VCDs until noon followed by a slight decrease in the late afternoon. This cycle is even stronger pronounced for the surface concentrations where the increase is steeper and the maximum is reached later at around 1 p.m. followed by a strong reduction of the surface concentrations to values close to 0. The fact that the observed variation is stronger for the surface concentrations indicates the influence of local effects. Summer exhibits similar diurnal cycles as spring but at generally higher levels fitting the annual characteristics of the atmospheric glyoxal abundance. For the VCDs, maximum glyoxal is found slightly later than in spring and again the diurnal cycle appears to be weaker compared to the surface concentrations. Also for HCHO, for both quantities pronounced diurnal cycles were found in summer (Fig. 10.2). Autumn shows a diurnal cycle similar to formaldehyde with an increase of the glyoxal abundance over the course of the day. Similar variations are found for VCDs and surface concentrations and no decline of the values is observed in the late afternoon. However, it should be noted that this effect might be masked by statistical artefacts as already explained for HCHO since in autumn morning and evening values are biased high by the strong weight of September measurements in these time bins. For summer, a low bias might be present in the mornings and evenings as outlined above. Finally, it should be stressed that differences in the characteristics of the diurnal cycles of glyoxal VCDs and near-surface concentrations are observed for spring and summer which is not the case for formaldehyde where both quantities show similar diurnal patterns.

The annual patterns of NO_2 vertical columns and surface concentrations are summarised in Fig. 10.5.

Despite the rural surroundings, the vicinity and the influence of the metropolitan city of London and its pollution becomes evident in the measured NO_2 at Bayfordbury with elevated NO_2 levels occurring throughout the year. This will become even more evident in Sect. 10.3 where the dependence of the observed

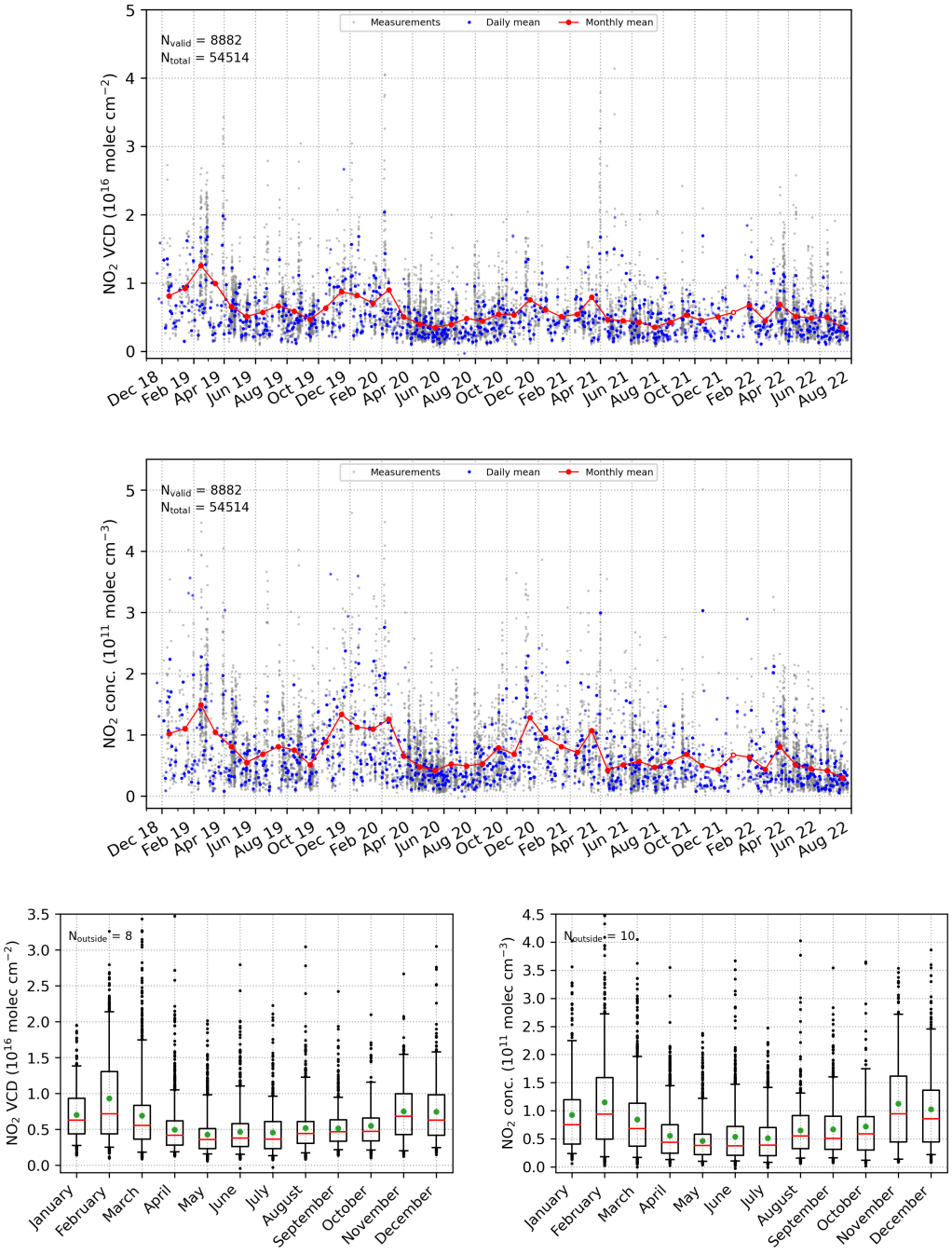


Figure 10.5: The two upper panels show time series of NO_2 VCDs and surface concentrations for Bayfordbury. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of NO_2 VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

NO_2 abundances on the wind direction is discussed. Nevertheless, highest NO_2 VCDs and near-surface concentrations are typically observed in winter, while lower pollution levels are found in summer. The latter is mainly connected to the shorter lifetime of NO_2 caused by enhanced photochemistry in summer. During winter, daily means exceeding 1.5×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3} can be obtained for VCDs and surface concentrations, respectively. In the summers, still daily values above 0.5×10^{16} molec cm^{-2} and 0.5×10^{11} molec cm^{-3} occur regularly, depending on the actual origin of the probed air masses. The influence of the origin of the air masses at Bayfordbury becomes also obvious by the large variability of the VCDs and surface concentrations, however, clearly larger variability can be observed for the near-surface concentrations. Further details of the effects of the wind direction on the measured NO_2 pollution are discussed in Sect. 10.3. Over the whole period, a weak decreasing trend can be observed in the measured NO_2 VCDs and surface concentrations (compare monthly means in the upper two panels). During the COVID-19 lockdown measures in spring 2020, only slightly lower pollution levels can be recognised.

The average annual variations of NO_2 VCDs (left) and near-surface concentrations (right) are presented in the two lower panels of Fig. 10.5 and basically confirm the previously outlined annual cycles with highest NO_2 in winter and minimum values in summer.

As for formaldehyde and glyoxal, also for NO_2 practically no differences between the results obtained with the two SF scenarios are observed for the VCDs, while generally lower near-surface concentrations are obtained using the variable SF. The same is valid for NO_2, nevertheless, the change of the surface concentrations is strongest for NO_2 amongst all retrieved species. This might be related to the fact that NO_2 is typically located very close to the ground. As mentioned above, more details are discussed in Sect. 10.5.

Figure 10.6 shows the diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for the four seasons.

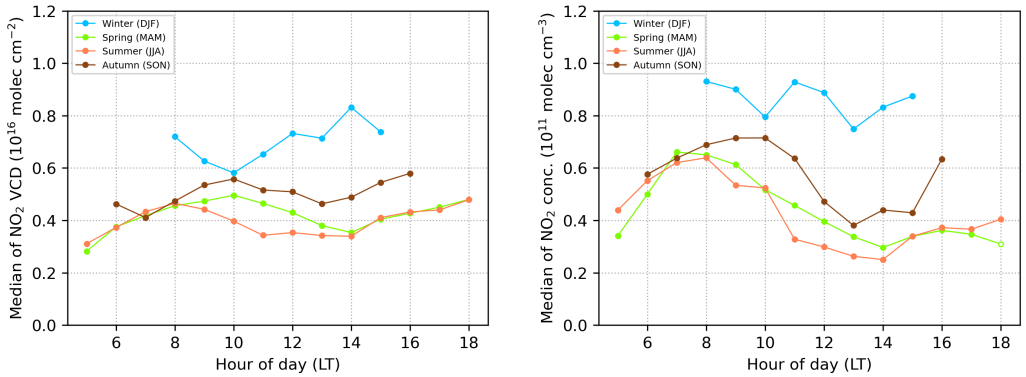


Figure 10.6: Diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for Bayfordbury for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

For both quantities, similar diurnal patters are found in spring, summer and autumn, whereby the patterns are more pronounced for the near-surface concentrations. In all of these seasons, an increase of the measured NO_2 abundance can be observed during the first half of the day peaking in the late morning. The exact time at which the maximum is reached varies between the seasons and is, especially in spring, slightly different for VCDs and surface concentrations. After reaching their maximum values, VCDs and near-surface concentrations drop to their minimum values. Again the times of the minima vary slightly between the seasons and the considered quantity. Towards sunset, a slight increase in the amount of NO_2 can be observed again. The finding that the diurnal patterns are stronger pronounced for the surface concentrations than for the VCDs and the differences of the times of the maxima and minima can probably be explained by at least two overlapping effects. Firstly, the planetary boundary layer rises in the course of the day leading to an extension of the layer into which the NO_2 is mixed. As a consequence, the VCDs are not affected by this effect, while the surface values are strongly influenced. Secondly, the surface concentrations are strongly affected by direct emissions, e.g. during the morning rush hour. The VCDs, on the other hand, are also influenced by transported pollution originating from London and its rush hour. Since these air masses have to be first transported to the measurement site, this leads to a time lag of the maxima and minima as well as to a smearing of the signal. In sum, this leads to a stronger diurnal cycle of the surface concentrations with earlier occurrences of the maximum values.

Highest NO_2 abundances are observed in winter. For this season, however, the interpretation of the diurnal cycle of NO_2 seems to be not so clear. No consistent pattern is found for the surface concentrations, while a similar diurnal cycle like for the other seasons can be observed for the VCDs. The reason for this difference is not clear but might be related to compensating effects of variations in the emissions and vertical mixing. However, due to shorter days, parts of the cycle are missing in the early morning and late afternoon.

Lastly, the results of the aerosol retrieval for Bayfordbury are discussed. The corresponding time series and seasonal variations of aerosol optical depth and near-surface aerosol extinction at 367 nm are depicted in Fig. 10.7. In addition to the MAX-DOAS results, the uppermost panel also includes AERONET AODs for comparison.

For both AOD and aerosol extinction, weaker annual cycles than for the trace gases are found. Higher values are typically found in spring (March and April) as well as in late summer/early autumn, while lower values are found during summer and winter. Both quantities show overall similar patterns and reasonable agreement to the AERONET AODs is observed. However, the lowest and highest values cannot be perfectly reproduced by the MAX-DOAS retrieval, e.g. in autumn and winter 2019.

The statistical analysis of the annual cycles of both quantities presented in the two lower panels of Fig. 10.7 confirms the findings outlined before. Highest AODs are observed in March and April and in late summer/early autumn with monthly means of around 0.2, while monthly means are below 0.15 during summer and winter. For the near-surface extinction, maximum and minimum values of around

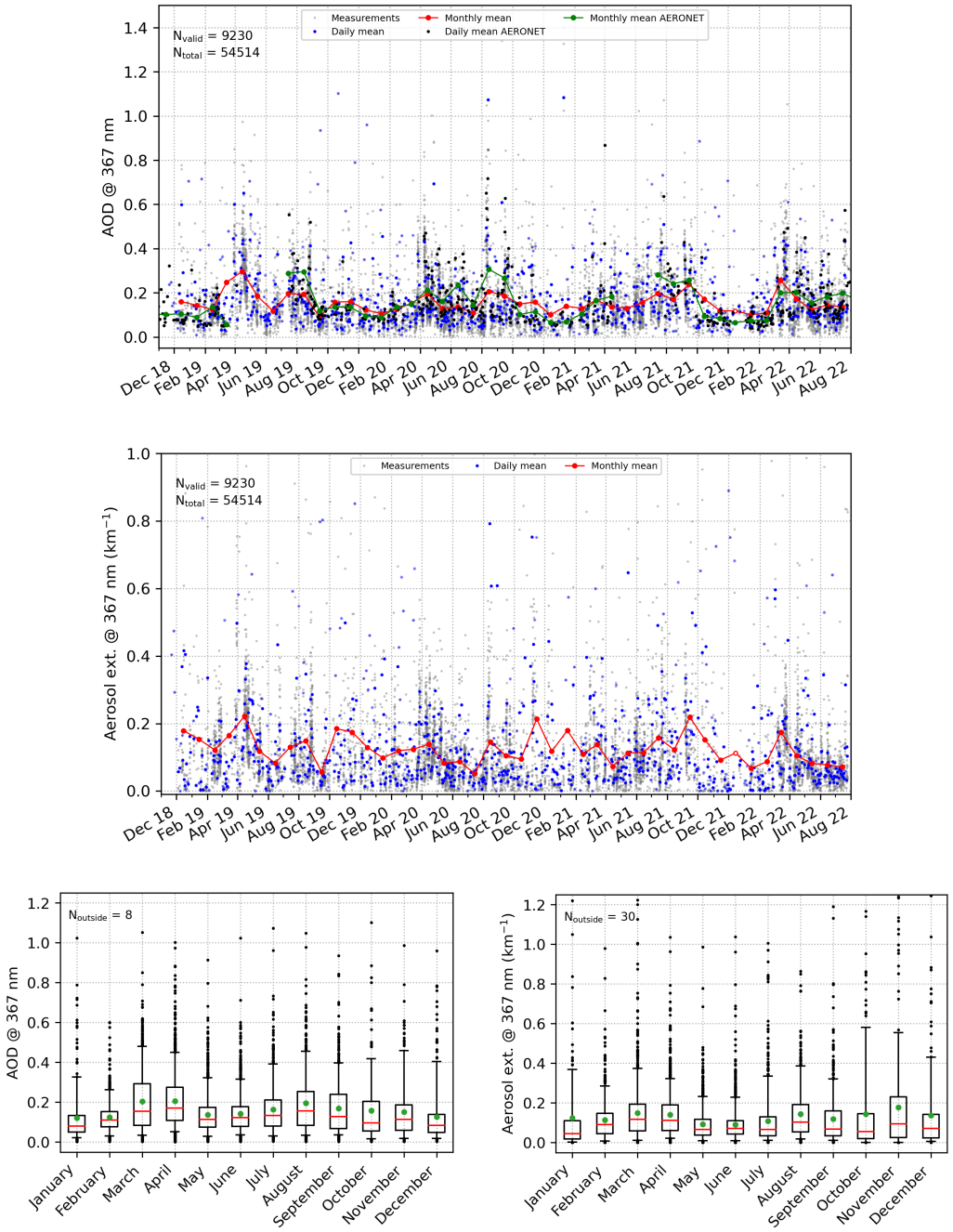


Figure 10.7: The two upper panels show time series of AOD and surface aerosol extinction at 367 nm for Bayfordbury. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and surface aerosol extinction at 367 nm (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

$0.15\,\mathrm{km}^{-1}$ and $0.1\,\mathrm{km}^{-1}$, respectively, are obtained in the respective seasons. Especially in late autumn and during winter, the considerably larger variability of the surface values becomes evident when looking at the 95th percentiles (upper whiskers) and the outliers (black circles). This observed large spread with many outliers towards high values considerably influences the means of these months. As a consequence, similar mean surface extinctions like in spring are obtained. However, the medians which are less affected by the large spread confirm the overall observed annual cycle with lower values in winter. The reason for the large spread in autumn and winter is most likely related to cloudy and foggy weather conditions at the measurement site (compare Sect. 4.2) leading to less valid and overall more variable MAPA retrievals. This effect is not found for the AODs since the total column is a more robust measure as compared to the vertical distribution of the extinction.

As outlined in detail in Sect. 3.3 and 9.2.1, the aerosol profile retrieval is performed using the measured O_4 dSCDs. Thus, the impact of the applied O_4 scaling factor on the results is larger than for the trace gas retrievals. While so far no significant influence of the SF on the presented trace gas VCDs could be observed, the retrieved AODs are considerably higher for a fixed SF of 0.8. Also the variability of the results is largely enhanced for the fixed SF. The same is valid for the aerosol extinction, however, this was also observed for the near-surface trace gas concentrations. Further, the agreement to the AERONET AODs is better for the variable SF. Despite these findings, it should be stressed once more that the overall annual behaviour as well as most of the results obtained in this work are not largely affected by the choice of the O_4 SF.

The diurnal cycles of AOD (left) and near-surface aerosol extinction (right) observed for the different seasons are presented in Fig. 10.8.

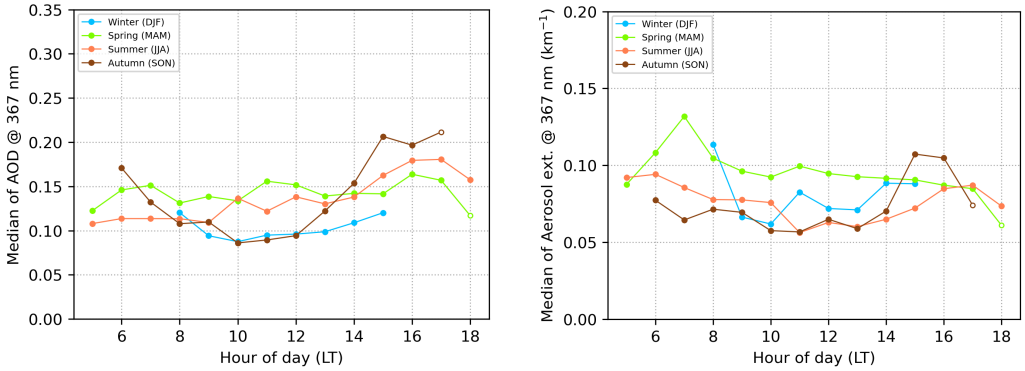


Figure 10.8: Diurnal cycles of AOD (left) and surface aerosol extinction (right) at 367 nm for Bayfordbury for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

For the AOD, only weak diurnal cycles occur in winter and spring, whereas a slight increase of the retrieved AODs in the course of the day can be observed in summer. In autumn, a U-shaped diurnal pattern is found where minimum values occur around noon and maximum values in the mornings and evenings. Parts of

this pattern might again be explained by the previously outlined dominance of September values in autumn time bins in the early morning and late afternoon leading to a high bias. However, the observed pattern is relatively strong (0.1 peak-to-peak) and appears to be, at least to some extend, real. The near-surface aerosol extinctions show no consistent diurnal patterns in winter and autumn where mainly low valid data availability and a larger variability affect the diurnal variations. Contrary to the diurnal patterns of the AODs, for spring and summer a slight decreasing trend from morning to afternoon is visible. Like for NO_2, this can be explained by the rise of the boundary layer height leading to higher reaching aerosol profiles in the course of the day. Similar to NO_2, in summer additionally a slight increase towards evening can be observed. Overall, the diurnal patterns of AODs and aerosol extinction are less pronounced compared to those of the trace gases and more difficult to interpret.

In addition to the species discussed in detail in this section, also HONO and water vapour retrievals were performed, yet both are not presented in the main part for Bayfordbury. For HONO, this is due to the difficulties of the retrieval and probably unreliable results for this station. As a concluding result, it can be stated that practically no enhanced HONO was measured at Bayfordbury. For water vapour, expected results showing a pronounced yearly cycle with high values in summer and low values in the (absolutely seen) dry winter are obtained, while almost no diurnal cycles are found.

10.3 INFLUENCE OF METEOROLOGICAL CONDITIONS

For some species, already the investigations of the temporal patterns discussed in the previous section suggested a close relation between their abundances and the temperature variations throughout the year. Therefore, in this section the dependence of the trace gas and aerosol abundances measured at Bayfordbury on the meteorological quantities temperature, wind speed and wind direction is investigated. Again, MAX-DOAS data were filtered according to the criteria defined in Sect. 9.3 and hourly means were calculated and used to minimise the scatter of the data. A description of how the weather data were treated can be found in Sect. 9.4.

Hourly mean HCHO VCDs (left) and near-surface concentrations (right) as function of temperature are presented in the uppermost panels of Fig. 10.9. Consistent to the observed annual cycles, for both quantities a distinct temperature dependence showing overall increasing formaldehyde levels with increasing temperatures can be observed. Up to temperatures of around 10 °C, almost no temperature dependence appears. In the temperature range between 10 and 20 °C, the observed increase of the HCHO values with increasing temperatures is relatively moderate and seems to be rather linear. Starting at temperatures of about 20 °C, the slope increases notably indicating a strong temperature dependence of the measured HCHO values. However, due to the rather cool climate at Bayfordbury not many measurements lie in this temperature range. A similar behaviour of obtained HCHO VCDs was also observed by Borovski et al. (2015) who performed measurements close to Moscow. Overall, the temperature dependence

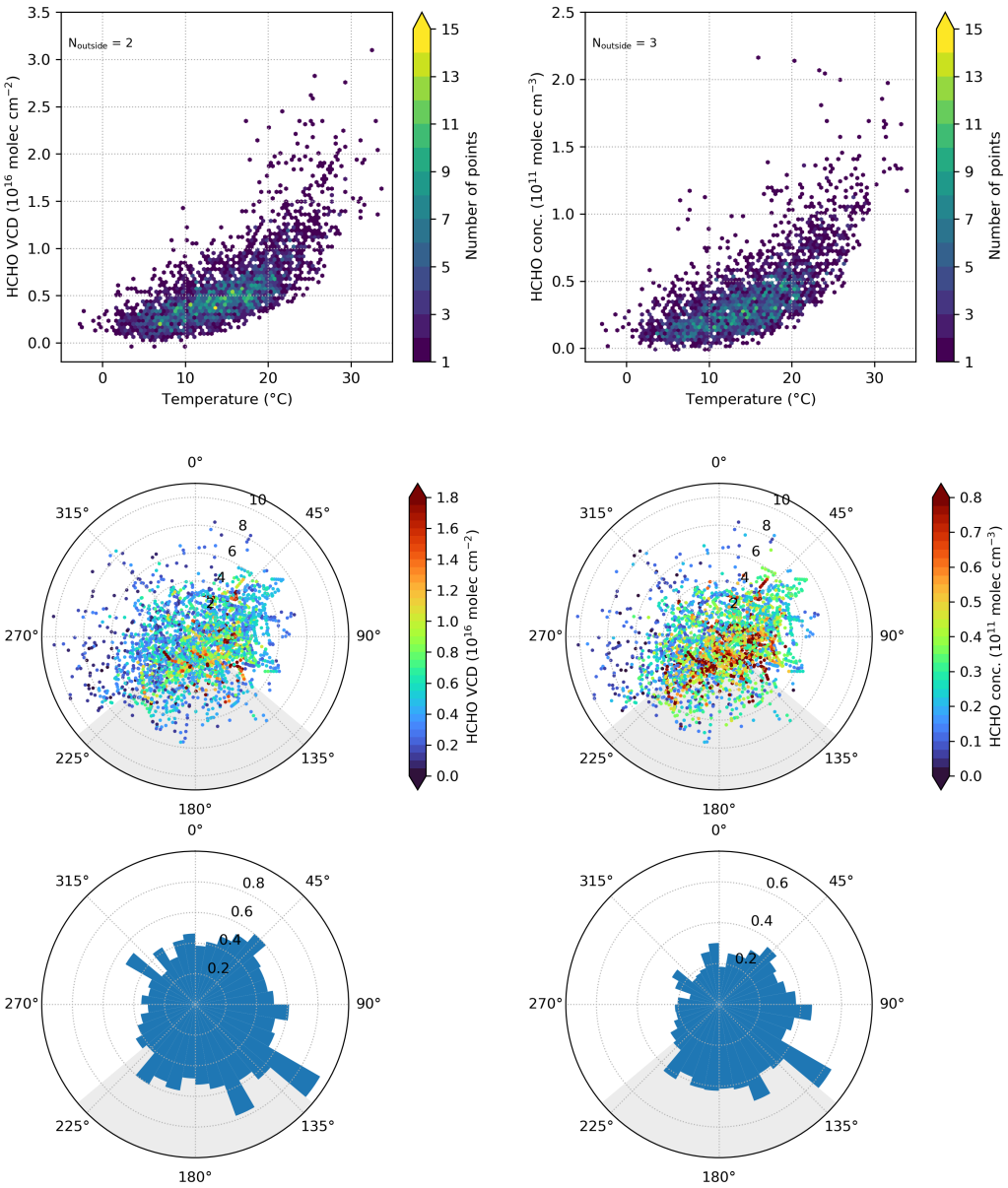


Figure 10.9: The two uppermost panels show hourly mean HCHO VCDs (left) and surface concentrations (right) measured at Bayfordbury as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HCHO VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HCHO VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius. The shaded areas mark the wind sector which is directly affected by the pollution of London.

is strongly non-linear when the complete temperature range is considered. This indicates the importance of biogenic VOC emissions for the HCHO abundances measured at Bayfordbury especially at high temperatures, i.e. in summer. In particular, the emissions of biogenic VOCs acting as precursors for HCHO show

a pronounced temperature dependence themselves which is valid for different classes of VOCs (Steiner and Goldstein, 2007). For instance, isoprene shows an exponential dependence on air temperature with an optimum at around 35 to 45 °C after which the emissions decrease again (the latter is not relevant at Bayfordbury). Further, also the emissions of monoterpenes and the direct emissions of oxygenated VOCs, including HCHO itself, show non-linear increases with increasing temperature (Steiner and Goldstein, 2007). Another aspect probably contributing to the strong temperature dependence of the HCHO abundances is that higher temperatures at mid-latitudes are typically associated with high solar radiation having a twofold effect. Firstly, isoprene and monoterpene emissions have shown to be also light-dependent, i.e increasing emissions under enhanced radiation (Steiner and Goldstein, 2007). Secondly, high solar radiation leads to very active photochemistry and thus an effective formation of HCHO from its precursors. Lastly, it should be stressed that quasi-equal patterns are observed for VCDs and surface concentrations.

As outlined in Sect. 4.2, depending on the prevailing wind direction the air masses at Bayfordbury can be originating from rather pristine regions or the heavily polluted city of London. Thus, next the dependence of the observed HCHO abundances at Bayfordbury on the wind characteristics is studied.

For that, first the obtained hourly mean formaldehyde VCDs and their corresponding near-surface concentrations were plotted in polar plots as shown by the two middle panels in Fig. 10.9. In these plots, the VCDs (left) and surface concentrations (right) are colour-coded, while the radius indicates the corresponding wind speeds in $\mathrm{m\,s^{-1}}$ and the angular position represents the wind direction. To get a more concise picture of the average angular distribution of the measured HCHO abundances, the hourly mean HCHO VCDs and surface concentrations were binned into 10°-bins according to their corresponding wind directions. Finally, medians were calculated for each of these wind sector bins and displayed in polar diagrams (lowermost panels of Fig. 10.9). As an orientation, the wind sector which is expected to be directly influenced by the pollution of London is marked by shaded areas.

Overall, enhanced formaldehyde abundances at Bayfordbury are measured under winds from a large sector spanning from north-east to south-west. In this sector, the median VCDs and surface concentrations are approximately twice as high as in the remaining sectors and reach values of around 0.5×10^{16} molec $\mathrm{cm^{-2}}$ and 0.35×10^{11} molec $\mathrm{cm^{-3}}$ for VCDs and surface concentrations, respectively. Both quantities show in general similar patterns, however, higher surface concentrations occur under southerly winds compared to easterly winds, slightly exceeding 0.4×10^{11} molec $\mathrm{cm^{-3}}$, while for the VCDs comparable values are found under both wind directions. The overall similar patterns of both quantities indicate the importance of secondary formation of HCHO from biogenic precursors, nevertheless, the higher surface concentrations under southerly winds (from London) suggest an additional effect of direct emissions and/or additional VOC emissions from the London area.

Typically highest HCHO VCDs and near-surface concentrations occur at rather low wind speeds and under south-easterly winds (sector from ca. 120 to 160°).

In this sector, the old harbour *London Docklands*, a major oil storage, several harbours and industries along the river Thames and its mouth, including a 732 MW gas-fired power station (Coryton) and a major LNG terminal (Isle of Grain), are located upwind from Bayfordbury. Especially, the latter two might be associated with emissions of methane and various other hydrocarbons and VOCs all yielding HCHO as an intermediate product during their oxidation. Remarkably, the absolute maxima of formaldehyde VCDs and surface concentrations are found under ca. 125°, the sector in which the gas-fired power station and the LNG terminal are situated at distances of around 50 and 60 km, respectively. The local maximum in the wind sector in which overall lower HCHO is measured (at roughly 300 to 310°) might be related to Luton Airport.

The overall angular pattern with higher formaldehyde abundances under north-easterly to south-westerly winds compared to the remaining sectors can probably be explained by a combination of several factors. Firstly, in general higher anthropogenic activities can be found in this sector since more towns, cities and industries are situated there. Here, also anthropogenic activities like harbours along the east-coast of Great Britain and the ship traffic in the North Sea between Great Britain and the European continent might play a role. Further, a possible influence of the pollution of western European countries like the Netherlands and Belgium cannot completely be ruled out since cities like Rotterdam and Antwerp, both having large harbours, have a distance less than 350 km to the observatory.

An interesting aspect here is that for NO_2 (compare Fig. 10.11) easterly and westerly winds lead to similar abundances. This difference in the angular patterns of both species might be a hint for the secondary formation of HCHO from precursors while the co-emitted NO_2 (the interconversion between the directly emitted NO and NO_2 happens quickly) might already be diluted and/or degraded. Further, enhanced abundances of both species are observed under southerly winds stressing the influence of the emissions of London on the measured HCHO abundances at Bayfordbury. Nevertheless, the maximum of the angular distribution of HCHO seems to be slightly shifted to the east as compared to NO_2 (compare lowermost panels in Fig. 10.9 and 10.11) corresponding to maxima observed under south-easterly and south-south-easterly winds for HCHO and NO_2, respectively. This indicates the importance of emissions of precursor substances for the formation of HCHO rather than direct emissions.

These observed dependencies suggest that the influence of anthropogenic emissions of both precursors for the secondary formation of HCHO and formaldehyde itself leads to higher abundances under easterly and southerly winds. Contrary to that, background levels of HCHO are measured under westerly and northerly winds. Besides the influence of different emission sources, also meteorological effects might have an important influence on the observed patterns. Typically westerly winds are associated with higher wind speeds, more clouds and rain and thus possibly lower temperatures and less solar radiation as compared to weather conditions under other wind directions. The importance of air temperatures and thus biogenic emissions of mostly precursor VOCs on the measured HCHO abundances has already been shown above.

Figure 10.10 shows the dependence of the glyoxal VCDs (left) and surface concentrations (right) measured at Bayfordbury on temperature (uppermost panels) as well as wind speed and direction (middle and lowermost panels).

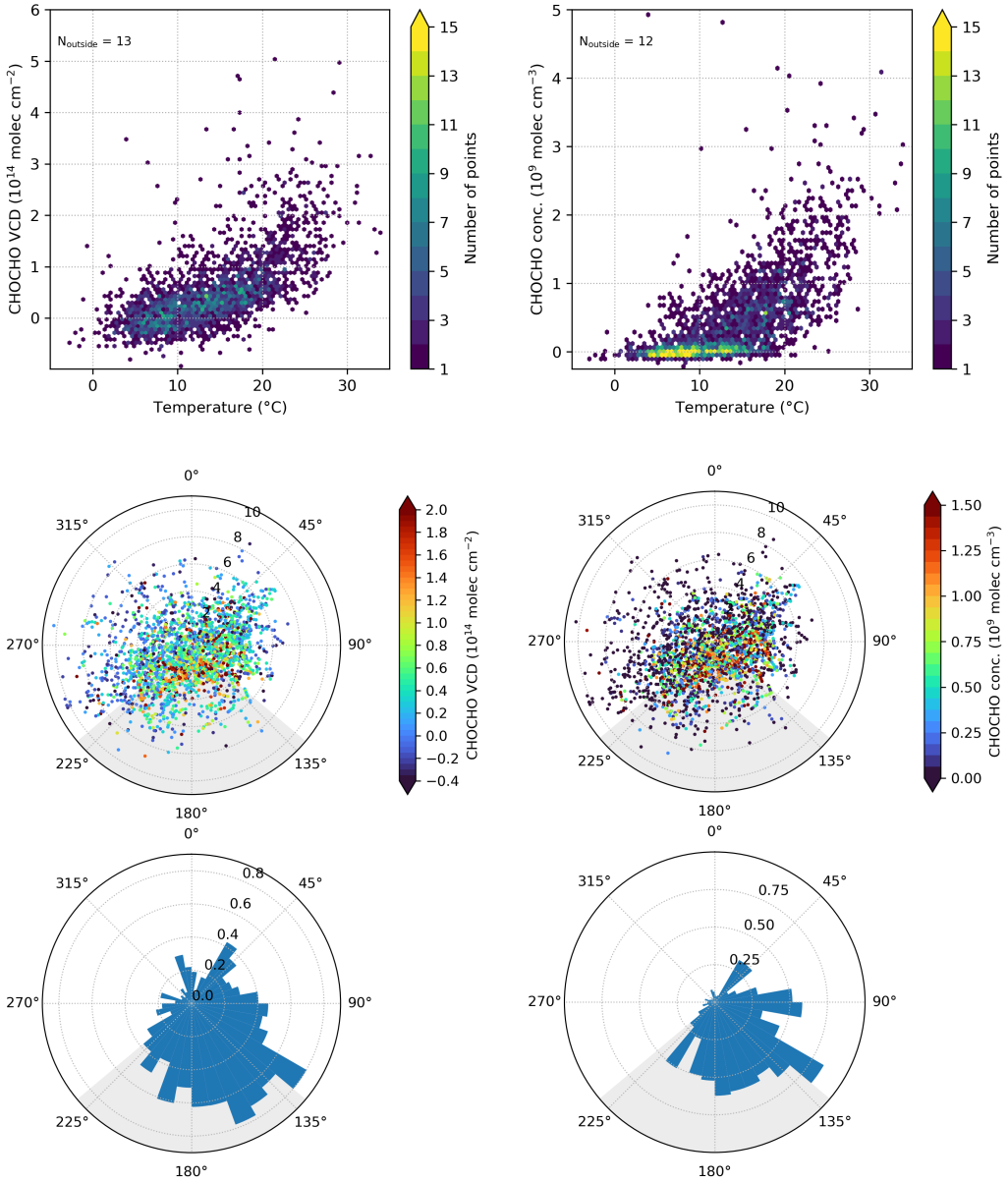


Figure 10.10: The two uppermost panels show hourly mean glyoxal VCDs (left) and surface concentrations (right) measured at Bayfordbury as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean glyoxal VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median glyoxal VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius. The shaded areas mark the wind sector which is directly affected by the pollution of London.

The overall temperature dependence of the glyoxal abundance is similar to the one of formaldehyde probably indicating similar sources. Consistent to the observed annual cycles, an increase of the glyoxal abundances with increasing temperature is found. Like for formaldehyde, the temperature dependence is strongly non-linear if the complete temperature range is considered. However, differences in the details can be observed between both species.

Firstly, below approximately 10 °C, corresponding to winter and late autumn, both VCDs and surface concentrations mainly scatter around 0. The temperature range between 10 and 20 °C is associated with an increase of the measured glyoxal abundances with increasing temperature. However, while the VCDs show a rather smooth increase in this temperature range, still a significant fraction of the near-surface concentrations is close or equal to 0. The findings in these two temperature ranges are different to formaldehyde. The observed discrepancy of the characteristics of both quantities in this temperature range probably indicates the absence of local (biogenic) glyoxal sources in late autumn and winter as already discussed in the previous section (compare Fig. 10.3). Like for HCHO, the slope increases considerably at around 20 °C. Nevertheless, it should be stressed that data coverage in this temperature range is sparse at Bayfordbury related to its climate. Similar to HCHO, the findings for the temperature dependence of the glyoxal abundances suggest that during summer (higher temperatures) biogenic emissions of VOCs acting as precursors are of major importance since the biogenic emissions themselves are highly temperature dependent. The slightly different temperature dependencies found for the glyoxal VCDs and near-surface concentrations are likely related to the fact that direct emissions are significantly less important for glyoxal compared to HCHO and that glyoxal is formed from a smaller subset of precursor VOCs (compare Sect. 2.4).

The most prominent feature of the wind rose plots for glyoxal are median VCDs and surface concentrations of practical 0 that occur under winds originating from a large sector ranging from south-west to north-east (compare lowermost panels of Fig. 10.10). Only minor enhancements of the glyoxal VCDs exist in this sector and might occur when summer background levels are measured at low wind speeds since they cannot be attributed to any distinct sources. Enhanced median glyoxal levels at Bayfordbury are observed when the site is influenced by winds originating from a sector spanning from north-east to south-west. Glyoxal VCDs and surface concentrations show overall similar angular patterns which are more smooth for the VCDs. In contrast to HCHO, for glyoxal clearly higher abundances occur under southerly winds where the measured VCDs are up to 1.5 times higher than under easterly winds. For the surface concentrations, the latter effect is weaker pronounced and probably partly hidden by the larger scatter.

Also for glyoxal, overall higher levels are measured under low wind speeds which is similarly pronounced for both quantities. The overall highest glyoxal VCDs and surface concentrations are obtained under south-easterly winds originating from a sector spanning from 120 to 160°, whereby the sector appears to be more confined at a range from 120 to 140° for the surface concentrations. As outlined above for formaldehyde, many potential sources lie in this sector. Nevertheless, given the larger differences in the glyoxal abundances between the

eastern and more southern sectors, the influence of anthropogenic sources located rather close to Bayfordbury seems to be more important than sources further away at the coast east of Bayfordbury. This is also supported by the finding that the angular patterns of glyoxal are apparently more similar to the ones of NO_2 (compare Fig. 10.11) than those of HCHO and NO_2 suggesting an important influence of general anthropogenic emissions for the presence of enhanced glyoxal. This might be related to emissions of aromatic compounds by vehicles which are high yield precursors of glyoxal (compare Sect. 2.4 as well as Volkamer et al., 2001 and Nishino et al., 2010). Further, the lifetime of glyoxal is typically shorter than the one of HCHO (compare Sect. 2.4). Nonetheless, similar to HCHO, the absolute maximum occurs under 125° again suggesting the importance of precursor substances emitted by methane handling industries. Lastly, also for glyoxal meteorological effects might affect the observed overall patterns.

The close relation between formaldehyde and glyoxal is further discussed in the upcoming sections.

NO_2 VCDs (left) and near-surface concentrations (right) as function of air temperature are presented in the two uppermost panels of Fig. 10.11.

No clear temperature dependence can be observed for NO_2 VCDs and surface concentrations. Only a slight tendency towards lower NO_2 abundances at higher temperatures is visible. Further, the majority of the highest VCDs and surface concentrations are observed at temperatures below ca. 15 °C. These findings fit the previously presented annual cycles with maximum NO_2 levels occurring during winter and occasional episodes of elevated NO_2 abundances that can occur throughout the year. Similar results are found for both quantities with slightly larger scatter for the near-surface concentrations.

Despite the observed annual cycle of NO_2 and the general tendency that highest NO_2 levels occur during winter at lower temperatures, the day-to-day variability observed throughout the year (compare upper panels in Fig. 10.5) can be explained by changing wind directions since the air masses probed at Bayfordbury can be originating from rather pristine regions or the heavily polluted city of London. The dependence of the measured NO_2 VCDs (left) and surface concentrations (right) on wind speed and direction is summarised by the middle and lower panels in Fig. 10.11.

Highest NO_2 abundances are measured under winds originating from a sector spanning from east to south-west indicating the importance of London and its emissions for the air quality at rural Bayfordbury. In this sector, median VCDs and surface concentrations exceed values of 0.6×10^{16} molec cm^{-2} and 0.8×10^{11} molec cm^{-3}, respectively. Significantly lower NO_2 levels occur in the remaining sectors that are not influenced by the pollution of London where values typically stay below 0.4×10^{16} molec cm^{-2} and 0.4×10^{11} molec cm^{-3}, respectively. For the VCDs, a second much weaker maximum occurs under north-westerly winds which is almost not noticeable for the surface concentrations probably indicating a source located more far away. Lowest NO_2 abundances occur under north to north-easterly winds. Besides this small difference, the observed patterns are similar for both quantities.

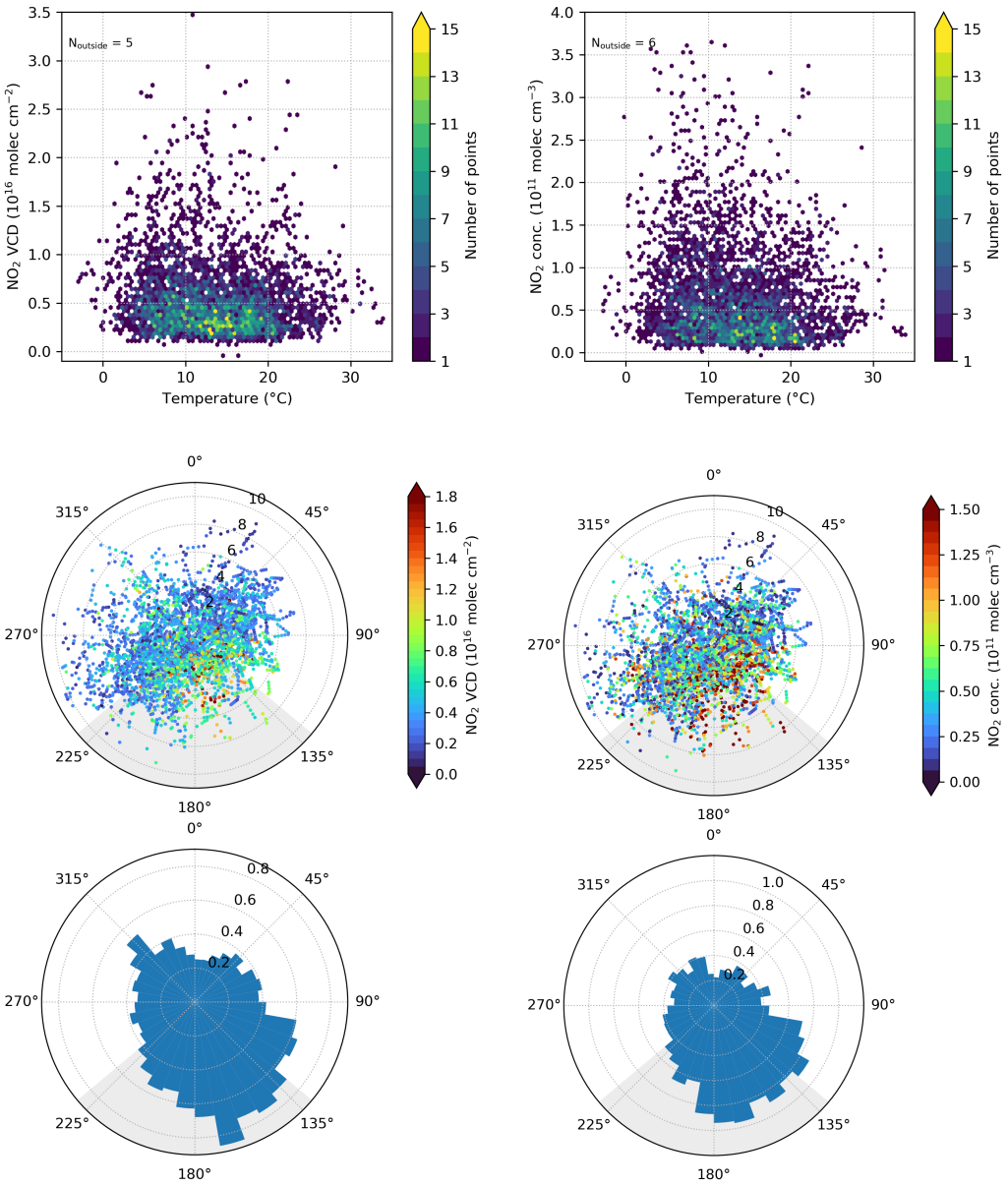


Figure 10.11: The two uppermost panels show hourly mean NO_2 VCDs (left) and surface concentrations (right) measured at Bayfordbury as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean NO_2 VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median NO_2 VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius. The shaded areas mark the wind sector which is directly affected by the pollution of London.

Like for glyoxal and formaldehyde, highest NO_2 abundances typically occur at low wind speeds which might be related to less effective mixing processes. Enhanced surface concentrations can occur under almost all wind directions at rather low wind speeds. While elevated NO_2 abundances are found in a rather

wide sector ranging from 90 to 230°, the maximum is observed in a more confined sector spanning from approximately 100 to 190°. As outlined above, especially in the eastern part of this sector several industrial sources east of London and along river Thames and its mouth region are located. The emissions of these industries and the mostly traffic related emissions of the city add up leading to the observed maxima. Also the city centre of London is situated within this more confined sector. The minor maximum which mainly occurs for the VCDs under north-westerly winds might be related to transported emissions from Northampton, Milton Keynes and Luton including its airport which all lie in this sector, at distances of 75 km, 55 km and 20 km, respectively. However, no effects from the Stansted and Heathrow airports can be observed.

Contrary to HCHO and glyoxal, easterly and westerly winds lead to similar NO_2 abundances. Details of this difference in the angular patterns of the different species were already discussed above. While the angular patterns of HCHO and NO_2 show several differences, the patterns found for glyoxal seem to be more similar to the ones of NO_2 (for details compare above).

Again meteorological effects may also influence the observed patterns in the NO_2 distribution. Most important might be higher wind speeds under westerly winds which are in general associated with lower NO_2 levels as shown in the middle panels of Fig. 10.11.

Finally, the effects of meteorological conditions on the measured AODs and surface aerosol extinctions at 367 nm are investigated. Both quantities are shown as function of temperature in the two uppermost panels of Fig. 10.12.

Both aerosol quantities do not show any temperature dependence. This finding fits the only weakly pronounced annual cycles which showed only slightly elevated AODs and surface extinctions in spring and late summer/early autumn compared to the rest of the year.

The dependence of the observed AODs (left) and surface extinctions (right) at 367 nm on wind characteristics is presented in the middle and lowermost panels of Fig. 10.12. In general, enhanced AODs and surface extinctions at 367 nm, similar to HCHO, are obtained under winds originating from a sector ranging from north-east to south-west. Median AODs and surface extinctions under winds from these sectors are approximately twice as high as in the remaining sectors leading to values reaching almost 0.2 and above 0.1 km^{-1} for AODs and surface extinctions, respectively. Both quantities show similar angular patterns, however, in contrast to the surface extinctions, slightly higher AODs occur under easterly winds compared to southerly winds where values only reach up to 0.15. This finding might indicate that under easterly winds the aerosol profiles reach higher into the lower troposphere with well-mixed and thus vertically well-distributed aerosol. A further effect might be the formation of SOA during the transport from more far away emission sources along the east-coast of Great Britain since also HCHO and partly glyoxal were enhanced in this sector. This effect is probably less important under southerly winds since here the dominant sources are located closer to Bayfordbury. Lastly, the distribution of the near-surface extinctions shows a larger variability than the one of the total AODs.

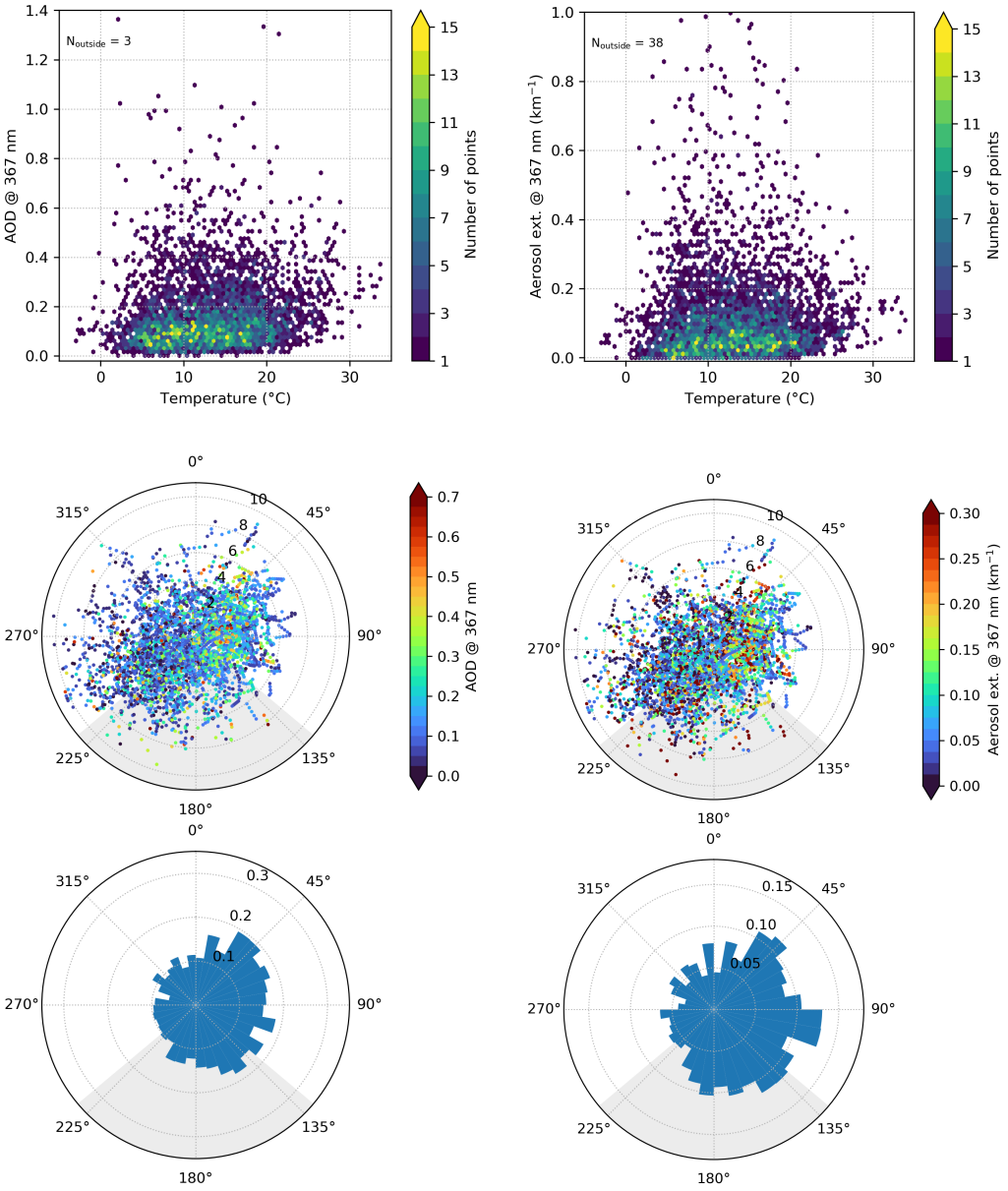


Figure 10.12: The two uppermost panels show hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) measured at Bayfordbury as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median AODs (left) and surface aerosol extinctions at 367 nm (right) obtained under different wind directions in bins of 10° are indicated by the radius. The shaded areas mark the wind sector which is directly affected by the pollution of London.

The observed distribution of the aerosol abundances fits the locations of the major sources, being the city of London, industries and harbours along the River Thames and possibly also harbours along the east-coast of Great Britain and the ship traffic in the North Sea as well as the European continent. However,

for the latter three, no influence was observed for NO_2. Contrary to that, in the remaining sector spanning from south-west to north-east considerably lower aerosol loads are found. The major reason for this is the absence of significant aerosol sources since, despite some smaller towns, rather pristine conditions prevail in this wind sector. Nevertheless, as outlined in detail for HCHO, again the possibility that meteorological effects have a strong influence on the observed angular distributions has to be stressed.

Finally, it can be stated that no strong dependence on the wind speed can be observed.

Again, for water vapour and its dependence on meteorological parameters the expected results are obtained. Firstly, an increase of the atmospheric water vapour content with increasing temperature is found. Secondly, no strong dependence on the wind characteristics is observed. Here, mainly south-westerly winds lead to slightly enhanced water vapour abundances which is caused by the fact that air masses originating from this direction are typically warmer and more humid since they were transported over the North Atlantic.

10.4 TRACE GAS CORRELATIONS

Despite differences in details, the previous studies for Bayfordbury suggest a close relation between formaldehyde and glyoxal as both species showed similar characteristics on annual and daily scales, e.g. both species have their annual maxima during the warm summer months. Furthermore, the dependencies on meteorological parameters showed similar patterns for both gases. As outlined in the last section, also anthropogenic emissions appear to have a significant influence on the atmospheric abundances of both species since enhancements of glyoxal and HCHO also occurred in wind sectors which also showed elevated NO_2 abundances. Given all these observations, this section examines the relation between HCHO, glyoxal, NO_2 and aerosol using correlations of the respective column and near-surface quantities. To get further insights into the relation between formaldehyde and glyoxal, the annual cycle of the ratio of both is studied at the end of this section.

In order to correlate the different species with each other, the corresponding MAX-DOAS data were first filtered using the criteria defined in Sect. 9.3. Next, hourly means were calculated to minimise the scatter of the data and obvious outliers were additionally excluded from the correlations. The number of excluded data points is provided in each correlation plot shown below.

Figure 10.13 shows the correlation of glyoxal and formaldehyde VCDs and near-surface concentrations in the upper and lower parts, respectively. Time series of hourly mean values of both species are shown in the left panels of each part. The right panels show the correlation plots in which the point densities are colour-coded. As an orientation, simple linear regressions were performed whose results are shown in red.

For both VCDs and surface concentrations, reasonable correlation between glyoxal and HCHO is found yielding similar R^2-values of 0.58 and 0.61, respectively. Also the linear regressions show similar results for both quantities exhibiting

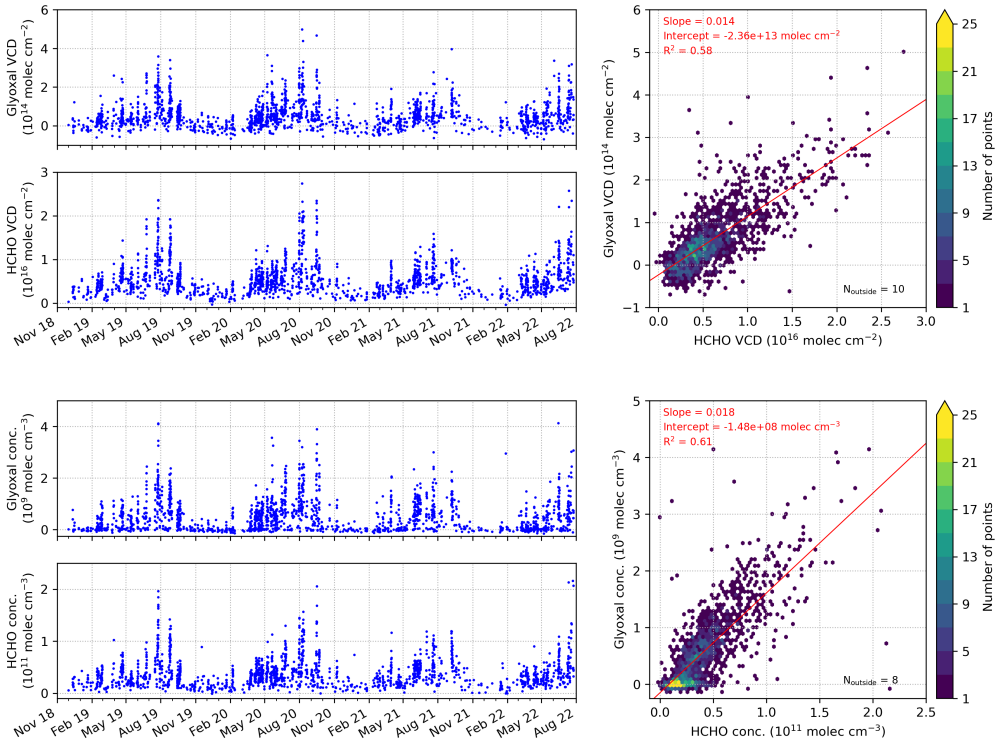


Figure 10.13: Comparison of glyoxal and formaldehyde VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The results of a simple linear regression are shown in red, while the number of excluded data points is indicated in the lower right corners.

slopes of 1.4 % and 1.8 %. The marginally larger R^2 and the slightly larger slope found for the surface concentrations might indicate different profile shapes for both gases, with the bulk of glyoxal located closer to the ground while the HCHO profiles seem to reach higher. Detailed investigations on the profile shapes of the different trace gases are performed in the next section.

The correlation plot of the surface values shows a high point density at glyoxal values close or equal to 0 which reflects results in autumn and winter. Especially in these seasons, for both quantities rather weak correlations are obtained since glyoxal values scatter around 0, whereas still non-zero HCHO abundances are measured[3]. The latter indicates that also during winter formaldehyde sources exist, while no sources of glyoxal appear to be present. In summer and spring, however, reasonable correlations similar to the overall correlations can be observed for both VCDs and surface concentrations.

These findings once more stress the close relation between both species, especially in summer.

To gain more insights into the effects of anthropogenic emissions on the formaldehyde abundances, HCHO and NO_2 VCDs and surface concentrations are

3 Seasonal correlation plots are not shown in the main part of this work. For the resulting R^2-values of the correlations on seasonal basis, refer to Table A.2 in the appendix.

correlated. The corresponding plots are presented in Fig. 10.14. Since no simple linear relation between both trace gases is found, no linear regressions were performed here.

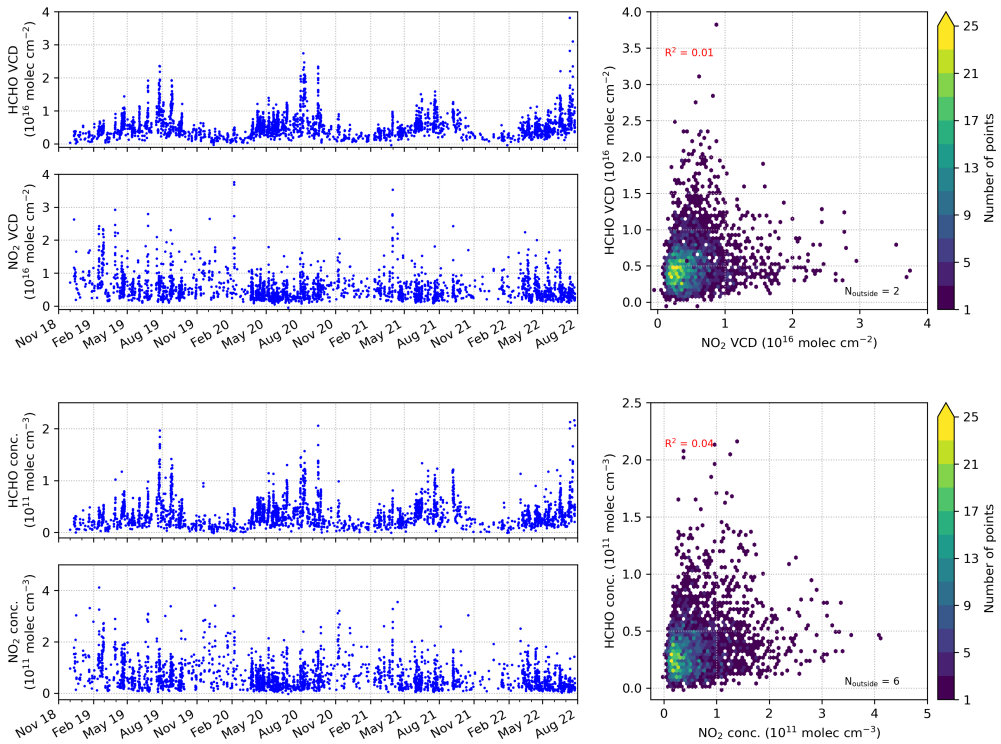


Figure 10.14: Comparison of formaldehyde and NO_2 VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Unlike the fairly good correlation between formaldehyde and glyoxal, more complex patterns emerge from the correlation of formaldehyde and NO_2. For both VCDs and surface concentrations, two major branches are visible corresponding mainly to winter and summer[4]. Values in between are obtained for spring and autumn.

The winter branches are associated with low to moderate HCHO abundances staying below approximately 0.75×10^{16} molec cm^{-2} and 0.5×10^{11} molec cm^{-3} for HCHO VCDs and surface concentrations, respectively. Reasonable correlations between both species are found for these branches and the highest values of formaldehyde occur under rather high NO_2 loads exceeding 1.5×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3}. These branches indicate that in winter anthropogenic emissions, of possibly both precursors and HCHO itself, determine the atmospheric formaldehyde levels since biogenic activity is low. As secondary formation is less effective under low radiation conditions and cold temperatures which prevail in winter, direct emissions of HCHO might have a large effect based on these simple correlations.

4 For the resulting R^2-values of the correlations on seasonal basis, refer to Table A.2 in the appendix.

Contrary to that, the summer branches show that highest formaldehyde occurs at only moderate NO_2 loads below 1×10^{16} molec cm^{-2} and 1.5×10^{11} molec cm^{-3}. These findings show that in summer, in addition to direct (biogenic) emissions, secondary formation from mostly biogenic precursors determines the HCHO abundances. Nevertheless, in summer the patterns are more complex and relatively weakly pronounced sub-branches exist which show moderate formaldehyde abundances at rather high NO_2 levels. These probably indicate episodes with less biogenic activity (e.g. cloudy scenes with less solar radiation) where anthropogenic emissions have a significant influence on the formaldehyde abundances even during summer.

The correlations of glyoxal and NO_2 VCDs and near-surface concentrations are shown in Fig. 10.15.

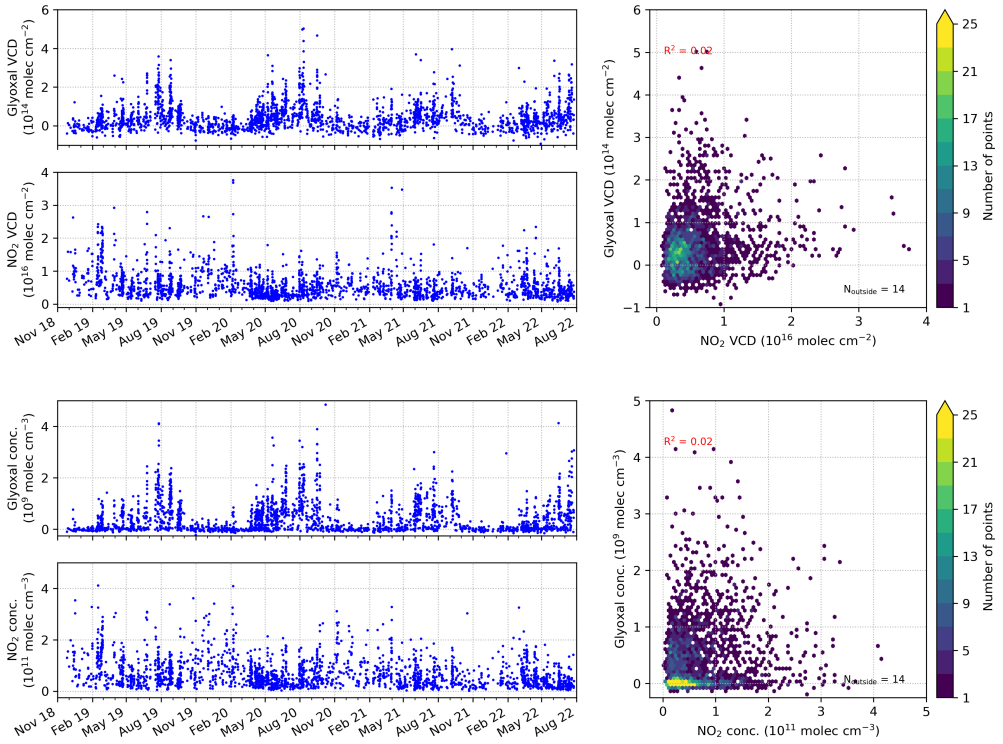


Figure 10.15: Comparison of glyoxal and NO_2 VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Similar to the correlation of formaldehyde and NO_2, also the correlation between glyoxal and NO_2 shows two distinct branches corresponding to summer and winter, respectively, while values in between are observed for spring and early autumn. Despite overall similar patterns for both quantities, the branch-pattern appears to be more extreme for the near-surface concentrations which show a large fraction of values close or equal to 0 corresponding to measurements in late autumn and winter (compare Sect. 10.2). For the VCDs, however, a larger scatter around 0 is observed leading to a less dense accumulation of data points.

Consequently, the winter branch only shows non-zero glyoxal values at very high NO_2 abundances exceeding 2×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3} for VCDs and surface concentrations, respectively. This shows that the glyoxal levels appear to be closer linked to high NO_2 loads than formaldehyde which was already seen for the dependencies on the wind direction presented before. Further, the winter branch suggests that no glyoxal sources exist during winter.

As expected, the patterns of the summer branches are very similar to the ones of the comparison between HCHO and NO_2. Again, highest glyoxal levels occur at moderate NO_2 loads below 1×10^{16} molec cm^{-2} and 1.5×10^{11} molec cm^{-3}. Also the sub-branches with moderate values at rather high NO_2 abundances are visible for glyoxal.

Next, the correlations of the different trace gases with the aerosol results are examined. Figure 10.16 presents the correlations between HCHO and aerosol column (upper part) as well as near-surface (lower part) quantities.

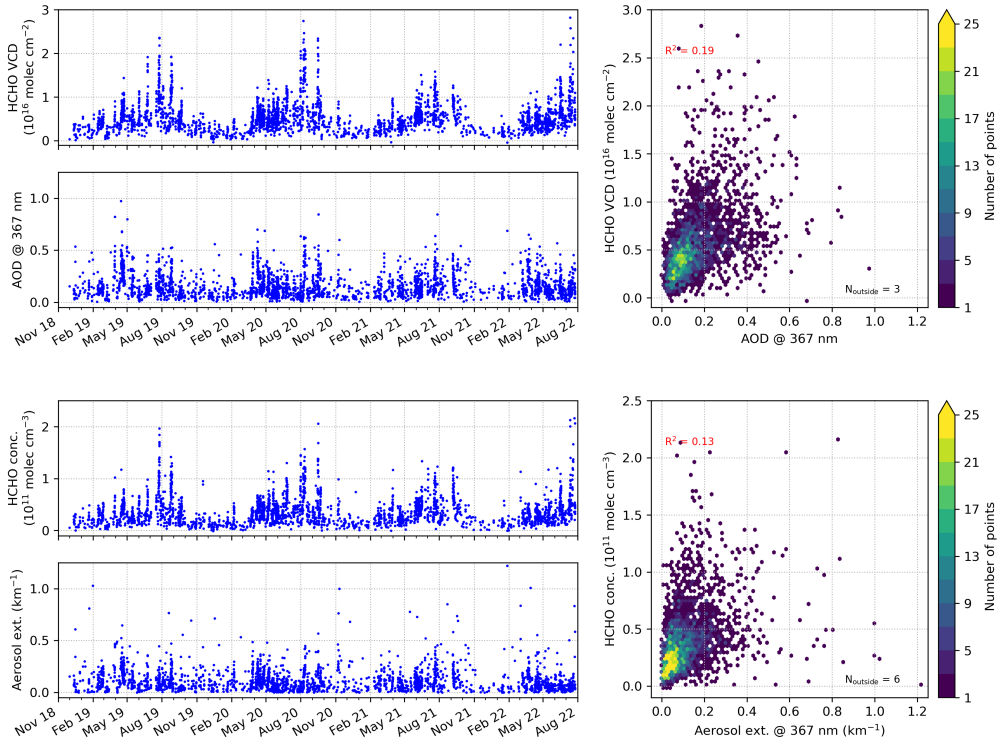


Figure 10.16: Comparison of formaldehyde and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

For both quantities, despite showing rather diverging patterns, increased aerosol loads are typically associated with enhanced HCHO abundances. This relation is slightly stronger pronounced for the column quantities of both species for which an R^2 of 0.19 is obtained. Separating the data into the different seasons mainly reflects the strong annual cycle of HCHO and separates the diverging correlation patterns into two branches yielding worse correlations in winter and rather good

correlations during summer. This applies in particular to the column quantities which yield an R^2 of 0.32 in summer. Overall, these observations are consistent with the annual cycles of both species as well as their dependencies on the wind direction which showed comparable angular patterns, e.g. common enhancements under easterly winds as presented in Fig. 10.9 and 10.12, suggesting similar sources and/or a possible effect of the formation of SOA.

Figure 10.17 shows the correlation between glyoxal and aerosol results.

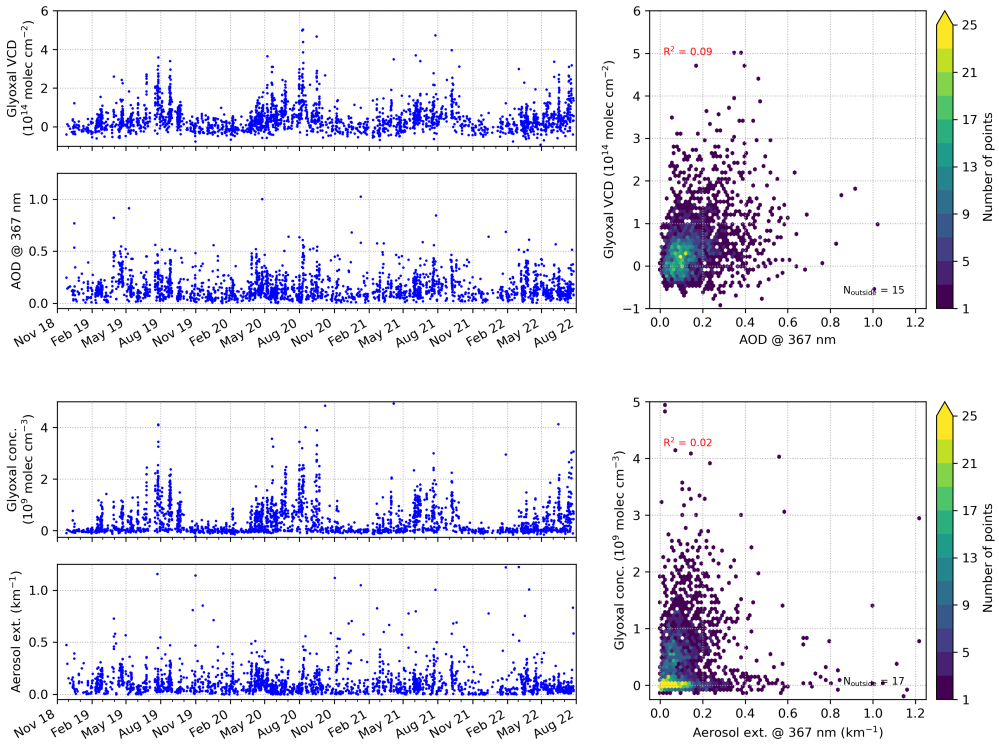


Figure 10.17: Comparison of glyoxal and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Given the similar annual cylces of HCHO and glyoxal, also for the correlation between the glyoxal and aerosol abundances typically higher aerosol loads are associated with higher glyoxal abundances. However, the glyoxal results show larger scatter. Thus, also the patterns of the correlations between glyoxal and aerosol show larger scatter and are, compared to the ones between HCHO and aerosol, more divergent (branch structure) indicated by low R^2-values of 0.09 and 0.02 for column and near-surface quantities, respectively. For the near-surface concentrations, again a large fraction of the data shows values close or equal to 0 (compare lower part) corresponding to measurements during late autumn and winter (compare Sect. 10.2) where almost no correlation between glyoxal and aerosol is found. The latter is also valid for the correlation between glyoxal VCDs and AODs. In summer, better correlation is found for columns and near-surface quantities with R^2-values of 0.15 and 0.06, respectively. Nevertheless, in sum the

correlation between the glyoxal and aerosol results is weaker and shows more scatter compared to the one of HCHO and aerosol. This can be explained by the angular patterns of the dependencies on the wind direction where largest enhancements of the glyoxal abundances are found under south-easterly and southerly winds while largest aerosol loads are observed under easterly winds (compare Sect. 10.3). Hence, HCHO and aerosol abundances are apparently closer related to each other than glyoxal and aerosol abundances.

Lastly, the correlation of the NO_2 and aerosol results is summarised in Fig. 10.18.

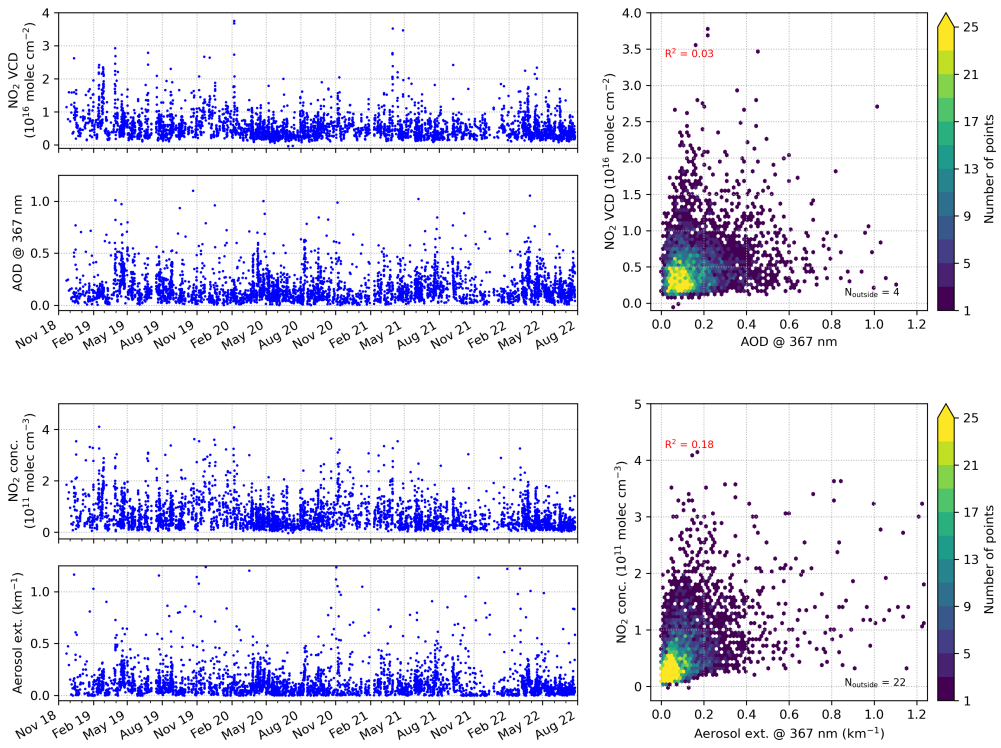


Figure 10.18: Comparison of NO_2 and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Like for the correlations of the HCHO and glyoxal abundances with the corresponding aerosol results, no simple linear relation between NO_2 and aerosol abundances is found and rather complex patterns emerge from the correlation plots. Compared to Fig. 10.16 and 10.17, these patterns diverge even more and apparently two branches exist for both column and near-surface quantities. One is associated with high NO_2 abundances at rather small aerosol loads and shows a strong increase of both NO_2 quantities when the respective aerosol quantity increases. The second branch is related with moderate NO_2 VCDs and surface concentrations at rather high aerosol loads whereby both NO_2 quantities increase only lightly when the aerosol abundance is increased. A comparison of the wind rose plots of both species (compare Sect. 10.3) shows that these branches can likely be linked to situations when the air masses probed at Bayfordbury origin-

ate from easterly directions which are associated with high aerosol loads but no strong enhancements of NO_2 and from London under south-easterly to southerly winds under which the maximum in the NO_2 angular distribution is found but no strong enhancements of aerosol occur. Under westerly winds, both species show no specifically high values. Such conditions are represented by values between the two main branches in the correlation plots. Separating the data into the different seasons mainly reflects the annual cycles of both species, whereby both branches are visible throughout the year yet their relative contributions vary. Altogether, these findings show that at Bayfordbury high NO_2 abundances are not always associated with high aerosol loads which probably has different reasons. In particular, both species can be co-emitted but processed differently in the atmosphere leading for instance to a degradation of NO_2 while the respective air mass is transported to Bayfordbury under easterly winds (compare considerations in Sect. 10.3). Contrary to that, emissions originating from London largely increase the NO_2 abundances under south-easterly to southerly winds since these emissions occur closer to the measurement site. Finally, the fact that the correlation pattern is slightly more confined for the near-surface quantities of both species probably indicates that these are related closer to each other.

To further investigate the relation between formaldehyde and glyoxal, the annual cycle of the ratio of both species is examined next. For this purpose, for each measurement with valid formaldehyde and glyoxal results (VCDs and near-surface concentrations), the ratio of both trace gases was calculated. It should be mentioned that from a total of 54 514 only 2013 data points remain after filtering for MAPA flags, SZA and overlapping HCHO and glyoxal results which corresponds to only 3.7 % remaining data. This low fraction is caused by a challenging glyoxal retrieval and the frequent occurrence of cloudy and rainy episodes at Bayfordbury (compare Sect. 4.2). Further, formaldehyde and glyoxal data which were included in the previous studies of the individual species might be excluded here in case results for one of the two species were not available for a measurement.

Average annual cylces in terms of monthly box-whisker-plots, as explained in detail in Sect. 10.2, are shown in Fig. 10.19 for VCDs (left) and surface concentrations (right).

Despite the large variability of the R_{GF}, for both quantities a pronounced annual cycle can be observed. In particular, a smooth annual variation becomes evident when monthly medians are considered. While larger variability is found for the VCDs in winter, during summer a larger variability can be seen for the surface concentrations. The observed annual cycle mainly reflects the annual characteristics of glyoxal which were found to be stronger pronounced than the annual cycle of formaldehyde. As a consequence, median values around 0 are obtained in late autumn and during winter. The slight increase of the glyoxal abundances in early spring is counteracted by a stronger increase of the formaldehyde abundances still leading to R_{GF}-values below 1 %. Maximum median ratios are obtained in late spring and during summer (May to September), where values of up to 1.2 % in August and 1.8 % in June are obtained for VCDs and surface concentrations, respectively. Caused by the abrupt drop of the glyoxal

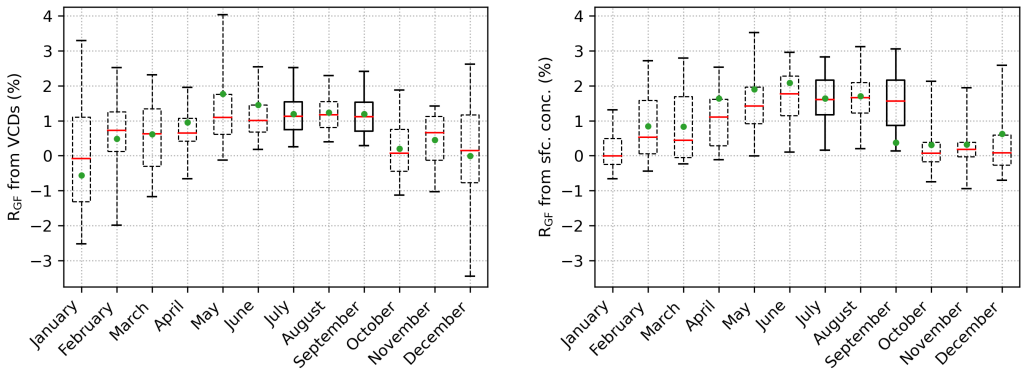


Figure 10.19: Monthly box-whisker-plots of the annual cycles of the ratio R_{GF} obtained from VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Dashed boxes indicate months with less than 5 % remaining data.

levels in October, the R_{GF}-values drop to their winter values around 0. Although the ratios exhibit a large overall variability, slightly higher values are obtained for the surface concentrations and annual median R_{GF}-values of 1.0 and 1.4 % are calculated for VCDs and surface concentrations, respectively. These in general higher ratios for the surface concentrations indicate that most of the glyoxal is located close to the ground, whereas the HCHO profiles seem to reach higher. Further details of the different profile shapes are discussed in the next section.

10.5 VERTICAL PROFILES

So far, VCDs and near-surface concentrations were studied. These investigations already suggested that differences in the vertical distributions of the different species exist. For example, from the trace gas correlations and the R_{GF}, it could be concluded that glyoxal typically seems to be concentrated quite close to the ground. Therefore, in this section vertical profiles of formaldehyde, glyoxal and NO_2 concentrations as well as profiles of aerosol extinction at 367 nm are investigated.

For that purpose, after applying the usual filters (compare Sect. 9.3), median profiles were calculated for each of the four seasons. The obtained seasonal profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction are presented in Fig. 10.20, where different colours mark the profiles of different seasons.

The vertical profile of formaldehyde concentrations shows its highest values up to a height of about 500 m, followed by a smooth decrease with increasing altitude. Note that the steep increase from the ground up to heights of around 300 m, which is found in a similar way for all species, is most likely an artefact caused by the specific profile parametrisation used in MAPA. Details of this behaviour were already discussed in Sect. 9.2. The formaldehyde profile reaches higher up compared to the other trace gases and the majority of the formaldehyde content

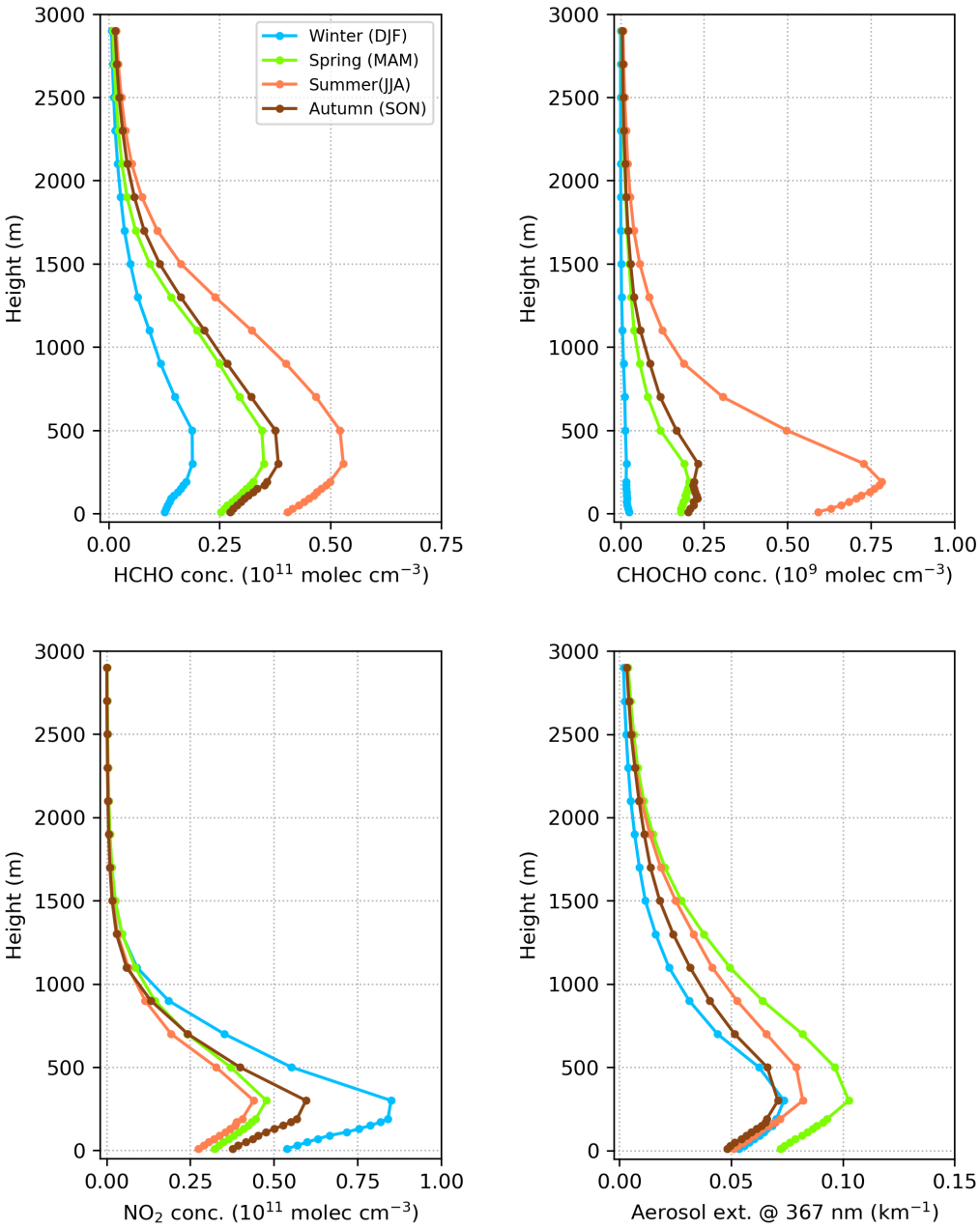


Figure 10.20: Seasonal median profiles of formaldehyde (upper left), glyoxal (upper right) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the Bayfordbury measurements. Profiles for the different seasons are indicated by different colours.

is distributed up to a height of around 1.5 km. For all seasons, similar relative median profile shapes are observed. Only in summer, the layer in which the bulk of formaldehyde is concentrated seems to be marginally thicker which might indicate the strong influence of secondary formation and stronger vertical mixing in the warm season. Nevertheless, the similarly high reaching profiles suggest that secondary formation may have an effect throughout the year. Here, it is

interesting to note that the considerations in the previous sections suggested that in winter anthropogenic emissions, including both HCHO itself and precursors, are of major importance for the HCHO abundances at Bayfordbury. The results for the profiles, however, indicate that direct emissions are probably less important than the secondary formation from (anthropogenic) precursor VOCs.

Compared to formaldehyde, the glyoxal profiles are shallower as their highest concentrations are found below a height of approximately 300 m and the majority of the glyoxal abundance is concentrated below around 1 km. This result is consistent with the findings of the previous section, where for example higher ratios between glyoxal and formaldehyde were found for the near-surface concentrations than for the VCDs. Similar median profiles are found for summer, spring and autumn although the latter two are more noisy since only little valid data are available for glyoxal, especially in autumn. Fitting the results from the earlier investigations, no enhanced glyoxal concentrations can be observed in winter.

Similar to glyoxal, the NO_2 concentration profiles are rather shallow and highest values are found in a layer reaching to roughly 300 m, while the majority of the NO_2 is again located below about 1 km. However, the decrease after reaching their maximum values is steeper for the NO_2 profiles as compared to glyoxal. For NO_2, the shallow profiles can be explained by the fact that it is emitted close to the ground and no secondary formation takes place. Again for all seasons similar relative median profile shapes are observed, while mainly the annual variability of the NO_2 loads is reflected. Further, no difference in the layer thickness can be observed between the seasons which is a surprising finding since typically the boundary layer reaches higher during summer. Probably the shorter lifetime in summer compensates the effect of stronger vertical mixing in the warmer period of the year.

The profiles of aerosol extinction at 367 nm are similar to the formaldehyde profiles and the highest values are found up to a height of about 500 m, followed by a smooth decrease. Further, the majority of the aerosol content is distributed below a height of around 1.5 km. Fitting the observed annual cycles, highest aerosol extinctions are found in spring, whereas the enhanced values in late summer/early autumn are masked by the lower values during the rest of both seasons (compare Sect. 10.2). Once more, only minor differences between the profile shapes and heights of the different seasons occur. However, for aerosol apparently a slightly shallower profile is found for winter indicating an overall shallower boundary layer.

As outlined earlier in this chapter, the main results of the different species are not largely affected by the choice of the O_4 scaling factor. For Bayfordbury, also basically no differences in the trace gas VCDs were observed, while consistently lower surface concentrations were obtained applying a variable SF. For the aerosol retrieval, the effect was strongest and also the AODs were smaller for the variable SF. As speculated earlier, this might indicate differences in the actual profile shapes and how the total column is distributed vertically.

A comparison of the seasonal median profiles obtained with a fixed SF of 0.8 (Fig. B.14 in the appendix) and a variable SF (Fig. 10.20) shows that in principle similar profile shapes and heights are obtained. However, a detailed look exhibits

that indeed slightly higher near-surface concentrations are obtained for all species using the SF of 0.8, while in the upper parts of the profiles slightly steeper decreases of the concentrations with increasing altitude towards 0 can be observed. Consequently, the two SFs actually lead to slightly different vertical distributions of the total column. As described above, this effect is very pronounced for NO_2 which is likely related to the fact that NO_2 is typically located relatively close to the ground. For the aerosol retrieval, however, this steeper decrease towards higher altitudes cannot be observed and both AODs and near-surface extinctions are higher for the fixed scaling factor. As shown before, the variable SF leads to less scatter, more valid retrievals and a better agreement between MAPA and AERONET AODs and therefore results for a variable scaling factor are shown in the main part of this work.

For water vapour, whose results are not shown here, relatively high reaching (up to about 2.5 km) vertical profiles are obtained. Similar relative profile shapes are found for the different seasons and the absolute values mainly reflect the annual variability of the atmospheric water vapour content.

10.6 COMPARISON TO TROPOMI

To round off the investigations for Bayfordbury, a brief comparison of the formaldehyde VCDs obtained by MAX-DOAS and the satellite instrument TROPOMI is provided.

A detailed description of the TROPOMI HCHO data set, containing daily VCDs averaged within a radius of 20 km around the station, and its handling was already presented in Sect. 5.3. For the comparison, MAX-DOAS HCHO VCDs measured $\pm$30 min around the local equator crossing time of TROPOMI (13:30 LT) and fulfilling the usual filter criteria were averaged on a daily basis. No further filtering was applied for the TROPOMI data. Besides the daily comparison, also a comparison on a monthly scale is performed. For that, monthly median HCHO VCDs were calculated for both data sets from the daily resolved data described above. Further, months with less than three days of overlapping MAX-DOAS and TROPOMI data are excluded from the monthly comparison which affected five months in the whole measurement period.

Figure 10.21 presents the daily (upper part) and monthly (lower part) comparisons, including time series of TROPOMI VCDs (upper left panels), MAX-DOAS VCDs (lower left panels) and the corresponding correlation plots in which the results of simple linear regressions are shown in red. In addition, the mean difference between TROPOMI and MAX-DOAS VCDs (Δ_{mean}) and the standard error of this mean (Δ_{sem}) are provided in the correlation plots.

In general, good agreement between both data sets is found for the daily formaldehyde VCDs yielding an R^2 of 0.51. Especially, under medium HCHO abundances both data sets agree relatively well, while larger absolute deviations are observed at high formaldehyde levels. For the latter, a significant underestimation of the HCHO VCDs by TROPOMI is visible. At low formaldehyde abundances, the TROPOMI results mainly scatter around 0 which is not the case for the MAX-DOAS data for which still background VCDs lar-

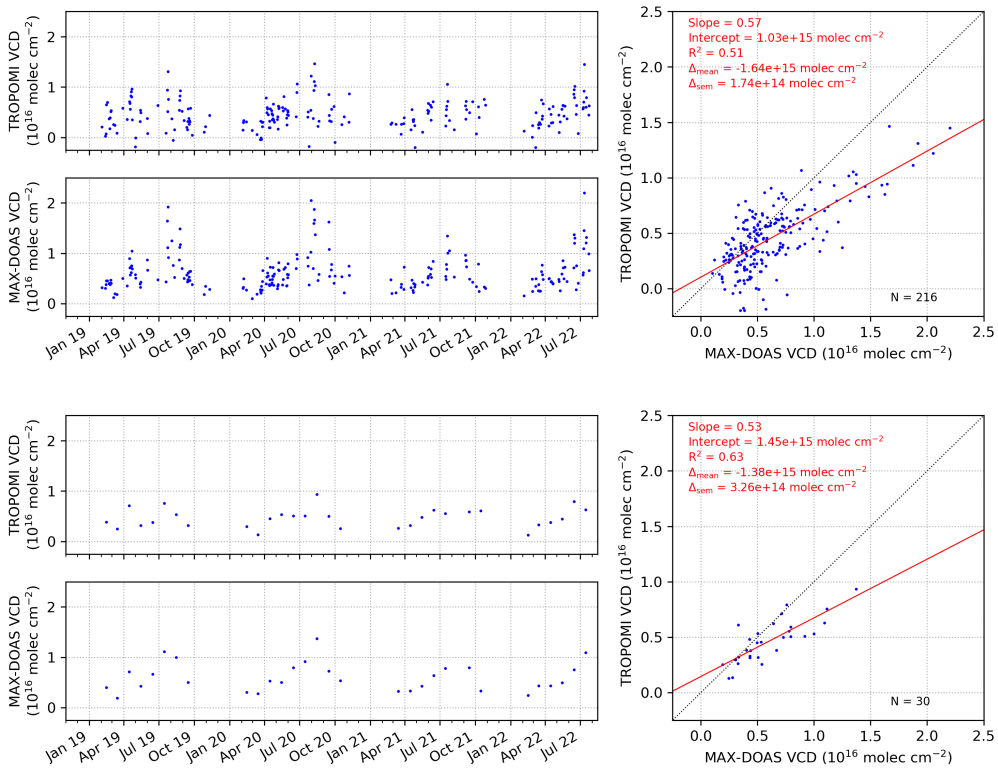


Figure 10.21: Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Bayfordbury on daily (upper part) and monthly basis (lower part). Each part consists of time series of TROPOMI (upper left panels) and MAX-DOAS (lower left panels) VCDs as well as the corresponding correlation plots (right panels). The results of a simple linear regression are shown in red and the number of data points used in the comparisons is indicated in the lower right corners.

ger than 0 are obtained. The mean difference between the two data sets is $(-1.64 \pm 0.17) \times 10^{15}$ molec cm^{-2} corresponding to 27 % of the total mean MAX-DOAS formaldehyde VCD (6.02×10^{15} molec cm^{-2}). The observed deviations can be explained by multiple reasons. Firstly, the sensitivities of both measurement types are different. While MAX-DOAS measurements are very sensitive to absorbers located close to the surface (compare Sect. 3.2), the sensitivity of satellite measurements decreases towards the ground. Here, only a simple comparison was performed and no averaging kernels which would partly account for the different sensitivities were applied. Further, the satellite ground pixels are typically larger than the sensitivity area of the MAX-DOAS measurements. Additionally, the satellite data were averaged 20 km around the MAX-DOAS station. As a consequence, despite the criteria outlined in Sect. 5.3, clouds might affect the TROPOMI measurements. The MAPA flags, however, serve as a rather strict cloud filter as examined before (compare Sect. 9.3) and only rather clear scenes at Bayfordbury are included in the comparison. The spatial averaging of 20 km could have a second effect, namely that a low bias is introduced when areas with lower formaldehyde abundances than at Bayfordbury are sampled.

As expected, the comparison performs better on a monthly basis yielding an R^2 of 0.63 and a slight decrease of the mean difference between both data sets to -1.38×10^{15} molec cm^{-2} corresponding to ca. 23 % of the mean MAX-DOAS HCHO VCD. However, the slope of the linear regression is almost unchanged compared to the daily comparison, still indicating an underestimation of the HCHO abundances by TROPOMI especially at high HCHO VCDs. Apparently, mainly the day-to-day variability is removed by considering monthly data. Looking at the time series exhibits that the TROPOMI measurements do not well reflect the pronounced annual cycle with increased HCHO abundances in summer which is found for the MAX-DOAS data. In addition, no data from November to January are included in the comparison caused by the quality criteria outlined previously (compare Sect. 5.3 and paragraph above). This is also true for the daily comparison where in total only three November days are included. Altogether, both data sets agree reasonably well with the limitation that the TROPOMI measurements systematically underestimate the HCHO VCDs measured by MAX-DOAS especially at elevated HCHO abundances. Possible reasons for the latter were discussed above.

In order to test in how far the comparison results are affected by the selection criteria applied to the data, comparisons were also performed using the MAX-DOAS results obtained for the fixed O_4 SF of 0.8 or other selection criteria. For the latter, the inclusion of MAX-DOAS data measured one hour around the TROPOMI equator crossing time as well as the inclusion of MAPA retrievals flagged with a warning were tested. An overview of the corresponding comparison results on a daily basis is provided in Table 10.1.

Table 10.1: Comparison results between daily TROPOMI and MAX-DOAS data at Bayfordbury for different selection criteria.

SCENARIO	QUANTITY				
	R^2	Slope	Intercept (molec cm^{-2})	Δ_{mean} (molec cm^{-2})	Δ_{sem} (molec cm^{-2})
Standard	0.51	0.57	1.03×10^{15}	-1.64×10^{15}	1.74×10^{14}
SF 0.8	0.38	0.50	1.31×10^{15}	-1.57×10^{15}	1.95×10^{14}
$\pm$60 min	0.45	0.61	6.91×10^{14}	-1.67×10^{15}	1.67×10^{14}
Warnings incl.	0.35	0.52	7.12×10^{14}	-2.08×10^{15}	1.57×10^{14}

The usage of the MAX-DOAS results obtained with the fixed O_4 SF has no large effect on the comparison results and similar mean differences between both data sets are found within their respective uncertainties (Δ_{sem}). However, less data are included for that comparison leading to slightly worse results for the linear regression (R^2 and slope). The usage of MAX-DOAS data one hour around the overpass leads to slightly more data points included in the comparison. Nevertheless, the comparison results are not largely changed. The inclusion of MAPA results flagged with a warning leads to almost twice as many data included in the comparison. However, the scatter is largely increased. While

mainly reasonable MAX-DOAS data are added to the comparison, most additional TROPOMI VCDs are negative or associated with relatively large effective cloud fractions leading to a larger mean difference between TROPOMI and MAX-DOAS HCHO VCDs and slightly worse regression results. This shows that despite the data selection outlined in Sect. 5.3, the exclusive usage of valid MAPA results provides an additional effective cloud filtering of the TROPOMI data. Further, the MAPA flags again prove to be quite strict in terms of cloud filtering (compare Sect. 9.3). In conclusion, the standard comparison approach described above is well justified and used for the further stations.

11

MAINZ, GERMANY

This chapter presents the results of the measurements performed in Mainz (Sect. 4.1), Germany. After an overview of the measurements and the particularities of the analysis in Sect. 11.1, annual and diurnal cycles of the main species HCHO, glyoxal, NO_2 and aerosol are discussed in Sect. 11.2. Dependencies of the abundances of these species on meteorological conditions are studied in Sect. 11.3, while the relation amongst the different species themselves is analysed in Sect. 11.4. The investigation of seasonal vertical profiles of the main species in Sect. 11.5 is followed by a short comparison of the MAX-DOAS formaldehyde results to the ones of TROPOMI in Sect. 11.6. Lastly, interesting HONO results are presented in Sect. 11.7.

11.1 OVERVIEW AND PARTICULARITIES OF THE ANALYSIS

At Mainz, the MAX-DOAS measurements started in June 2013 and were running until March 2022, when major instrumental issues occurred and the measurements stopped. Data from 18 June 2013 to 09 March 2022 are used in this work. This time range covers eight full annual cycles and includes lockdown measures due to the COVID-19 pandemic in 2020 and 2021.

Despite in general good data coverage, several gaps exist in the time series. These are caused by various instrumental issues ranging from simple PC crashes to larger problems with the CCD-camera and the spectrometer themselves. A larger data gap occurred from mid of June to December 2021, where the CCD-camera had to be repaired by the manufacturer. Despite this, most of the time simple restarts were sufficient to continue the measurements and only few cases required greater efforts. The latter sometimes included slight manipulations of the instrument, e.g. re-installation of the shutter, leading to changes of the spectral properties of the instrument. This resulted in six major calibration regimes[1] in which the optical properties of the instrument were stable and for which different spectral calibrations and SLFs were used for the DOAS analysis. An overview of the respective time periods, the resulting spectral ranges and the corresponding spectral resolutions in terms of the FWHM (at 341.5 nm) is provided in Table 11.1.

Given the spectral range covered by the Mainz instrument, the standard fit settings of glyoxal and NO_2 listed in Tables 9.2 and 9.3 had to be adjusted. For all calibration regimes, the DOAS fits for both species were performed using a fit interval from 410 to 457 nm that barely covers the main glyoxal absorption band (compare Fig. 7.18). Brief investigations to check the effect of this modification of the fit settings using data from the the ATTO 80 m instrument for two selected

1 Between 27 March 2018 and 13 May 2018 measurements exist but the calibration was shifted by 7 nm towards longer wavelengths. Since no HCHO fit consistent with the other periods is possible with this setup, this time range was not considered in this work.

Table 11.1: Overview of the spectral characteristics of the Mainz instrument in the six calibration regimes.

PERIOD	PARAMETER	
	Wavelength range	Spectral resolution
13/06/2013 – 31/05/2017	320 – 457 nm	0.72 nm
12/06/2017 – 24/10/2017	321 – 458 nm	0.72 nm
19/11/2017 – 13/03/2018	321 – 458 nm	0.73 nm
14/05/2018 – 14/09/2018	322 – 460 nm	0.72 nm
26/09/2018 – 16/12/2021	323 – 461 nm	0.73 nm
16/12/2021 – 09/03/2022	323 – 460 nm	0.71 nm

days which show significant glyoxal signals were conducted. These studies showed that on average 5 to 10 % smaller glyoxal dSCDs compared to the standard interval are retrieved with the shorter fit interval. Therefore, an underestimation of the glyoxal dSCDs can be expected for the Mainz measurements as well. Since Mainz typically experiences relatively high NO_2 abundances and its strong absorption features are well distributed over the considered spectral range, the effect on the NO_2 fit is expected to be much smaller.

Further, this fit interval led to implications in the calculation of the two PSAs accounting for the Ring effect. Since the Ring effect is caused by inelastic Raman scattering (compare Sect. 3.1.5.1), the calculation of the Ring PSAs requires a certain wavelength margin at the edges of the spectral range for which the PSAs are calculated. For the first three calibration regimes, these wavelength margins were not large enough and caused errors on the edges of the calculated Ring spectra which in turn led to large structures in the DOAS fit. As a solution, the universal FRS of the affected calibration regimes which were used to calculate the Ring PSAs were artificially extended using a literature solar reference spectrum (taking the spectral properties of the Mainz instrument into account). This approach then allowed to calculate Ring PSAs without those boundary effects.

Since no water vapour residual structures were identified, no PSAs accounting for water vapour residual structures as described in Sect. 7.2.1 were included in the glyoxal and NO_2 fits.

For all other species, the standard fit settings were used.

All standard settings as outlined above could be used for the MAPA retrievals for which an elevation bias of 1°, determined according to Sect. 9.2.2, was applied.

For the atmospheric investigations, in principle data from the whole measurement period could be used after applying the standard selection criteria as described in Sect. 9.3. However, during these studies it became evident that the difficulties of the spectral retrieval of glyoxal outlined above lead to very large errors of the glyoxal dSCDs. As a consequence, also the VCDs and near-surface concentrations obtained with MAPA are hardly interpretable. Thus, it was de-

cided to only present annual and diurnal cycles of glyoxal for which these errors mainly average out, but to exclude glyoxal from the further atmospheric investigations for Mainz. As for glyoxal, also for NO_2 the shorter fit window described above was used for the analysis of the Mainz measurements. However, no strong effects on the NO_2 results were observed.

Finally, it should be mentioned that only a rather small fraction of the MAPA retrievals yielded valid results for HCHO (compare Table 9.8). In March 2019, the number of individual scans per spectrum was increased leading to less noise which in turn leads to a larger fraction of valid HCHO retrievals. This becomes also evident in the annual overview plots presented in the next section.

11.2 ANNUAL AND DIURNAL CYCLES

The annual and diurnal cycles of formaldehyde, glyoxal, NO_2 and aerosol results at Mainz are discussed in this section. For that, the MAPA results were first filtered using the criteria described in Sect. 9.3 and then visualised on annual and diurnal bases as explained in detail for Bayfordbury in Sect. 10.2.

The annual cycles of HCHO VCDs and near-surface concentrations are presented in Fig. 11.1, where the two uppers panels show the full time series and the two lower panels present monthly box-whisker-plots.

Both VCDs and surface concentrations show pronounced annual cycles with enhanced HCHO abundances during the summer months and lowest values in the winter periods. While this cycle is very clear for the VCDs, more variability is observed for the surface concentrations. The differences of the absolute summer values between the years are related to the highest temperatures that are reached in the different years. Highest temperatures at Mainz during the measurement period occurred in the summers of 2015 and 2019 which are also associated with highest HCHO VCDs clearly exceeding 3×10^{16} molec cm^{-2}. The observed temperature dependence is slightly weaker pronounced for the surface concentrations, but still visible. Further details of temperature dependencies are discussed in Sect. 11.3.

The statistical representation of the annual cycles of HCHO VCDs (left) and surface concentrations (right) depicted in the two lower panels of Fig. 11.1 confirms the observed annual variations. Lowest HCHO VCDs occur in January with a monthly mean of 0.45×10^{16} molec cm^{-2}, while the monthly mean surface concentrations in the months from November to May are comparable and vary slightly around 0.5×10^{11} molec cm^{-3}. Starting at their minimum values, the VCDs increase in the further course of the year and reach their maximum in June and July, where monthly means of around 1.2×10^{16} molec cm^{-2} are obtained. The surface concentrations show a slightly different pattern as the monthly means stay at their minimum values during spring and show an abrupt increase from May to June towards their maximum values of ca. 0.9×10^{11} molec cm^{-3} in summer. These differences in the patterns of both quantities possibly indicate that the increased VCDs in spring are caused by non-local HCHO sources. As already seen for Bayfordbury, the range of HCHO abundances is larger in summer compared to the rest of the year (compare whiskers). After reaching their respective max-

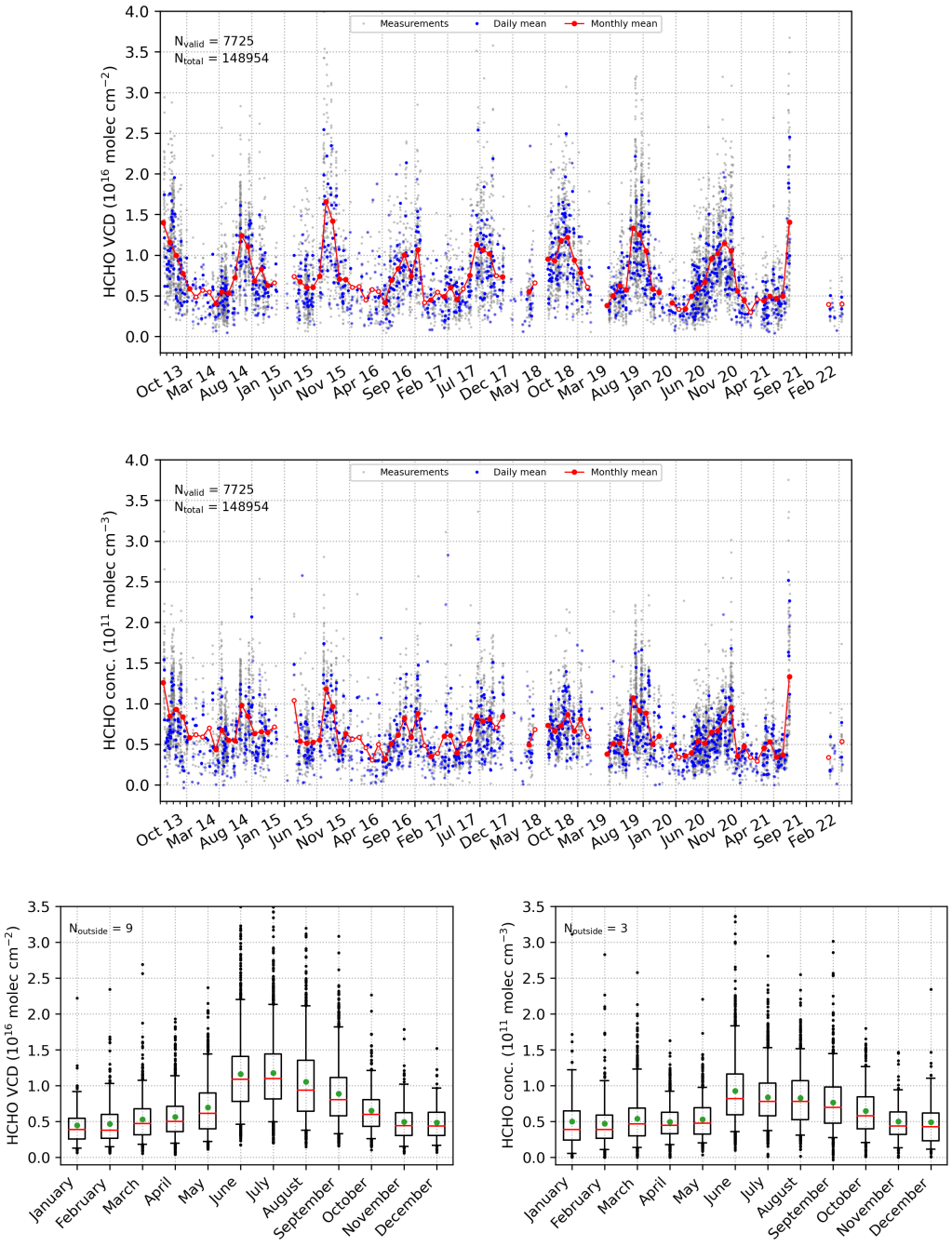


Figure 11.1: The two upper panels show time series of HCHO VCDs and surface concentrations for Mainz. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of HCHO VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

ima, both quantities show rather smooth decreases in late summer and autumn reaching values comparable to their respective annual minima in November. The amplitude of the observed annual cycles is larger for the VCDs which might again indicate that the bulk of the HCHO measured at Mainz during summer is not emitted locally. A further effect might be the rise of the boundary layer leading to a distribution of the HCHO over a larger column in summer. Lastly, it should be mentioned that compared to Bayfordbury the annual minimum and maximum values of both quantities are larger at Mainz. Further, the period in which the HCHO abundances are highest starts earlier and lasts longer. While the findings for summer are most likely related to longer and on average warmer summers at Mainz, the higher winter values probably show the influence of direct HCHO emissions caused by the closeness of the city to the location of the instrument at Mainz.

Like for Bayfordbury, no significant differences in the annual characteristics of the HCHO results are observed if a fixed SF of 0.8 is used. Only marginally larger values are obtained for both quantities using the fixed SF, whereby the effect is slightly stronger for the surface concentrations, again indicating slight differences in the profile shapes obtained with the two different SFs (compare Sect. 11.5).

The diurnal cycles of HCHO VCDs (left) and surface concentrations (right) are shown in Fig. 11.2, where different colours indicate results obtained for the four seasons.

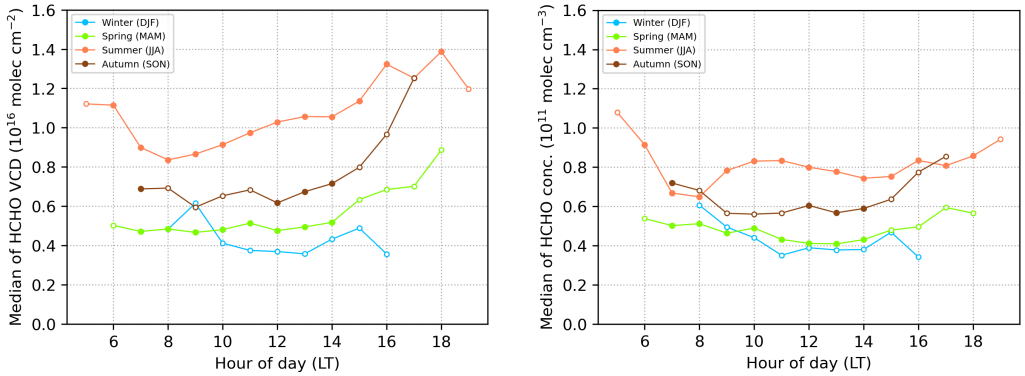


Figure 11.2: Diurnal cycles of HCHO VCDs (left) and surface concentrations (right) for Mainz for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

For the surface concentrations, none of the four seasons shows a significant diurnal cycle. The enhanced concentrations towards morning and evening in autumn can be explained by the dominance of September measurements in these bins as explained in detail for Bayfordbury. A similar effect might cause the high HCHO concentrations during the early morning hours in summer where June values dominate which are higher than the ones in late summer (August). Also for the VCDs, no strong diurnal cycle is observed in winter, where only slight enhancements at 9 a.m. and 3 p.m. occur which might indicate rush hour effects. However, in winter all time bins are flagged for having less than 5 % valid MAPA retrievals and therefore the results have to be considered with care.

Contrary to that, rather clear diurnal variations of the HCHO VCDs are found in spring, summer and autumn. Here, constant increases during the course of the day are observed. Again statistical effects might influence the values in the early mornings and late evenings. For instance, in summer measurements performed in June dominate where the days are longest and the HCHO abundances higher compared to the ones found in late summer, i.e. August (for detailed explanations of these statistical effects, refer to Sect. 10.2). However, even if these periods of the day are ignored, strong and systematic diurnal cycles remain. The fact that, despite the clear diurnal cycles for the VCDs, no diurnal cycles are found for the surface concentrations might have two major reasons. First, this probably indicates the dominance of secondary formation of HCHO with precursor species being transported upwards and forming HCHO not close to the surface but at higher altitudes. Second, the development of the boundary layer height in the course of the day probably increases the VCDs, while the near-surface concentrations stay rather constant since the HCHO is distributed over larger parts of the atmospheric column.

Annual cycles of glyoxal VCDs and surface concentrations are presented in Fig. 11.3.

Unlike formaldehyde and contrary to Bayfordbury, no consistent annual patterns are observed for glyoxal at Mainz. Both quantities mainly show huge scatter and a suspicious decreasing trend starting in June 2017, when first instrumental issues required maintenance work. The latter is particularly pronounced for the surface concentrations and is also observed in the glyoxal dSCDs which are directly obtained from the DOAS fit. Despite several investigations, the reason for this decreasing trend could not be clearly identified. However, no irregularities in the fitted intensity offsets nor the spectral shifts of the DOAS analysis were found. Since the decreasing trend only started in June 2017 and several strange findings, e.g. sudden jumps, occur in the glyoxal results afterwards, it was decided to only include glyoxal data until 31 May 2017 for the further atmospheric studies.

This limited data set was used for the generation of the monthly box-whisker-plots which are displayed in the two lower panels of Fig. 11.3 for VCDs (left) and near-surface concentrations (right). Still, both quantities do not show conclusive results given their very large uncertainties (compare whiskers).

Also the diurnal cycles of both quantities which are presented in Fig. 11.4 were generated using the limited data set. No clear diurnal cycles can be observed for the VCDs, whereas overall decreasing trends are found for the surface concentrations in all seasons. In principle, this could be explained by a constant VCD throughout the day in combination with an increasing boundary layer height. However, in view of the difficulties of the spectral retrieval described above and the inconclusive results for the annual cycles it was decided to exclude glyoxal from the further atmospheric investigations for Mainz. Lastly, it should be mentioned that also the analysis with a fixed O_4 SF of 0.8 did not yield better results.

Although the same fit interval is used, the NO_2 results are of higher quality. The annual cycles of NO_2 VCDs and near-surface concentrations are summarised in Fig. 11.5.

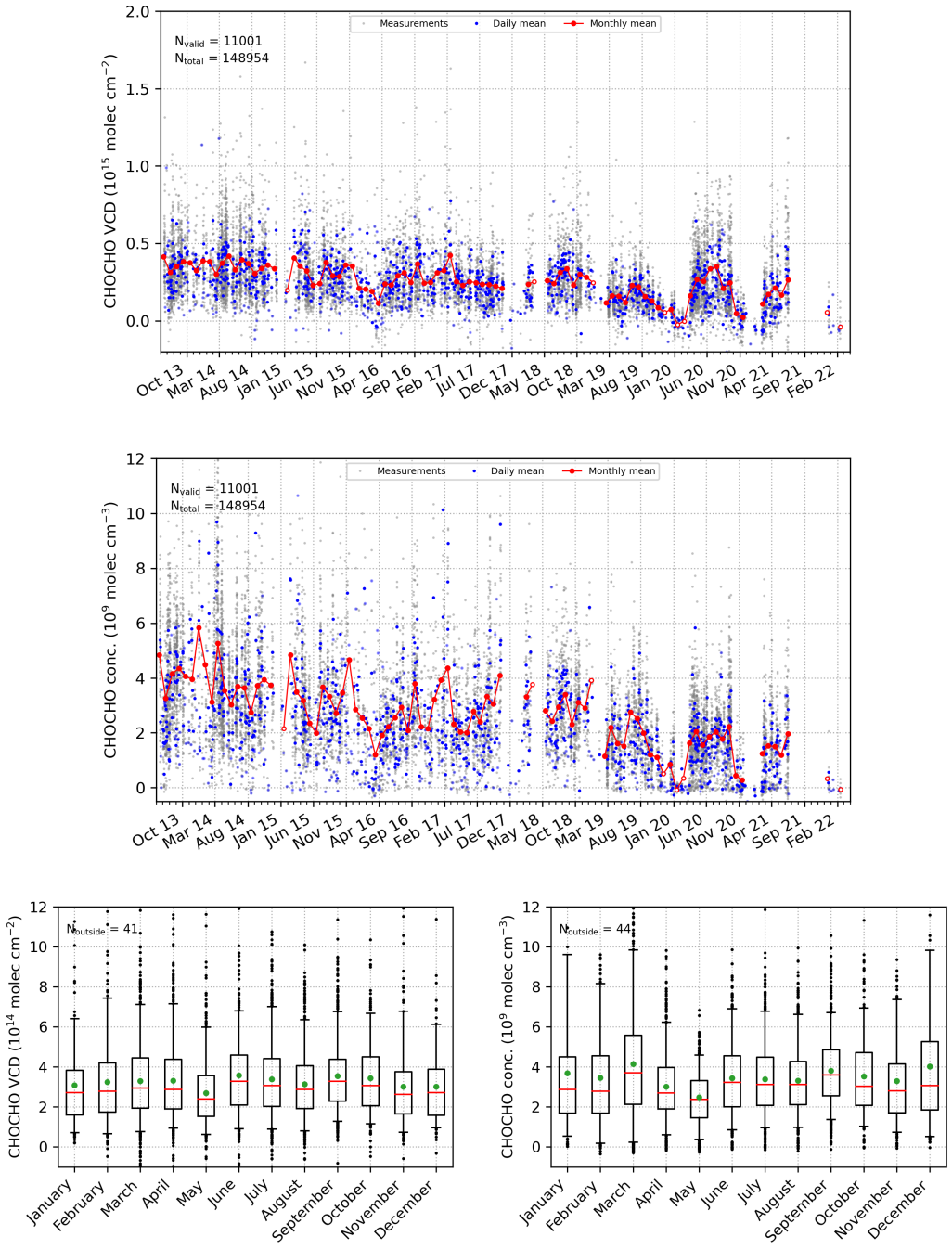


Figure 11.3: The two upper panels show time series of glyoxal VCDs and surface concentrations for Mainz. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of glyoxal VCDs (left) and surface concentrations (right), where only data until 31 May 2017 were included. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

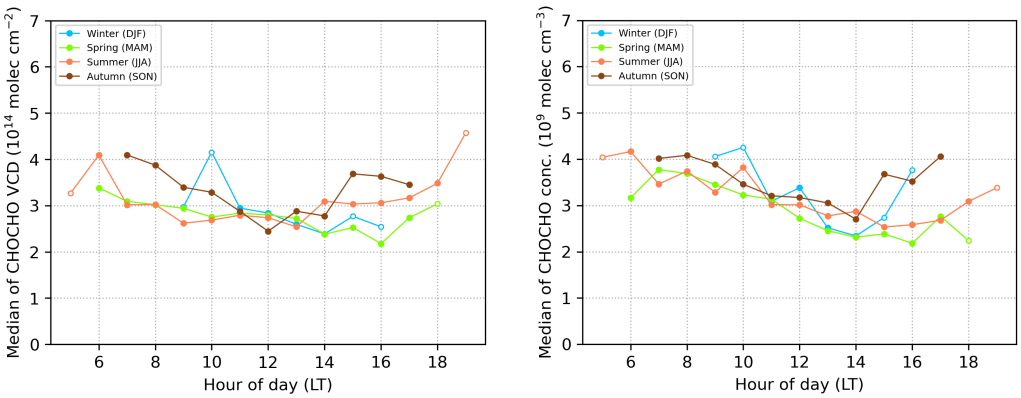


Figure 11.4: Diurnal cycles of glyoxal VCDs (left) and surface concentrations (right) for Mainz for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season. Note that only glyoxal data until 31 May 2017 were considered for this plot.

Given its location at the western edge of the densely populated Rhine-Main region, high NO_2 pollution levels at Mainz occur throughout the year. Interestingly, Mainz shows the highest NO_2 abundances of all stations included in this work, especially for the surface concentrations. Nevertheless, similar annual cycles with highest and lowest NO_2 abundances during winter and summer, respectively, are observed for both VCDs and surface concentrations. In winter, daily mean VCDs and surface concentrations regularly exceed values of 2×10^{16} molec cm^{-2} and 3.5×10^{11} molec cm^{-3}, respectively. During summer, daily means of 0.5×10^{16} molec cm^{-2} and 0.5×10^{11} molec cm^{-3} or even higher occur frequently. Further, a high day-to-day variability of the NO_2 abundances which is similar for both quantities is observed indicating the importance of the origin of the air masses probed by the Mainz instrument. Details on the effect of the wind characteristics on the measured trace gas abundances are discussed in Sect. 11.3. Similar to Bayfordbury, a weak decrease of the NO_2 VCDs and surface concentrations can be observed over the entire measurement period (compare monthly means in the two upper panels). Again this trend masks a possible reduction of the NO_2 abundances during the COVID-19 lockdown measures in spring 2020 where no significant effect can be observed.

The outlined annual cycles with highest NO_2 VCDs and near-surface concentrations in winter and minimum NO_2 pollution levels in summer are also confirmed by the statistical representations of the annual patterns of both quantities in the two lowermost panels of Fig. 11.5. Also the overall higher NO_2 abundances compared to Bayfordbury (compare Fig. 10.5) become evident in these plots. While the highest NO_2 VCDs and surface concentrations at Bayfordbury are observed between November and February, the period of enhanced NO_2 is longer in Mainz covering the months from October to March.

The same general patterns of the NO_2 results are yielded with the fixed O_4 SF of 0.8. For the VCDs, essentially no differences between both SF scenarios are found while lower surface concentrations are obtained with the variable SF.

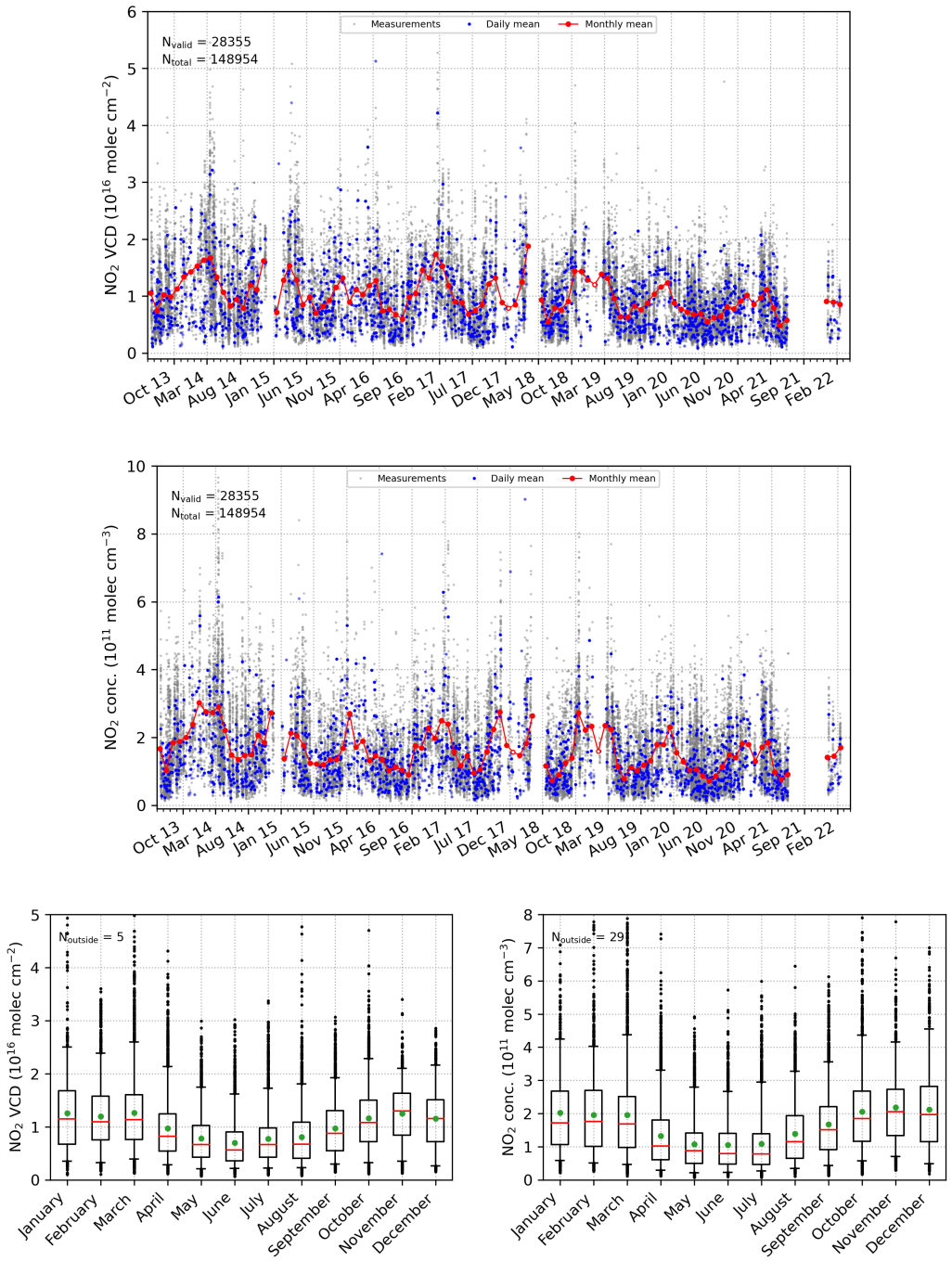


Figure 11.5: The two upper panels show time series of NO_2 VCDs and surface concentrations for Mainz. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of NO_2 VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

Median diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for the four season are presented in Fig. 11.6.

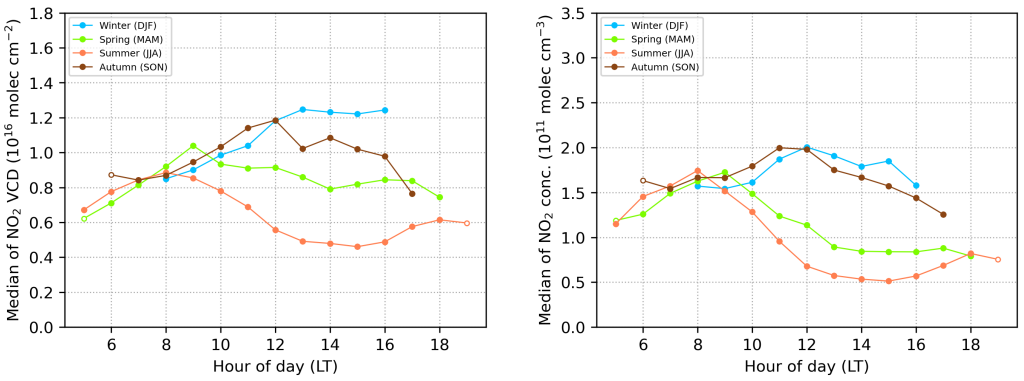


Figure 11.6: Diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for Mainz for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

Both quantities show similar diurnal cycles in spring and summer, whereby the amplitude of the observed variations is higher for the surface concentrations. Further, for the VCDs the diurnal cycle is less prominent in spring as compared to summer. Nevertheless, in both seasons an increase of the measured NO_2 abundances is found during morning reaching their maxima in the late morning. After that, VCDs and surface concentrations decrease and reach their minimum values in early afternoon. Towards sunset, a slight increase of the observed NO_2 abundances is found. It should be mentioned that contrary to the observations made for Bayfordbury the maxima and minima of both quantities occur simultaneously (within a season). This indicates the immediate proximity of the city to the measurement site in Mainz and that variations in the emissions, e.g. rush hour effects, are simultaneously seen in both quantities. Thus, the influence of transport processes on the diurnal variations of NO_2 can on average be neglected. The fact that the diurnal patterns are weaker pronounced for the VCDs can be explained in a similar way as for Bayfordbury and basically indicates an effect of the rising boundary layer in the course of the day and upward mixing of the emitted NO_2. For further details, refer to Sect. 10.2.

Consistent to the observed annual cycles, higher NO_2 abundances are observed in autumn and winter which show similar NO_2 levels and diurnal variations. In these two seasons, both quantities increase during the morning, peak around noon and decrease over the course of the afternoon. For the VCDs, however, the decrease during afternoon is not visible in winter, where afternoon values stay at the level of the noon time peak. Obviously, the diurnal cycles in autumn and winter significantly differ from the diurnal cycles observed in spring and summer. The reasons for these differences are not fully clear but possibly indicate different characteristics of emissions, vertical mixing processes and tropospheric chemistry during the respective parts of the year. At Bayfordbury, only winter showed different diurnal cycles. This difference to the findings for Mainz might

be related to differences in the meteorological conditions since winters are more moderate at Bayfordbury.

Next, annual cycles of AOD and near-surface aerosol extinction at 367 nm are presented in Fig. 11.7 which also includes AERONET AODs for comparison.

While essentially no annual cycle can be seen in the AOD, highest near-surface aerosol extinctions typically occur in late autumn, winter and early spring. For the AODs, reasonable agreement to the AERONET measurements is found, whereby lowest and highest AODs are not perfectly matched by the MAX-DOAS retrieval as already seen for Bayfordbury. The agreement between both data sets is slightly worse at Mainz compared to Bayfordbury. One possible reason for this might be strong horizontal gradients in the aerosol distribution close to the city as suggested by studies including all viewing directions of the used 4-azimuth MAX-DOAS instrument (Remmers et al., 2019). As both instruments do not necessarily look in the same azimuthal direction, different air masses might be probed by the MAX-DOAS instrument and the AERONET sun-photometer, respectively.

The monthly box-whisker-plots of AOD and near-surface aerosol extinction confirm the previous findings. Almost no annual cycle is found for the AOD and monthly means close to 0.2 or slightly below are obtained throughout the year. Only during winter faintly lower monthly mean values of around 0.16 occur which might also be caused by statistical effects as less valid MAPA retrievals are obtained during winter. Also the variability, indicated by the whiskers, stays constant during the entire year. The near-surface aerosol extinction, on the other hand, shows a weak annual cycle with highest monthly means of approximately $0.22\,\mathrm{km}^{-1}$ in the months from November to March. Lower values occur during summer, where the monthly means are only ca. $0.12\,\mathrm{km}^{-1}$. Also the variability of the obtained values is larger during the colder part of the year for the aerosol extinctions. The observed discrepancy of the annual patterns of AODs and near-surface extinctions can be explained by an on average higher boundary layer height during the warmer parts of the year. Since the total AOD remains mainly constant throughout the whole year, the layer into which the aerosol is mixed is thicker during summer leading to slightly lower near-surface aerosol extinctions. Similarly to Bayfordbury, the large spread of the near-surface extinctions in late autumn, winter and early spring is probably caused by the frequent occurrence of cloudy sky conditions. For details, refer again to Sect. 10.2.

Consistent to the aerosol results for Bayfordbury, significantly higher values are obtained for both AODs and near-surface extinctions if the fixed O_4 SF of 0.8 is applied (for trace gases, only the surface concentrations are significantly affected). However, it should be stressed that the general findings do not depend on the choice of the SF and that more valid MAPA retrievals are obtained with the variable SF. Lastly, also for Mainz the scatter of the results obtained with the fixed SF is larger.

Figure 11.8 depicts the median diurnal cycles of AOD (left) and near-surface aerosol extinction (right) in the different seasons.

For the AOD, no consistent diurnal cycle occurs in winter, while overall slight increases of the retrieved AODs over the course of the day are observed for the remaining seasons. Note that again the time bins in the early mornings and

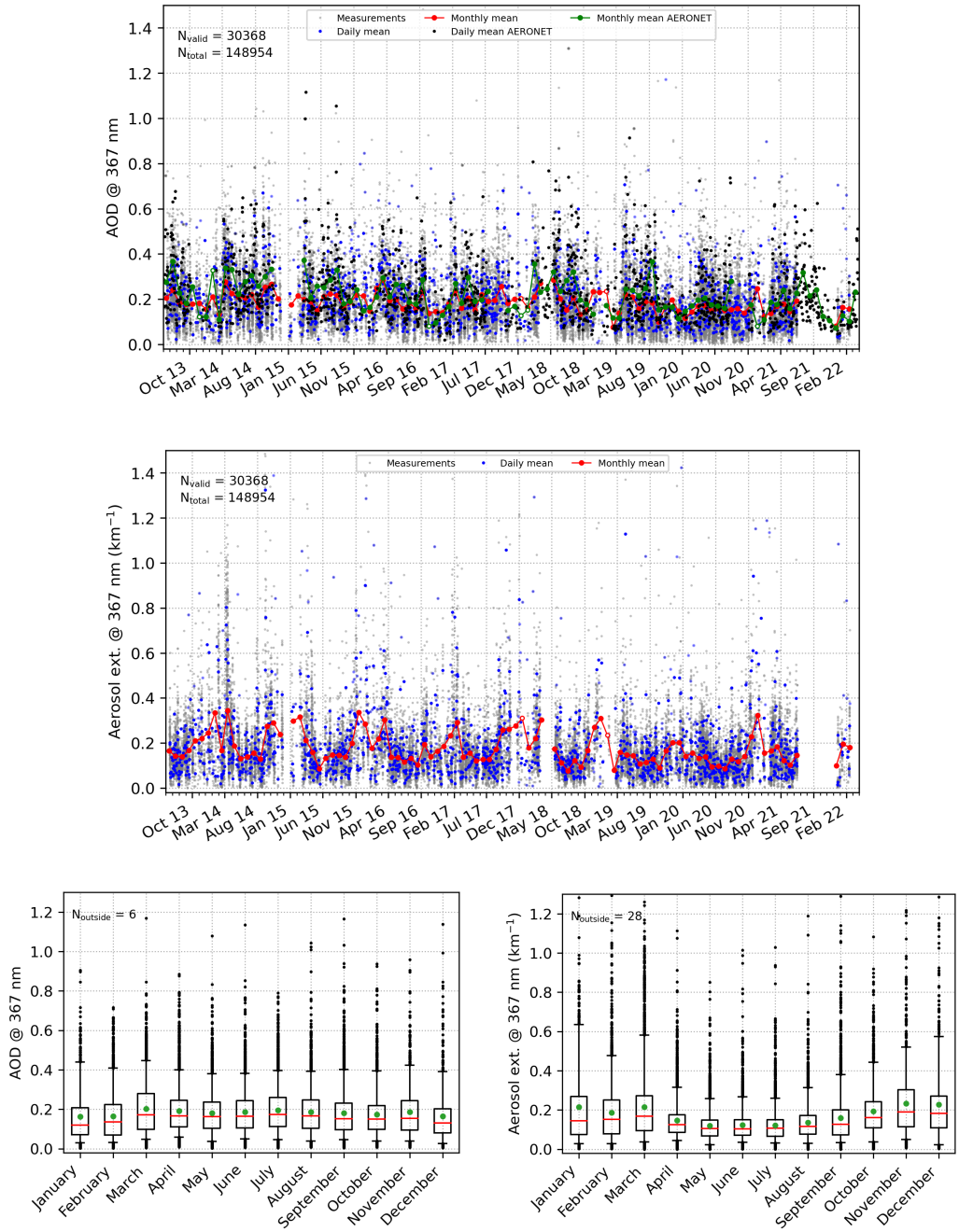


Figure 11.7: The two upper panels show time series of AOD and surface aerosol extinction at 367 nm for Mainz. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and surface aerosol extinction at 367 nm (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

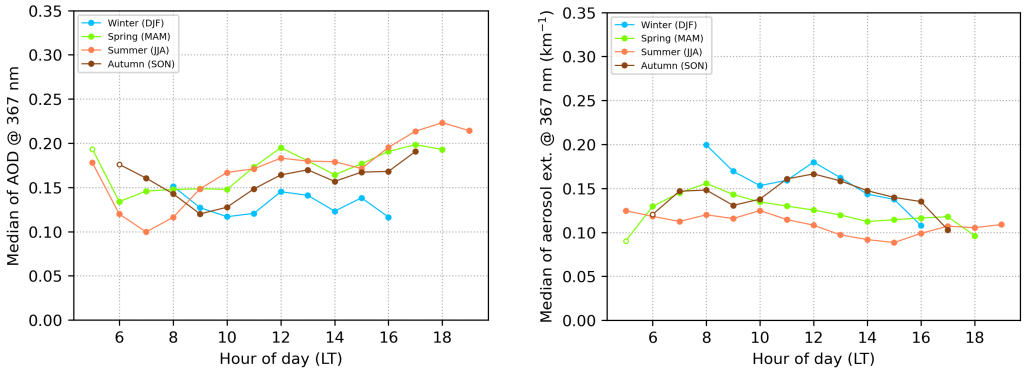


Figure 11.8: Diurnal cycles of AOD (left) and surface aerosol extinction (right) at 367 nm for Mainz for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

late evenings can be affected by statistical artefacts as outlined above which is valid for both quantities. Contrary to the findings for the AODs, the near-surface aerosol extinctions show overall decreasing trends in all seasons. For autumn, additionally, a slight increase in the morning towards noon is visible. For spring and summer, the findings are consistent to Bayfordbury and can most likely be explained by the development of the boundary layer leading to higher reaching profiles during the course of the day.

Besides the species discussed in detail in this section, also HONO and water vapour retrievals were performed. An overview of the main findings for HONO is provided in Sect. 11.7. Similar to Bayfordbury, expected results were obtained for water vapour which exhibits a strong annual cycle with high values in summer and low values in the (absolutely seen) dryer winter months. Finally, almost no diurnal cycles are found for H_2O.

11.3 INFLUENCE OF METEOROLOGICAL CONDITIONS

Next, the dependence of the trace gas and aerosol abundances at Mainz on air temperature, wind speed and direction will be discussed. Before the data were visualised, they were filtered applying the criteria described in Sect. 9.3 and hourly means were calculated.

The two uppermost panels of Fig. 11.9 show hourly mean HCHO VCDs (left) and near-surface concentrations (right) as function of temperature.

For both quantities, a strong highly non-linear temperature dependence showing increasing HCHO abundances with increasing temperatures is found which is consistent to the annual cycles that showed highest HCHO during summer. In general, the observed temperature dependence is very similar to the one found for Bayfordbury. For instance, three temperature zones with similar characteristics can be defined, namely up to ca. 10 °C, 10 to 20 °C and above 20 °C, which show almost no, moderate and strong temperature dependencies, respectively. Given the longer measurement period and the continental influence on the climate at Mainz,

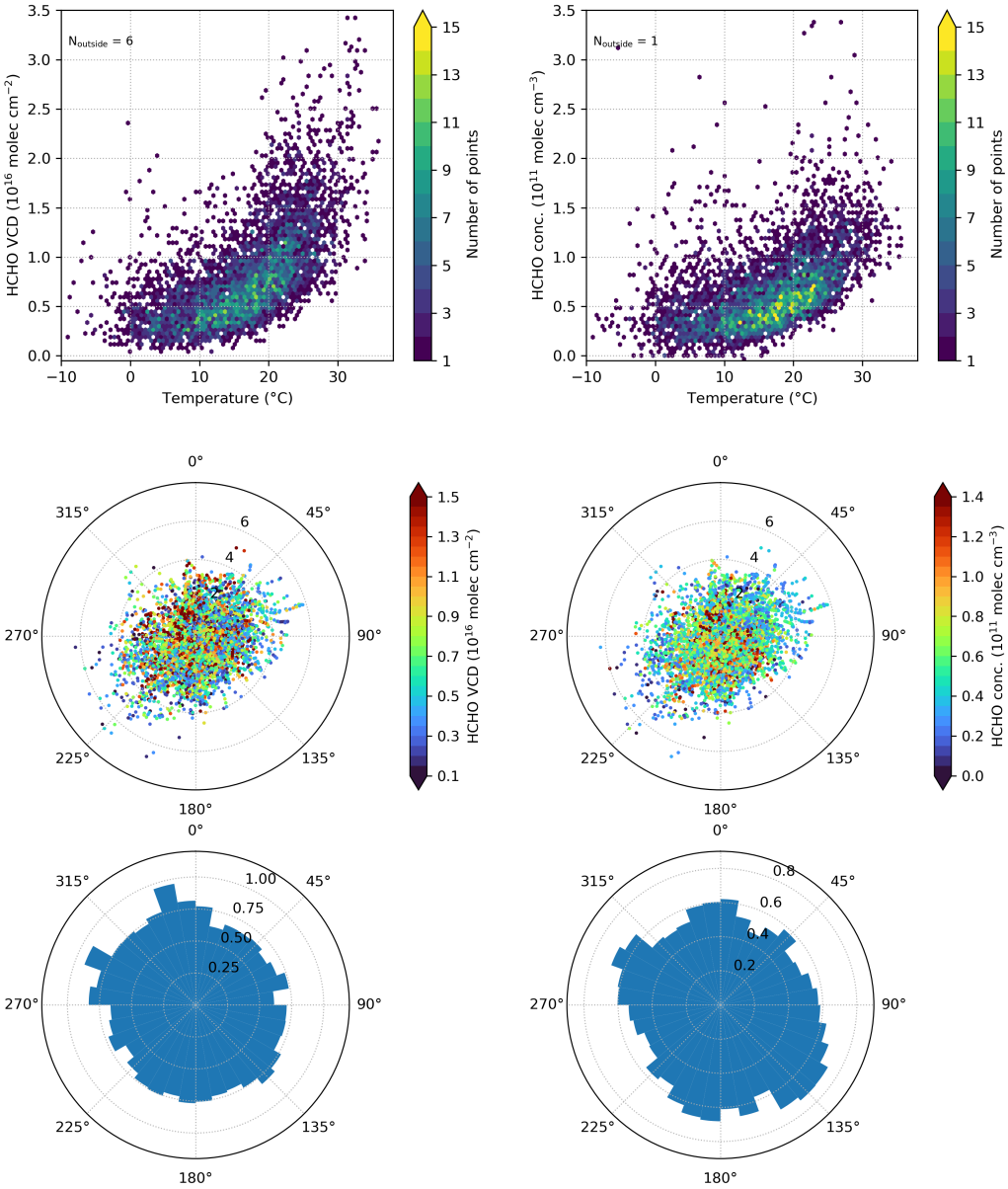


Figure 11.9: The two uppermost panels show hourly mean HCHO VCDs (left) and surface concentrations (right) measured at Mainz as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HCHO VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HCHO VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

more measurements in the temperature range above 20 °C are available compared to Bayfordbury. Altogether, the observed strong temperature dependence once more stresses the importance of secondary formation of HCHO from mostly biogenic precursor species. Lastly, it should be mentioned that for Mainz small differences in the temperature dependencies of VCDs and surface concentrations

are found. In particular, the non-linearity of the temperature dependence seems to be weaker for the surface concentrations. However, no conclusive explanation for this finding can be given so far.

The middle panels of Fig. 11.9 present hourly mean HCHO VCDs (left) and near-surface concentrations (right) measured at Mainz as functions of wind speed and direction, while the lower panels depict general angular distributions in terms of median values in bins of 10°. For a detailed description of these plots, refer to Sect. 10.3.

Compared to Bayfordbury, the angular distributions of the HCHO abundances at Mainz are more homogenous (compare lower panels) and only slight enhancements are observed under some wind directions. This finding probably indicates that at Mainz the air temperature has a stronger influence on the occurring HCHO abundances as compared to the wind direction. Further, possible anthropogenic sources such as the city of Mainz lie in the direct vicinity of the instrument unlike to Bayfordbury where anthropogenic sources are found at larger distances from the measurement site.

In general, enhanced HCHO abundances are measured under westerly to northerly (sector from ca. 270 to 360°) and easterly to southerly (sector from 90 to 210°) winds. In these sectors, median VCDs and surface concentrations are around or slightly above 0.75×10^{16} molec cm^{-2} and 0.6×10^{11} molec cm^{-3}, respectively. In the remaining sectors, abundances slightly, yet notably, below these values are found. Overall, similar angular patterns are found for both quantities again indicating the importance of secondary formation on the HCHO abundances. However, differences in the details of the angular distributions of both quantities can be observed.

For the VCDs, the observed enhancements are rather weak and more pronounced in the north-westerly wind sectors, while only marginally enhanced values are obtained in the south-easterly sectors. Contrary to that, the influence of the wind direction is generally more prominent for the surface concentrations. These show largest enhancements in the south-easterly wind sectors, while slightly smaller enhancements are found in the north-westerly sectors. These differences between both quantities suggest an additional effect of local emissions of both HCHO itself and anthropogenic precursors which influence the near-surface concentrations stronger. A similar effect was observed at Bayfordbury (compare Sect. 10.3). Nevertheless, it should be stressed that the effects of different wind directions are generally smaller at Mainz compared to Bayfordbury which probably indicates that mostly biogenic emissions dominate the HCHO abundances, especially during summer. Lastly, typically higher HCHO abundances occur at low wind speeds which is stronger pronounced for the surface concentrations (compare middle panels).

Given the observation that at Mainz the temperature dependence, and thus biogenic emissions, mainly determines the HCHO abundances and the rather homogenous angular distribution of the measured HCHO, the observed slight enhancements under specific wind directions can be explained by additional sources located in these sectors. For the following discussions, it is important to

note that in this work data from the viewing direction towards the city centre of Mainz at an azimuth angle of 51° were used.

Under north-westerly winds, the air masses probed by the instrument are affected by the city of Wiesbaden leading to enhancements of the HCHO VCDs and near-surface concentrations in this sector. Further, both quantities show local maxima at around 295 and 345° which are probably related to industrial facilities including some chemical factories and a cement plant situated along the Rhine river north-west of the centre of Mainz. These sources located within a few kilometres to the instrument particularly lead to an enhancement of the near-surface concentrations. The latter are also clearly elevated under south-easterly winds, where air masses are mostly influenced by anthropogenic emissions of Mainz itself (compare angular distribution of NO_2 in Fig. 11.10) and another cement factory at ca. 6 km from the instrument. Further, some medium-sized towns lie in this wind sector. The sectors more to the east which show marginal enhancements for the surface concentrations (almost none for the VCDs) might be affected by Frankfurt airport and the Rhine-Main region in general. The increase of the VCDs is generally weaker pronounced probably indicating that the total HCHO column is less affected by local emissions, similar as already found for Bayfordbury (compare Fig. 10.9). As a consequence, also sources located farther away from the measurement site might play a role for the VCDs. For the north-westerly wind sector, which showed the largest enhancements of the VCDs, the Taunus mountain range which is largely covered by trees is most likely an important source of biogenic VOCs acting as HCHO precursors. Furthermore, an effect of the densely populated Ruhr region cannot be completely ruled out. However, no such effect of the Rhine-Main region on the HCHO VCDs is observed.

In the remaining sectors, generally less anthropogenic activities are present and no obvious distinct sources of biogenic VOCs can be found. Thus, HCHO abundances are mainly dominated by background levels which are varying throughout the year. As outlined in detail for Bayfordbury, the described dependencies on the wind direction might implicitly also contain other dependencies. For instance, westerly winds are typically associated with higher wind speeds, more clouds and rain and hence possibly lower temperatures and less solar radiation as compared to the weather conditions under other wind directions.

Figure 11.10 summarises the dependence of the measured NO_2 abundances at Mainz on temperature (upper panels) as well as wind speed and direction (middle and lowermost panels).

No clear temperature dependence is found for both NO_2 VCDs and surface concentrations. Nevertheless, both quantities show a slight tendency towards lower NO_2 abundances at higher temperatures. These findings are consistent with the observed annual cycles with on average lower NO_2 pollution levels during summer compared to winter, while single enhancements can occur throughout the entire year.

Similar to Bayfordbury, the large day-to-day variability of the NO_2 abundances can be mainly attributed to changing wind directions and thus different source regions of the probed air masses. The dependence of NO_2 VCDs (left) and surface

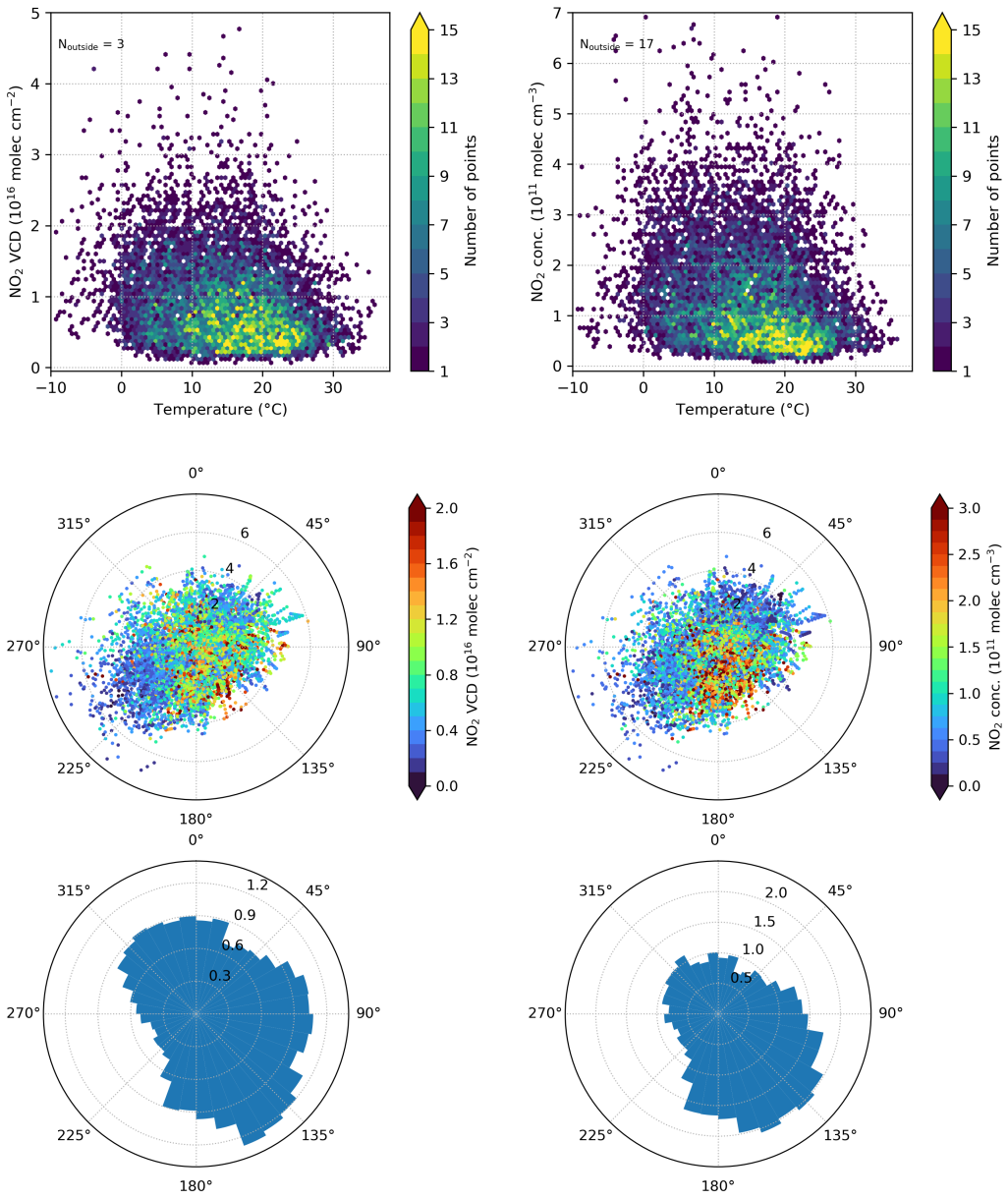


Figure 11.10: The two uppermost panels show hourly mean NO_2 VCDs (left) and surface concentrations (right) measured at Mainz as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean NO_2 VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median NO_2 VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

concentrations (right) measured by the Mainz instrument on wind speed and direction is presented by the middle and lower panels of Fig. 11.10.

While the VCDs show high median values in a large sector from north-west to south-south-west, three more confined sectors exist in which the NO_2 VCDs are particularly high. These sectors range approximately from 290 to 20°, 60 to 110°

and 120 to 200°, where the air masses probed under such wind directions are mainly affected by the city of Wiesbaden, the Rhine-Main region together with Frankfurt airport and Mainz itself, respectively. The sector showing enhanced near-surface concentrations is more confined ranging from east-north-east to south-south-west indicating that emissions from sources located at larger distances from the measurement site affect the surface concentrations less. For both quantities, lowest NO_2 abundances occur under south-westerly to westerly winds. This sector is mainly rural with no significant anthropogenic sources. Nevertheless, it should be noted that especially the median surface concentrations of ca. 0.75×10^{11} molec cm^{-3} in this sector are still quite high (close to the elevated levels occurring at Bayfordbury).

Both quantities show their absolute maxima under south-easterly winds (sector from 120 to 170°), where median values of around 1.2×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3} are obtained for VCDs and near-surface concentrations, respectively. Air masses probed under winds from this sector are mainly affected by Mainz and its mostly traffic related emissions. These relatively "fresh" emissions lead to a strong simultaneous enhancement of both quantities, while the relative increase is larger for the surface concentrations. For the VCDs, also an influence of long-range transport of pollution originating from Ludwigshafen and Mannheim, both located roughly 55 km to the south (ca. 165°), cannot be ruled out. Sectors more to the east, especially from 60 to 110°, are influenced by the Rhine-Main region with its cities and towns including the city of Frankfurt and its airport. The surface concentrations show clearly weaker enhancements in this sector, where the values drop below 1.5×10^{11} molec cm^{-3} towards easterly and north-easterly wind directions. The VCDs, on the other hand, are strongly elevated in this sector and still show values of around 1×10^{16} molec cm^{-2}. Apart from that, the VCDs are also enhanced under north-westerly to northerly winds, whereas the surface concentrations show only marginally higher values compared to the south-westerly wind sector where their absolute minimum occurs. This northerly sector is mainly influenced by Wiesbaden and industries along the Rhine river as pointed out above.

The differences in the angular patterns of VCDs and surface concentrations under north-westerly to northerly and easterly winds are consistent to similar observations made for the Bayfordbury measurements and can be explained by a combination of several effects which have different importances under different wind directions. First, near-surface concentrations are primarily determined by local emissions, whereas sources located at larger distances mainly lead to enhancements of the VCDs as mixing processes distribute the emitted NO_2 within the boundary layer. Additionally, parts of the NO_2 might already be degraded. Further, also the emission height might play a role. In the northerly wind sector, a power plant is located at the Rhine river whose emissions are probably mixed towards higher altitudes very quickly. Under easterly winds, a major emission source is Frankfurt airport at a distance of ca. 25 km. Starting aircraft ascend very quickly and emit a large fraction of their NO_2 over a large altitude range leading mainly to an increase of the VCDs.

Again, also the dependence of wind speed on wind direction probably affects the angular distribution of NO_2. Most important might be higher wind speeds under westerly winds which are in general associated with lower NO_2 abundances as becomes evident by the middle panels of Fig. 11.10.

Lastly, the influence of meteorological conditions on the aerosol quantities is investigated. The upper panels of Fig. 11.11 show AOD (left) and near-surface aerosol extinction (right) at 367 nm as function of temperature.

Neither of the two aerosol quantities shows a distinct temperature dependence. Nevertheless, the near-surface extinctions show a slight tendency of higher values at lower temperatures with the majority of the highest near-surface extinctions occurring at temperatures below ca. 20 °C. For the AODs, the bulk of values around 0.2 occurs at higher temperatures. These two findings fit the observed annual cycles with faintly lower monthly mean AODs in winter and highest near-surface aerosol extinctions occurring from November to March.

The middle and lower panels of Fig. 11.11 summarise the dependence of AOD (left) and near-surface aerosol extinction (right) at 367 nm on wind speed and direction.

Similar to HCHO, the observed angular patterns are rather homogenous for both aerosol quantities and only relatively small variations in the angular distributions are found. Highest AODs are found under westerly to northerly winds (ca. 270 to 20°), where median values up to approximately 0.2 are obtained. No clear enhancements are found in the remaining sectors from north-east to south in which median AODs are quite constant and typically stay below 0.2. Lowest AODs (around 0.15) are observed under southerly to south-westerly winds (ca. 170 to 270°). The near-surface aerosol extinctions show a different angular distribution, having their maximum values under east-south-easterly to southerly winds (ca. 100 to 200°) which lead to near-surface extinctions of around 0.15 km^{-1}. Rather similar extinctions are found in the remaining sectors where the values scatter around 0.12 km^{-1}. Also the near-surface extinctions show their lowest median values under south-westerly winds, where no distinct aerosol sources are located. Almost no effect of the wind speed can be observed (compare middle panels of Fig. 11.11) for both quantities. Overall, the angular patterns of HCHO VCDs and AODs are rather similar, however, no minimum is found under south-westerly winds for HCHO. The angular distributions of the surface quantities of both species show larger discrepancies.

The observed differences in the angular distributions of both aerosol quantities can probably be explained by the distance of possible emission sources of aerosol or their precursors to the measurement site. While rather fresh emissions of the city of Mainz lead to enhancements of the near-surface extinctions under east-south-easterly to southerly winds, no strong enhancements of the surface extinctions are found under more easterly winds since the major emission sources in this sector are located several tens of kilometres away. Under these more easterly winds, air masses are mainly influenced by the cities and towns of the Rhine-Main region, its heavy traffic and Frankfurt airport. Here, the aerosol is already well-mixed into the lower tropospheric column and as a consequence no particular enhancement of near-surface aerosol can be observed. This effect

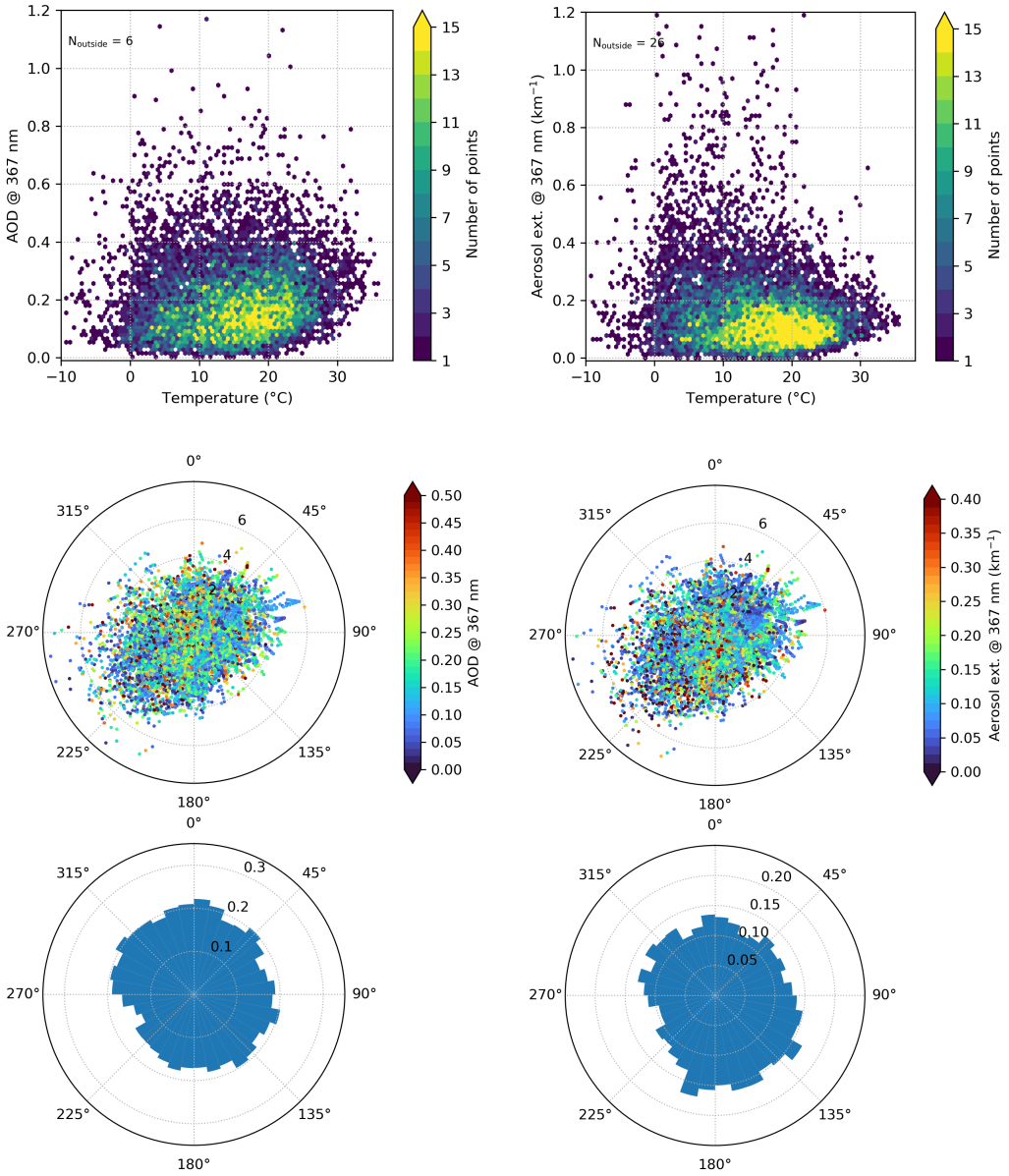


Figure 11.11: The two uppermost panels show hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) measured at Mainz as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median AODs (left) and surface aerosol extinctions at 367 nm (right) obtained under different wind directions in bins of 10° are indicated by the radius.

leads to similar AODs that are observed in a wide sector spanning from ca. 20 to 170°. The reason for the additional enhancement of the AODs under westerly to northerly winds is not fully clear. One possible explanation might be the formation of SOA since also the HCHO abundances are enhanced in this wind sector probably caused by forest emissions from the Taunus mountain range.

Further, several industries and the city of Wiesbaden are located in this wind sector.

Lastly, it should be mentioned that despite the rather homogenous angular distributions of both aerosol quantities, the presented findings do not rule out horizontal gradients in the aerosol distributions. As outlined above, the less good agreement to AERONET data can likely be attributed to strong horizontal gradients occurring at Mainz.

Also at Mainz, expected results are obtained for the dependencies of the measured atmospheric water vapour content on the meteorological conditions. An increase of the retrieved H_2O VCDs and surface concentrations with increasing temperatures is found, while no strong dependence on the wind characteristics can be observed. Here, only slightly larger abundances are obtained for both quantities under westerly winds compared to easterly winds as air masses originating from these sectors are typically warmer and more humid. The results for HONO are briefly discussed in Sect. 11.7.

11.4 TRACE GAS CORRELATIONS

To further study the influence of anthropogenic emissions on the HCHO abundances at Mainz and to generally gain more detailed insights into the relation between HCHO, NO_2 and aerosol, their VCDs/AODs and near-surface concentrations/extinctions are correlated amongst each other. For this, hourly means of both quantities were used and obvious outliers were removed. For a more detailed explanation of how the data were handled, refer to Sect. 10.4.

It should be stressed that no correlations between glyoxal and the other species are presented since the glyoxal results were excluded from atmospheric studies which go beyond annual and diurnal cycles because of insufficient data quality. Further, no investigations of the ratio of glyoxal to formaldehyde are included for the Mainz measurements.

Figure 11.12 shows the correlation of HCHO and NO_2 VCDs and near-surface concentrations in the upper and lower parts, respectively.

Overall, the observed patterns in the correlation of HCHO and NO_2 abundances are similar to Bayfordbury indicating similar general characteristics of both stations. Also for Mainz, rather complex patterns showing two major branches are found for both quantities. These branches mainly correspond to winter and summer while values in between are obtained for spring and autumn. Despite these similarities, it should be mentioned that the separation into the two major branches is less clear for Mainz which might be related to the fact that anthropogenic emission sources are in general closer to the measurement site compared to Bayfordbury.

The winter branches of both quantities show only low to moderate HCHO VCDs and surface concentrations which mainly stay below approximately 1×10^{16} molec cm^{-2} and 1×10^{11} molec cm^{-3}, respectively. Largest HCHO abundances occur at high NO_2 pollution levels exceeding ca. 1×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3}. For these branches, reasonable correlations between HCHO and NO_2 are found for both quantities which are slightly worse compared to

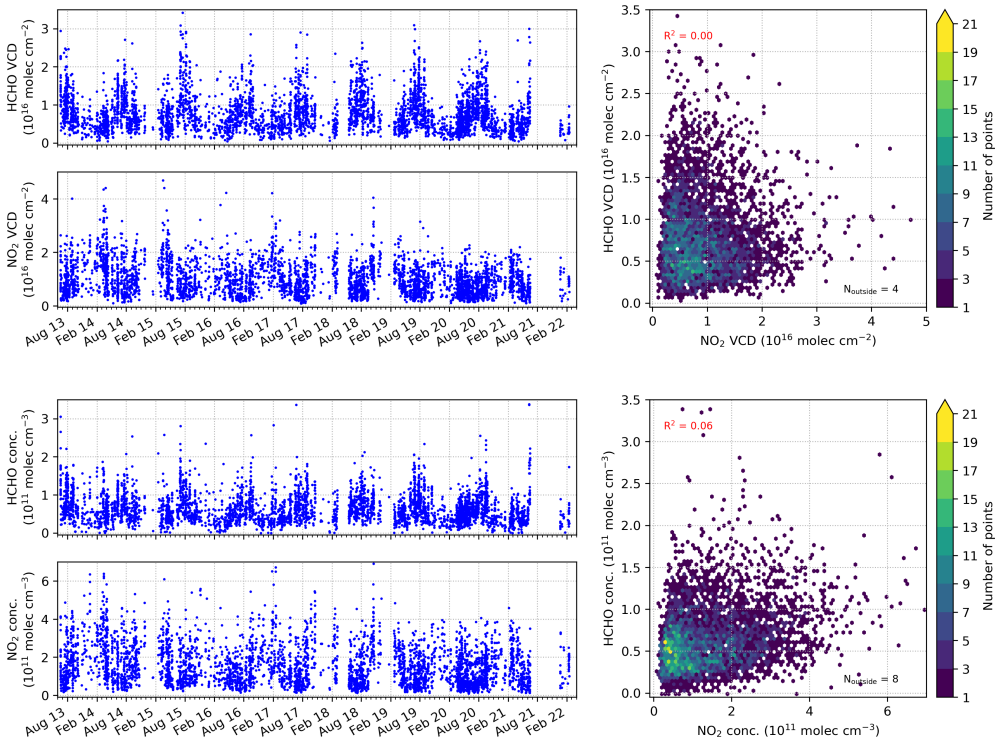


Figure 11.12: Comparison of formaldehyde and NO_2 VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Bayfordbury[2]. Further, the correlation between both species is better for the near-surface concentrations. As local emissions typically affect the surface concentrations stronger than the whole lower tropospheric column, this indicates a significant influence of local anthropogenic emissions on the HCHO abundances since biogenic activity is low during winter. Here, both direct HCHO emissions as well as emissions of precursor substances might play a role.

Like at Bayfordbury, the summer branches of VCDs and surface concentrations exhibit more complex patterns. Nevertheless, it becomes evident that highest HCHO VCDs occur at moderate to low NO_2 abundances below roughly 1×10^{16} molec cm^{-2} which again indicates that generally biogenic emissions determine the formaldehyde abundances in summer. On the other hand, a sub-branch exists which is associated with moderate HCHO VCDs occurring at slightly enhanced NO_2 abundances. For the near-surface concentrations, this sub-branch is even stronger pronounced, while the bulk of elevated HCHO surface concentrations still occurs at NO_2 surface concentrations below ca. 1×10^{11} molec cm^{-3}. These sub-branches indicate that also during summer anthropogenic emissions can have a significant additional influence on the measured HCHO abundances at Mainz. These sub-branches seem to be stronger pronounced at Mainz compared

2 For the resulting R^2-values of the correlations on seasonal basis, refer to Table A.2 in the appendix.

to Bayfordbury which can be explained by the fact that anthropogenic sources are generally closer to the measurement site at Mainz.

Correlations of formaldehyde and aerosol column and near-surface quantities are presented in Fig. 11.13.

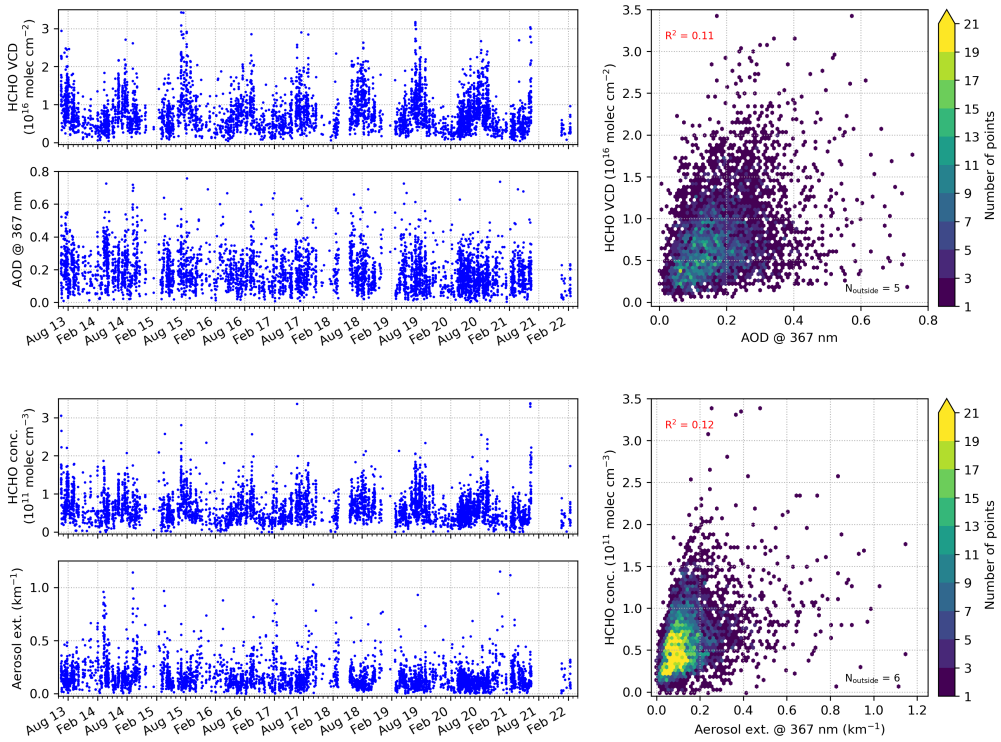


Figure 11.13: Comparison of formaldehyde and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

For both quantities, rather complex patterns and no high correlations between both species are found when the whole data sets as presented in Fig. 11.13 are considered. Nevertheless, a weak tendency that generally higher aerosol loads are associated with higher HCHO abundances is visible for both quantities. This also holds for a separation of the data into the different seasons. For the VCDs/AODs, mainly the variations according to the presented annual cycles with no major changes in the correlations are found. For the near-surface quantities of both species, however, reasonable correlations are obtained in the individual seasons as mainly the pattern found in the lower part of Fig. 11.13 is split up in summer and winter branches. These show enhanced HCHO surface concentrations at moderate near-surface extinctions in summer and low to moderate HCHO surface concentrations at generally higher surface extinctions in winter. Since these correlations between the near-surface quantities of both species are higher in winter, these findings stress that local anthropogenic emissions have a significant impact on the HCHO (and aerosol) abundances during this period of the year.

Lastly, Fig. 11.14 depicts the correlations of NO_2 and aerosol results.

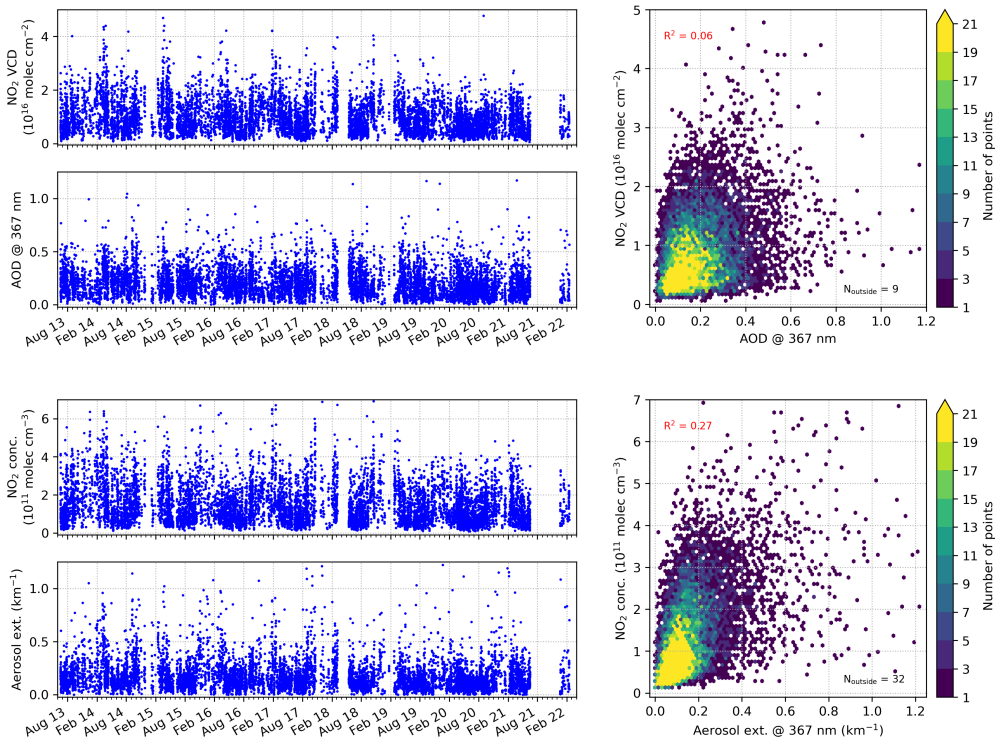


Figure 11.14: Comparison of NO_2 and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

In general, similar patterns as for the correlations between HCHO and aerosol results are obtained. However, the surface quantities of NO_2 and aerosol show a slightly better correlation than observed above with higher NO_2 abundances occurring at higher aerosol loads. Also the separation into different seasons does not change the overall patterns nor the correlations significantly for both column and surface quantities. The findings for the near-surface quantities indicate that locally NO_2 and aerosol have mostly common sources which was also suggested by rather similar angular patterns of the dependencies on the wind direction. For the column quantities of both species, already these angular distributions showed rather different patterns indicating that elevated NO_2 VCDs are not always associated with enhanced AODs.

11.5 VERTICAL PROFILES

Figure. 11.15 shows median vertical profiles of formaldehyde and NO_2 concentrations as well as aerosol extinction at 367 nm for each of the four seasons indicated by different colours. Similar to Bayfordbury, the profiles of all species are affected by an artefact in the lowermost layers as outlined in detail in Sect. 9.2 and 10.5.

Formaldehyde shows its highest concentrations up to a height of approximately 300 m after which the concentrations smoothly decrease with increasing altitude.

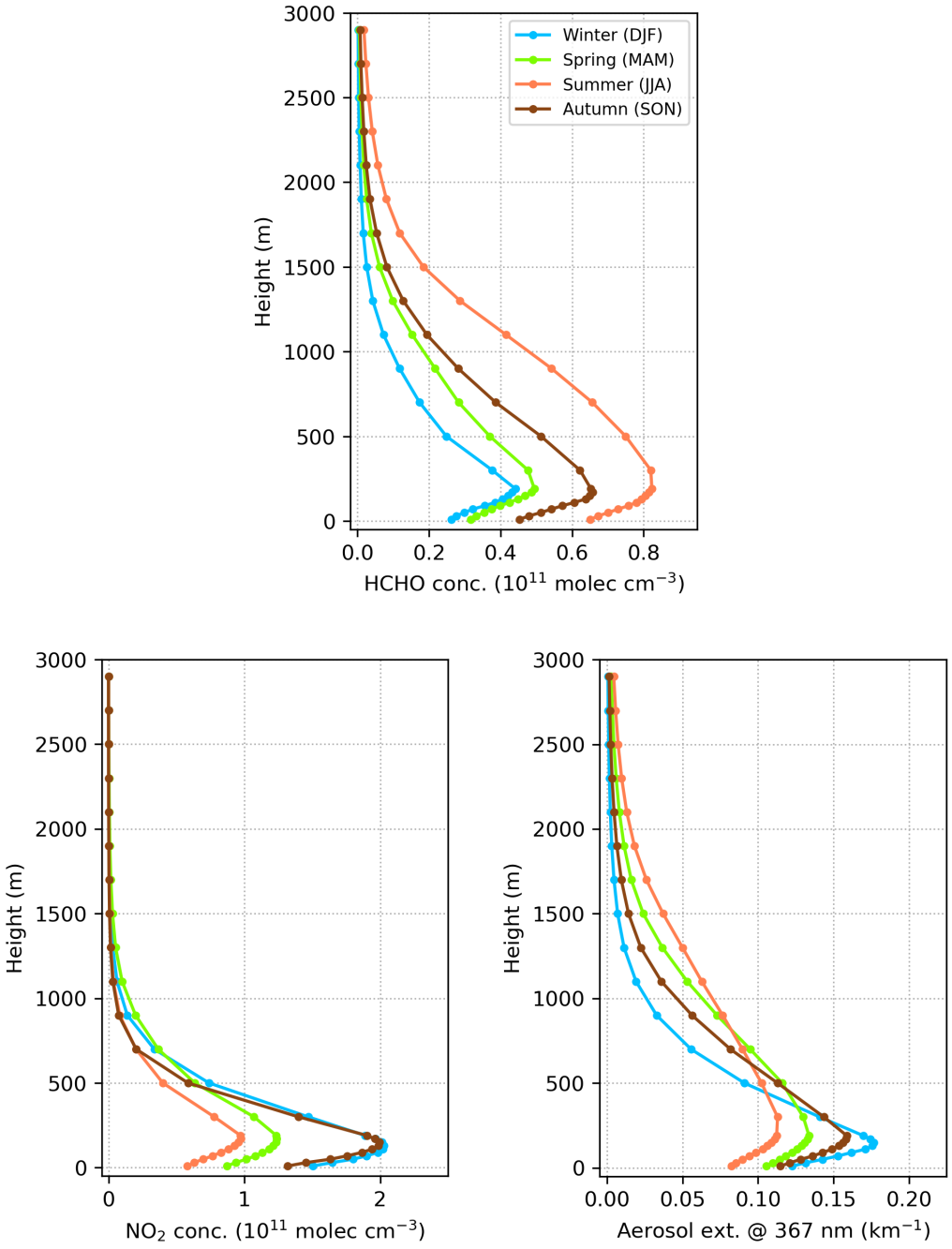


Figure 11.15: Seasonal median profiles of formaldehyde (upper panel) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the Mainz measurements. Profiles for the different seasons are indicated by different colours.

The median HCHO profiles reach higher than the profiles of NO_2, whereby the majority of the HCHO content is distributed up to a height of around 1.5 km. Besides the variations of the HCHO concentrations between the seasons, the layer in which the maximum of the HCHO is concentrated is obviously more confined in winter and autumn, while the vertical distribution is vertically more extended

in summer. Further, the height where the maximum concentration is found is slightly lower in winter than in summer. These two findings can be explained by enhanced vertical mixing leading to typically higher boundary layer heights in summer and the fact that local anthropogenic emissions play a larger role in winter. Despite similar profile heights, the layer in which the maximum HCHO occurs seems to be vertically more extended at Bayfordbury compared to Mainz in all seasons. This likely indicates that possible sources directly emitting HCHO and/or the sources of its precursors are located closer to the instrument in Mainz compared to Bayfordbury. Since this difference between both stations is largest pronounced in winter, this also suggests that during winter the effect of direct HCHO emissions might be larger at Mainz.

The median profiles of NO_2 are substantially shallower than the HCHO profiles and highest NO_2 concentrations are found up to a height of only about 200 m, while the majority of the NO_2 is located below about 1 km. These shallow profiles can be explained by emission sources at the ground and that the conversion from NO to NO_2 happens quickly within a few minutes. For all seasons, overall similar relative profile shapes are obtained mainly reflecting the annual variability of the atmospheric NO_2 content. However, the layer in which the bulk of NO_2 is concentrated seems to be slightly thicker in spring and summer. Further, the decrease of the concentrations after reaching their maximum is less steep in the warmer seasons of the year. These two findings most likely indicate an on average higher boundary layer and enhanced mixing during summer and spring. However, these effects are only weakly pronounced and, similar to Bayfordbury, parts of the effect of stronger vertical mixing might be counteracted by shorter NO_2 lifetimes in the respective seasons. Despite this similarity between the two stations, the NO_2 is located closer to the ground at Mainz and the layer in which it is mainly concentrated is thinner again caused by the fact that the emission sources are located closer to the instrument at Mainz. Finally, the maximum of the NO_2 concentrations at Mainz is more than twice as high as at Bayfordbury, whereby in winter values of 2.03×10^{11} and 0.85×10^{11} molec cm^{-3} are observed at the two stations, respectively.

For the profiles of aerosol extinction at 367 nm, large differences between the seasons are found. While the maximum aerosol extinctions are found up to a height of about 300 m in spring, summer and autumn, the maximum is found closer to the ground in winter (ca. 200 m). Also the height up to which the majority of the aerosol content is distributed varies significantly between the seasons, where values of around 1.25 km and 2 km are found for winter and summer, respectively. Values in between are obtained for spring and autumn. Further, the thickness of the layer containing the bulk of aerosol is larger in the warmer parts of the year compared to winter. These differences in the relative profile shapes can be explained by variations in the vertical mixing strengths throughout the year leading to different boundary layer heights. Lastly, larger near-surface extinctions are found in winter compared to summer, with values in between for spring and autumn. The observed differences in the relative profile shapes between the seasons explain the findings for the annual cycles of AODs and near-surface aerosol extinctions (Fig. 11.7). These showed almost constant monthly mean

AODs throughout the year, whereas the near-surface extinctions were higher from November to March. Thus, it seems that higher reaching aerosol profiles compensate the lower near-surface extinctions during summer, while shallower profiles are counteracted by larger near-surface extinctions during the colder seasons of the year. At Bayfordbury, almost no variations of the relative profile shapes between the different seasons were observed and the layer containing the bulk of the aerosol was vertically more extended compared to Mainz in all seasons, while the general profile height containing the majority of the aerosol load is similar for both stations. The large differences of the annual patterns observed for the aerosol profiles at Mainz and Bayfordbury and the generally thinner layers occurring at Mainz can mainly be explained by two factors. Firstly, the emission sources are closer to the instrument at Mainz. Secondly, the amplitude of the air temperature variations throughout the year is smaller at Bayfordbury leading to less differences between the seasons.

As outlined in Sect. 11.2, also for the Mainz measurements consistently lower near-surface concentrations are obtained using a variable O_4 SF compared to a fixed one of 0.8, while only marginal differences are observed for the trace gas VCDs. For the aerosol retrieval, both quantities show lower values applying a variable SF. These findings can be explained by slight differences of the obtained vertical distribution of the total trace gas column using either SF scenario. A detailed explanation of this effect was already provided for Bayfordbury in Sect. 10.5.

Similar to Bayfordbury, relatively high reaching (up to about 2.5 km) vertical profiles are obtained for water vapour. All seasons show very similar relative profile shapes, whereby the absolute values mostly reflect the annual variability of the atmospheric water vapour content.

11.6 COMPARISON TO TROPOMI

In this section, a brief comparison between HCHO VCDs measured by the MAX-DOAS instrument at Mainz and the TROPOMI satellite instrument is presented. It should be noted that TROPOMI HCHO VCDs are only available from May 2018 onwards as outlined in Sect. 5.3. Despite the over eight years of MAX-DOAS data available, this leads to a data overlap of only three to four years given the sparse MAX-DOAS sampling towards the end of the measurement period (compare Sect. 11.1). Details of how both data sets were treated for the comparison are described in Sect. 10.6. Comparisons on daily (upper part) and monthly (lower part) basis are shown in Fig. 11.16.

Reasonable agreement between both data sets with an R^2 of 0.39 is found for the daily mean HCHO VCDs. Best agreement is observed for moderate formaldehyde abundances, while the TROPOMI measurements underestimate especially high VCDs. Further, a substantial fraction of the TROPOMI VCDs scatters around 0, whereas non-zero background or moderate HCHO abundances are obtained for the MAX-DOAS measurements in these cases. These observed differences between both data sets lead to a mean difference of $(-2.55 \pm 0.31) \times 10^{15}$ molec cm^{-2} which corresponds to 33 % of the total mean MAX-DOAS HCHO VCD around the

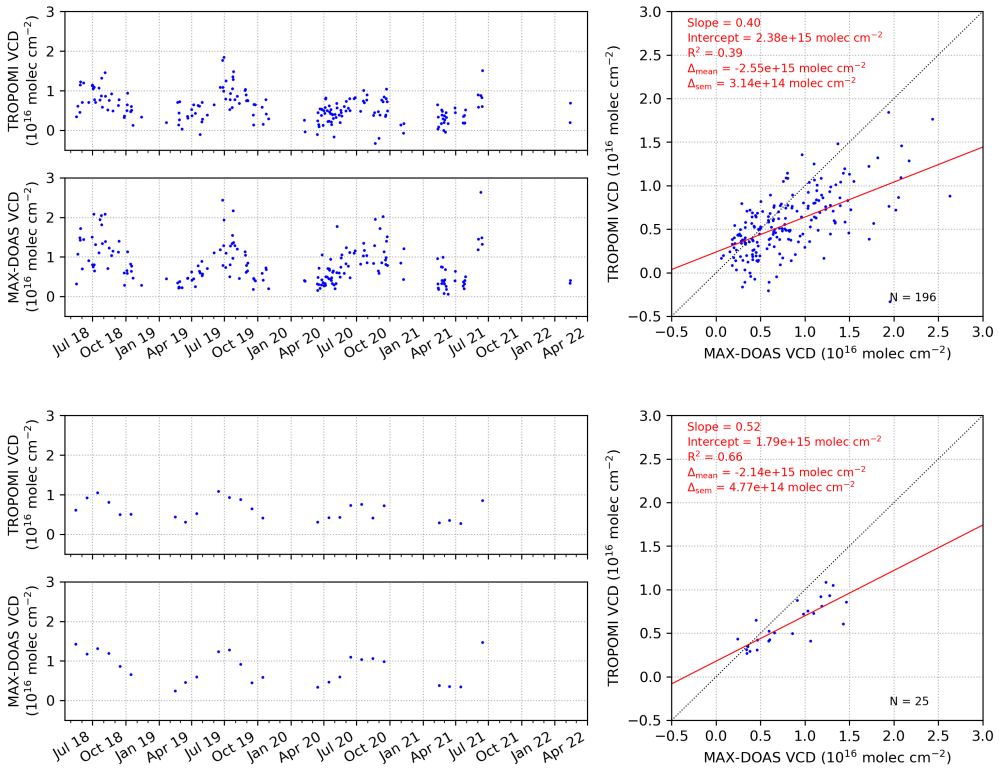


Figure 11.16: Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Mainz on daily (upper part) and monthly basis (lower part). Each part consists of time series of TROPOMI (upper left panels) and MAX-DOAS (lower left panels) VCDs as well as the corresponding correlation plots (right panels). The results of a simple linear regression are shown in red and the number of data points used in the comparisons is indicated in the lower right corners.

satellite overpass (7.66×10^{15} molec cm^{-2}). As described in detail for Bayfordbury, these deviations can have several reasons including different sensitivities of both measurement types to trace gases located close to the ground which are not accounted for in this simple comparison. Further, the size of the satellite ground pixels and the fact that satellite data were averaged 20 km around the MAX-DOAS station play a role. For instance, the sky can be completely clear over Mainz (according to the MAX-DOAS retrieval) while there might be still clouds in other parts of the (averaged) TROPOMI pixels. In addition, a low bias might be introduced when areas with lower HCHO abundances than at Mainz are sampled by TROPOMI which is probably sometimes the case for the Mainz measurements.

As expected, the comparison on a monthly scale yields a better agreement between both data sets. The R^2 is larger (0.66) and also the slope of the regression line is increased, while the mean absolute difference decreases to $(-2.14 \pm 0.48) \times 10^{15}$ molec cm^{-2} corresponding to ca. 28 % of the total mean MAX-DOAS HCHO VCD. These improvements are mainly achieved by the fact that VCDs scattering around 0 have less weight in the calculation of monthly medians. However, the pronounced annual cycle with clearly enhanced HCHO VCDs during summer is still not well represented in the TROPOMI data even for monthly medians.

This leads to a distinct underestimation of the largest HCHO VCDs. Note that, given the selection criteria of both MAX-DOAS and TROPOMI data (compare Sect. 5.3 and 9.3), only sparse data are included in the comparison from October to February and no data at all from December and January.

Overall, the comparison performs slightly worse compared to Bayfordbury. A possible simple explanation might be smaller spatial gradients at rural Bayfordbury.

Also for Mainz, comparisons were performed using the retrieval results for a fixed O_4 SF of 0.8 or other selection criteria. For the latter, the usage of MAX-DOAS data recorded ± 60 min around the local equator crossing time for the comparison and the inclusion of MAPA retrievals flagged with a warning were tested. An overview of the corresponding results of the daily comparisons is provided in Table 11.2.

Table 11.2: Comparison results between daily TROPOMI and MAX-DOAS data at Mainz for different selection criteria.

SCENARIO	QUANTITY				
	R^2	Slope	Intercept (molec cm^{-2})	Δ_{mean} (molec cm^{-2})	Δ_{sem} (molec cm^{-2})
Standard	0.39	0.40	2.38×10^{15}	-2.55×10^{15}	3.14×10^{14}
SF 0.8	0.39	0.39	2.30×10^{15}	-2.54×10^{15}	3.26×10^{14}
± 60 min	0.40	0.43	2.22×10^{15}	-2.51×10^{15}	2.53×10^{14}
Warnings incl.	0.27	0.35	1.95×10^{15}	-4.13×10^{15}	2.44×10^{14}

While the usage of a fixed SF has almost no effect on the comparison results, faintly better results are obtained when data ± 60 min around the local equator crossing time are included in the comparison. Most likely, the latter is caused by better statistics since more data points are included in the comparison (196 vs. 274). Lastly, including also MAPA retrievals which are flagged with a warning clearly worsens the agreement between both data sets.

11.7 FURTHER RESULTS: HONO

While practically no enhanced HONO was measured at Bayfordbury, interesting results were found for Mainz which are briefly presented in this section.

Figure 11.17 shows average annual cycles of HONO VCDs (left) and near-surface concentrations (right) as monthly box-whisker-plots.

Despite the relatively large scatter of the results, distinct annual cycles of the HONO abundances are found for both quantities. These annual variations are very similar to the ones observed for NO_2 (compare Fig. 11.5) indicating a close relation between both trace gases. Hence, enhanced HONO abundances are found from October to March, with monthly means around 1.5×10^{14} molec cm^{-2} and 2.5×10^{9} molec cm^{-3} for VCDs and near-surface concentrations, respectively.

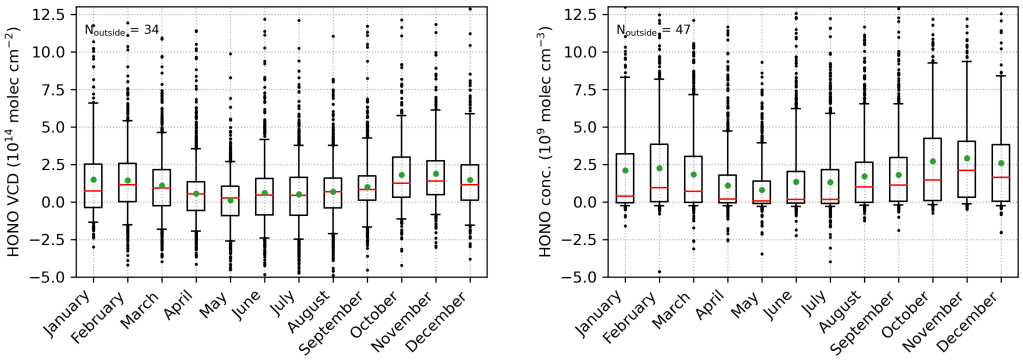


Figure 11.17: Box-whisker-plots of the annual cycles of HONO VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

Clearly lower HONO is found during the rest of the year, where the monthly means are only around 0.5×10^{14} molec cm^{-2} and 1.0×10^{9} molec cm^{-3} for the two quantities. It should be noted that, especially for the near-surface concentrations, the large scatter of the results with many outliers towards high values largely affects the calculation of monthly means. However, also the monthly medians show distinct annual variations for both quantities. The observed annual variation can mainly be explained by the fact that HONO is closely related to NO_2 whose abundances are higher in winter. Further, it is quickly photolysed under the influence of solar radiation which is especially important during summer (Atkinson, 2000). Nevertheless, it should be noted that only little MAPA retrievals yield valid results during winter, while more valid results are obtained in summer where values are closer to 0.

Similar results are obtained using a fixed O_4 scaling factor which mainly leads to slightly larger abundances during winter, while no clear tendency is found in summer. Also, the scatter is further enhanced compared to the results for a variable SF. These findings hold for both VCDs and near-surface concentrations.

The diurnal cycles of HONO VCDs (left) and surface concentrations (right) are presented in Fig. 11.18, where different colours indicate results obtained for the four seasons.

For both quantities, distinct diurnal cycles are found in all seasons which are stronger pronounced for the near-surface concentrations. The general patterns of these diurnal variations are similar in all seasons. Largest HONO abundances occur during the early morning which rapidly approach 0 within a few hours. Minimum HONO surface concentrations are reached around 9 a.m. during spring and summer and around noon in autumn and winter, while minimum VCDs are typically observed a few hours later. As mentioned above, this rapid decrease can be explained by rapid photolysis of HONO under the influence of solar radiation providing an OH source in the early morning which can be considered as initiator of urban photochemistry (Atkinson, 2000). This finding is consistent

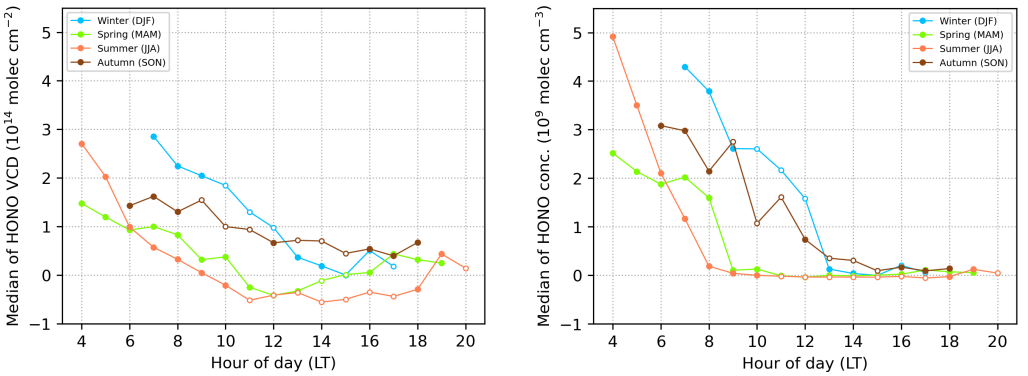


Figure 11.18: Diurnal cycles of HONO VCDs (left) and surface concentrations (right) for Mainz for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

with other studies, e.g. Perner and Platt (1979), Hendrick et al. (2014) and Wang et al. (2017a, 2020). The shift of the diurnal patterns between the different seasons reflects the changing day lengths. Interestingly, morning HONO abundances are similar in winter and summer as well as in spring and autumn. However, both quantities reach values close to 0 in spring and summer, whereas still non-zero HONO abundances are measured during afternoon in autumn and winter. Slightly negative values are obtained for spring and summer in the afternoon. Although this is an interesting finding, this observation can be explained by the fact that the VCDs can become negative in summer (and partly in spring) but the near-surface concentrations overall not which is caused by the specific parametrisation of the MAPA retrieval (compare Sect. 3.3). While the decrease towards 0 is very steep for the near-surface concentrations, it is less steep for the VCDs. Together, the findings firstly suggest that HONO is primarily located very close to the surface which is reasonable since it is photolysed very quickly. Secondly, during autumn and winter small HONO amounts are probably present in the lower tropospheric column throughout the day, being a possible additional OH source. Lastly, no increase towards sunset, and thus higher SZAs, is observed[3] which is consistent to other studies (e.g. Hendrick et al., 2014; Wang et al., 2017a, 2020) and indicates that HONO is mainly formed during night after sunset.

To investigate the dependence of the measured HONO abundances at Mainz on meteorological conditions, Fig. 11.19 shows HONO VCDs and near-surface concentrations as functions of temperature (upper panels) as well as wind speed and direction (middle and lowermost panels).

Given the large scatter of the HONO results, both quantities show almost no temperature dependence. Nevertheless, a slight decrease of the VCDs (upper left panel) with increasing temperature is found which is consistent with the observed annual cycles. For the surface concentrations, on the other hand, no such pattern is observed. A possible explanation for this difference is probably the

3 Note that the increase towards evening which is visible for the VCDs in autumn and summer might be caused by statistical effects.

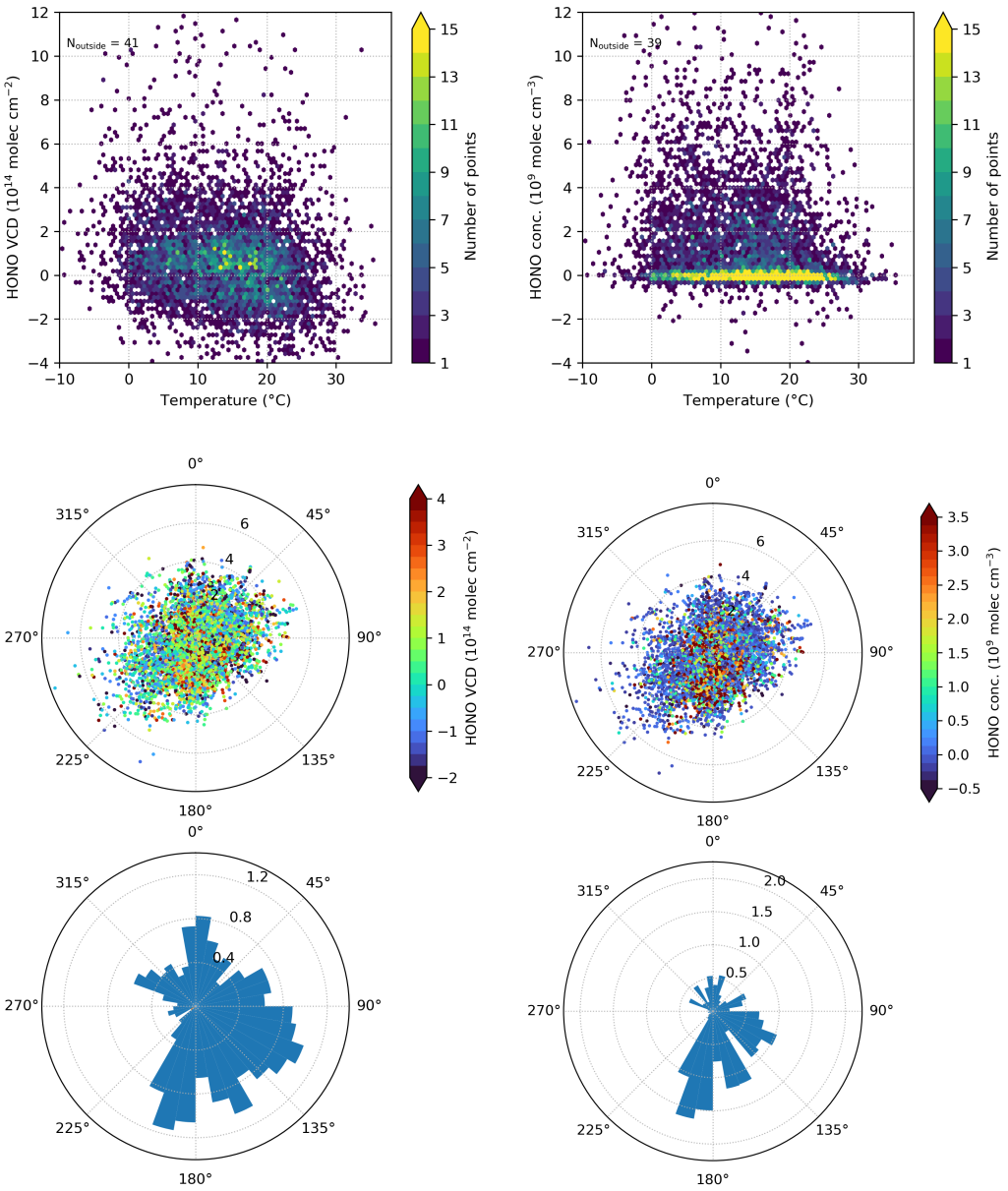


Figure 11.19: The two uppermost panels show hourly mean HONO VCDs (left) and surface concentrations (right) measured at Mainz as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HONO VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HONO VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius. Negative medians are marked in orange.

larger scatter for the near-surface concentrations compared to the VCDs (compare whiskers in Fig. 11.17).

The HONO VCDs show high medians in a large sector spanning (clockwise) from north-west to south-south-west again indicating its close relation to NO_2.

Thus, also for HONO three more confined sectors can be defined in which the VCDs are particularly high, namely 290 to 20°, 60 to 140° and 150 to 210° (compare Fig. 11.10). In these sectors, the probed air masses are mainly influenced by the city of Wiesbaden, the Rhine-Main region together with Frankfurt airport, and Mainz itself, respectively. However, compared to the angular patterns of NO_2, more variability is found for HONO and the obtained angular patterns are less smooth. For the near-surface concentrations, rather extreme angular patterns are found showing largely enhanced HONO surface concentrations under easterly to south-south-westerly winds, while significantly smaller medians are found in the remaining wind sectors. For instance, the marginal enhancements in the more northerly wind sectors are three times smaller than the maximum values under south-south-westerly winds. For both quantities, lowest abundances occur under south-westerly to westerly winds. No clear dependence of the HONO abundances on wind speed is observed for both VCDs and near-surface concentrations (compare middle panels).

The angular patterns of VCDs and surface concentrations potentially indicate a strong HONO source in the south-south-westerly sector since both quantities show their absolute maxima in a sector ranging from about 180 to 210°. Parts of the city of Mainz lie in this sector as well as more rural areas with agricultural fields. No easy attribution to one distinct source is possible with the approach used in this work. Under winds from north-west to south-east (clockwise), HONO appears to be already mixed into larger parts of the lower tropospheric column since the relative enhancements of the VCDs are larger than the ones of the surface concentrations in these wind sectors. Nevertheless, anthropogenic activities seem to be of major importance for the observed abundances since practically no HONO is measured under south-westerly to westerly winds where mainly rural areas are located. As HONO is not the main focus of this work, no deeper investigations on possible sources are performed.

Lastly, the vertical distribution of HONO is investigated. Figure 11.20 depicts median vertical profiles of HONO concentrations, whereby different colours indicate the results for each of the four seasons.

Compared to the NO_2 profiles (Fig. 11.15), the median profiles of HONO are even shallower with the maximum concentrations occurring up to a height of approximately 120 m, while the majority of HONO is located below about 750 m. Also the decrease of the concentrations above their maximum is steeper compared to NO_2. For autumn and winter, similar relative profile shapes are obtained while during spring and summer only slightly enhanced concentrations occur close to the ground. Although enhanced morning values (compare Fig. 11.18) are observed also in the warmer part of the year, the values are close to or equal 0 during large parts of the day leading to overall low median concentrations (compare Fig. 11.17). Also the observation that higher concentrations are found for the profiles in autumn than in winter is caused by the usage of median profiles instead of means (compare red lines and green dots in Fig. 11.17). Nevertheless, the main findings and the relative profile shapes are not affected by these statistical effects which can be understood. The very shallow profiles of HONO are caused by at least two factors. Firstly, enhanced HONO mainly occurs in the early morning

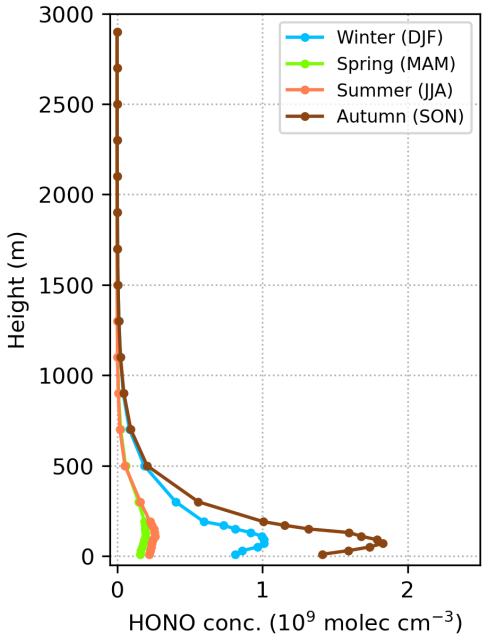


Figure 11.20: Seasonal median profiles of HONO concentrations obtained for the Mainz measurements. Profiles for the different seasons are indicated by different colours.

hours right after sunset where the boundary layer is still rather shallow. In the further course of the day, when the boundary layer rises, HONO abundances are much lower and not many valid profiles are obtained (compare many open circles in Fig. 11.18). Secondly, HONO is photolysed very quickly. As a consequence, its atmospheric lifetime during day time is rather short and it is not mixed into higher atmospheric layers.

In conclusion, several interesting aspects of HONO are observed at Mainz. However, HONO is not the main focus of this work and thus no detailed interpretation is given here. Since HONO chemistry is very complex, still not well understood and many open questions exist, further studies should focus on HONO measurements at Mainz.

12

MOHALI, INDIA

This chapter discusses the results of the measurements conducted at Mohali (Sect. 4.3), India, whereby the focus is put on HCHO, glyoxal, NO_2 and aerosol results. Following the structure of the two previous chapters, first a general overview of the measurements including particularities of the analysis is provided in Sect. 12.1. After a discussion of annual and diurnal cycles of the main species HCHO, glyoxal, NO_2 and aerosol in Sect. 12.2, dependencies of the abundances of these species on meteorological conditions are investigated in Sect. 12.3. The relation amongst the different species themselves is studied in Sect. 12.4, while vertical profiles of the main species are presented in Sect. 12.5 on a seasonal basis. A brief comparison of the MAX-DOAS formaldehyde results to TROPOMI is provided in Sect. 12.6, while interesting HONO results are presented in Sect. 12.7 which concludes this chapter.

12.1 OVERVIEW AND PARTICULARITIES OF THE ANALYSIS

Following first MAX-DOAS measurements by Kumar et al. (2020), the measurements at Mohali presented in this work started on 24 April 2019. Although the measurements are in principle still running (as of September 2023), only data until 31 July 2022 are considered in this work. This time range covers almost three full annual cycles and includes lockdown measures due to the COVID-19 pandemic in 2020.

Thanks to excellent local and remote support, only a few data gaps occurred in 2019 which were caused by minor instrumental issues. After that, the measurements were very stable and only marginal data gaps in the order of days exist. Logging of the spectrometer temperature started only in July 2020 and showed stable values of 15 °C over the whole period. Also in the period before the start of the logging there is no evidence of unstable temperatures, for instance no unexpected patterns in the fit RMS and intensity offsets occurred. Given this stable spectrometer temperature, one set of DC and offset spectra was used for the correction of the measured spectra and the same spectral calibration and SLF parameters could be used for the DOAS retrievals for Mohali.

All standard fit settings according to Sect. 9.1 could be used for the spectral analysis and only one minor spectral gap from 345.3 to 345.7 nm to deal with a hot pixel had to be introduced in the HCHO, HONO and O_4 fits. Also the MAPA retrievals were conducted using the standard settings described in Sect. 9.2.1. The applied elevation bias was 1° and determined as explained in Sect. 9.2.2. It should be noted that daily horizon scans monitoring the elevation calibration were implemented only in July 2020, however, no indications of a significantly different elevation calibration before are given. Finally, all data could be used for the atmospheric investigations after they were filtered according to Sect. 9.3.

Another interesting aspect is that the MAPA retrievals for Mohali yield the largest fraction of valid MAPA results amongst all five instruments used in this work, e.g. more than 25 % of the inversions are valid for HCHO (compare Table 9.8). This is probably caused by the fact that spectra at Mohali were recorded using a total integration time of only 30 s, while 60 s were used at all other stations operating a TubeMAX-DOAS instrument[1] (compare Table A.1 in the appendix). This shorter integration time might lead to less atmospheric variations within one elevation sequence and thus to more valid retrievals. However, no investigations on that topic were performed but offer an interesting starting point for further improvements of the MAX-DOAS measurements.

12.2 ANNUAL AND DIURNAL CYCLES

To provide a general impression of the results for Mohali, annual and diurnal cycles of HCHO, glyoxal, NO_2 and aerosol are presented in this section. For that, the data were visualised as outlined in detail in Sect. 10.2 after the usual filters (compare Sect. 9.3) were applied.

Figure 12.1 shows the resulting annual cycles of HCHO VCDs and near-surface concentrations, where the two upper panels show full time series and the two lower panels depict monthly box-whisker-plots. It should be noted that, given the shorter integration times for the spectra recorded at Mohali, more individual elevation sequences are recorded in the same time. Thus, slightly different sampling filters as for the other stations were applied for the calculation of daily and monthly means used for the visualisation of the full time series. For daily means, days with less than five valid data points were marked by light colours, while monthly means were only calculated for months with at least 20 valid individual measurements. Lastly, months with less than 60 valid single measurements were labelled with open circles.

Both quantities show distinct annual variations which are strongly influenced by the monsoon cycle. Minima are found in late winter/early summer[2] and maximum HCHO abundances typically occur in the pre- and post-monsoon months, while local minima are observed during monsoon in July and August. Apparently, the pre-monsoon maximum is less pronounced, but still visible, for the surface concentrations (details are discussed below). Given the subtropical climate with rather high temperatures, maximum HCHO VCDs above 3.5×10^{16} molec cm^{-2} can occur at Mohali. Further, the post-monsoon season of 2020 shows an exceptional enhancement of the near-surface concentrations compared to other years, while the VCDs show only a slightly larger enhancement than in the remaining years.

The observed annual variations of the HCHO abundances are generally confirmed by the statistical representation of the annual cycles of HCHO VCDs (left) and surface concentrations (right) depicted in the two lower panels of Fig. 12.1. However, differences in the detailed annual patterns of both quantities

1 The 4-azimuth MAX-DOAS instrument operated at Mainz uses a different approach of determining the integration time.

2 Note the different definition of the seasons at Mohali (compare Sect. 4.3).

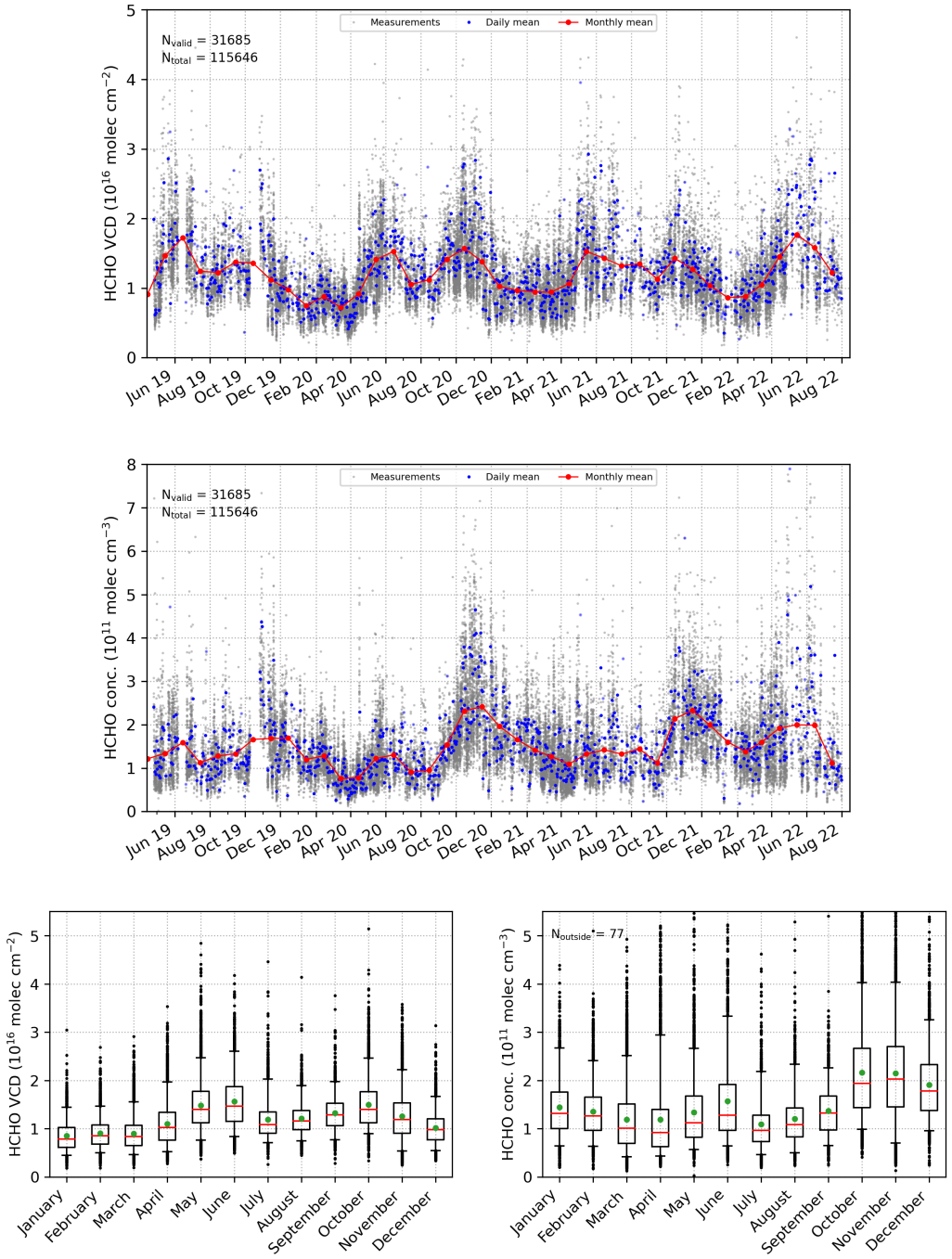


Figure 12.1: The two upper panels show time series of HCHO VCDs and surface concentrations for Mohali. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 60 valid data points. The lower panels depict box-whisker-plots of the annual cycles of HCHO VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

emerge from their respective monthly box-whisker-plots. The VCDs show their lowest monthly means from January to March, where similar values of around 0.9×10^{16} molec cm^{-2} are obtained in all three months. On the other hand, a slight decreasing trend in this time period is observed for the near-surface concentrations which show their minimum in March and April with means of ca. 1.2×10^{11} molec cm^{-3}. Towards the monsoon season, as the temperatures rise in early summer, both quantities indicate an increase of the HCHO abundances and reach maxima in the pre-monsoon month June. Here, monthly means of 1.55×10^{16} molec cm^{-2} and 1.55×10^{11} molec cm^{-3} are found for VCDs and surface concentrations, respectively. For the VCDs, this value corresponds to the annual maximum which is not the case for the surface concentrations. During the monsoon season (July and August), both quantities show local minima. While the monsoon values are comparable to or even below the minimum in March and April for the near-surface concentrations, the VCDs stay above their absolute minimum during winter. During post-monsoon, an increase of the HCHO abundances is observed which peaks in October. For the VCDs, values similar to pre-monsoon are found, while the absolute annual maximum of approximately 2.15×10^{11} molec cm^{-3} is reached for the surface concentrations. The latter show no decrease in November and only a slight decrease during winter towards March as outlined in the beginning. Contrary to that, for the VCDs a decrease is observed from November on towards the overall minimum during late winter.

As mentioned above, the observed variations during the course of the year are closely related to the annual monsoon cycle which is, besides variations of temperatures and wind patterns (compare Sect. 4.3), tightly connected with anthropogenic activities. For instance, large-scale crop residue burning occurs in the pre- and post-monsoon months, while winter is associated with private fires for domestic heating and cooking (compare Sect. 4.3). These activities apparently have a significant impact on the observed HCHO levels and partly explain the slight differences in the annual patterns of VCDs and near-surface concentrations. In particular, the emissions of private fires during winter in combination with a rather shallow boundary layer mostly affect the lower atmospheric layers leading solely to an enhancement of the surface concentrations. Further, direct emissions might have an important influence, since low ambient temperatures prevail and sunlight is limited, which probably do not lead to a large enhancement of the HCHO VCDs. The almost opposite behaviour with the maximum VCDs occurring in pre-monsoon months, while maximum monthly means of surface concentrations are observed post-monsoon, can be explained by higher temperatures and almost uninterrupted solar radiation in the pre-monsoon season. This leads to high biogenic and anthropogenic activities together with very effective photochemistry yielding very high HCHO abundances. In the post-monsoon season, on the other hand, temperatures and sunlight are already decreasing. Thus, the anthropogenic sources gain more importance. Also differences in the boundary layer heights probably contribute to these behaviours and more of the total HCHO amount is concentrated in the lowest atmospheric layers. Lastly, an effect of the type of burned crop residue cannot be ruled out since during pre-monsoon mainly wheat residue fires occur, whereas paddy residue fires are dominant in

October (Kumar et al., 2020). The fact that the local minimum during the monsoon months is stronger pronounced for the near-surface concentrations is probably caused by only biogenic emissions determining the atmospheric HCHO content as anthropogenic activities are low during that period and temperatures are still rather high.

The findings obtained within this work are consistent with the investigations by Kumar et al. (2020) who also reported maximum HCHO VCDs during pre- and post-monsoon months and lowest abundances in March. Also the absolute numbers agree reasonably well. Kumar et al. (2020) reported HCHO VCDs of around 1.72×10^{16} molec cm^{-2} in May and June, while values of ca. 1.55×10^{16} molec cm^{-2} were found in this work. Slight differences occur during winter where higher values during December were reported by Kumar et al. (2020). Minimum HCHO abundances were found in their work in March (ca. 0.6×10^{16} molec cm^{-2}), while comparable VCDs are observed from January to March here. These differences are probably caused by slight fluctuations of the specific characteristics in different years.

As seen earlier, almost no differences in the general characteristics of the HCHO results are observed for the fixed O_4 SF of 0.8. However, slightly larger values are obtained for both quantities using the fixed SF, whereby the effect is stronger for the surface concentrations and seems to have different strengths for different seasons. The latter was not seen for the two previous stations and already indicates larger differences in the profile shapes between the seasons for Mohali. Further details are discussed in Sect. 12.5.

Diurnal cycles of formaldehyde VCDs and near-surface concentrations for the different seasons are shown in Fig. 12.2. Note that a deviating definition of the seasons is applied for Mohali compared to the mid-latitude stations Bayfordbury and Mainz. Further, different sample criteria were used and only 1-hour bins with more than 20 valid retrievals are displayed, while no modifications of the relative sampling filter were made.

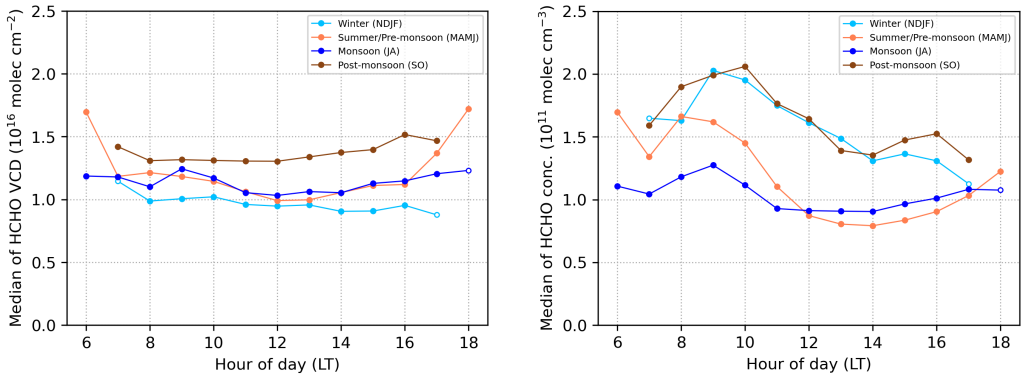


Figure 12.2: Diurnal cycles of HCHO VCDs (left) and surface concentrations (right) for Mohali for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

For the VCDs, almost no diurnal cycle is found for any of the seasons indicating almost constant HCHO abundances throughout the day. Only in summer margin-

ally lower VCDs are observed around noon and a slight increase of the HCHO VCDs towards late afternoon in the post-monsoon season is visible. The largely enhanced values in the early morning and late evening hours during summer are statistical artefacts caused by the dominance of May and June measurements in these bins, where days are longest and HCHO VCDs are highest. For the near-surface concentrations, pronounced diurnal variations are found in all seasons. An increase of the HCHO surface concentrations during the early morning hours is found and the maximum is reached in the late morning hours (9 to 10 a.m.). After that, the surface concentrations decrease reaching their minimum values in early afternoon. For all seasons except winter, slight increases of the HCHO surface concentrations can be observed towards evening. Note that possibly parts of these increases are statistical artefacts. While the general patterns are similar in all seasons, the amplitude is strongest in summer/pre-monsoon and weakest during monsoon. The diurnal cycles of the near-surface concentrations can be explained by a combination of different effects. Firstly, anthropogenic (traffic) emissions peak during the rush hours in the morning and the afternoon. Secondly, the sources apparently exceed the sinks in the morning resulting in pronounced maxima of the near-surface concentrations, while the loss processes of HCHO likely dominate in the afternoon where high solar radiation (photolysis) and high OH concentrations prevail leading to rather short lifetimes. Consequently, no peak in the HCHO surface concentrations is visible in the afternoon. Only towards sunset, as both solar radiation and OH concentrations decrease, slight increases of the near-surface concentrations are found in some seasons. Another aspect contributing to the pronounced morning peak is that crop residue fires which likely act as strong HCHO sources typically occur in the morning (Kumar et al., 2020). A consequence of the latter is also that the amplitude of the diurnal variations of the near-surface concentrations is weakest in the monsoon season. Here, the crop residue fires cease and no strong enhancement is visible during the morning hours resulting in a rather weak diurnal cycle in that season. Lastly, also boundary layer dynamics have to be considered. While the boundary layer is rather shallow in the morning leading to an accumulation of HCHO close to the ground, less HCHO is concentrated in the lowermost atmospheric layers in the afternoon since enhanced mixing and a higher boundary layer distribute the HCHO over larger parts of the tropospheric column. Hence, the VCDs stay rather constant over the course of the day, whereas the near-surface concentrations are strongly varying.

As shown below, the observed diurnal variations of the HCHO surface concentrations are similar to the ones of NO_2 (compare Fig. 12.6) again indicating a close relation of the measured HCHO abundances to anthropogenic emissions. However, the maxima seem to be slightly shifted by plus/minus one to two hours between both species.

Lastly, it should be mentioned that Kumar et al. (2020) found similar cylces for the near-surface concentrations of HCHO, however, a discrimination of the concentrations in the height ranges from 0 – 200 m and 200 – 400 m was made. The latter led to slightly different findings for the lowest 200 m in winter, where the profiles are rather shallow (compare Sect. 12.5).

The annual cycles of glyoxal VCDs and near-surface concentrations are presented in Fig. 12.3.

The general patterns of the annual variations of glyoxal VCDs and near-surface concentrations are very similar to the ones observed for HCHO. Strongly influenced by the annual monsoon cycle, maximum glyoxal abundances occur during the pre- and post-monsoon months, while lower values are typically found during monsoon. Lowest glyoxal abundances are measured in late winter/early summer. Both quantities show overall similar annual variations, however, similar to HCHO, the pre-monsoon maximum is weaker compared to the post-monsoon maximum for the surface concentrations. The post-monsoon season of 2020 shows an exceptional enhancement of the near-surface concentrations compared to other years, while the VCDs show only a slightly larger enhancement than in the remaining years. This was already observed for HCHO and stresses the special nature of the 2020 post-monsoon season.

The statistical representations of the annual cycles of glyoxal VCDs (left) and surface concentrations (right) are depicted in the two lower panels of Fig. 12.3 and mainly confirm the annual variations described above. Despite showing overall very similar characteristics, slight differences between the annual cycles of glyoxal VCDs and near-surface concentrations as well as compared to HCHO can be identified. A first finding is that for glyoxal the average annual cylces of VCDs and surface concentrations are more similar than the cycles of both quantities of HCHO (compare Fig. 12.1). This suggests overall shallower glyoxal profiles with the bulk of glyoxal located closer to the ground compared to HCHO. Details of different profile shapes are discussed in Sect. 12.5. Slightly different than observed for HCHO, both quantities show their minima simultaneously in early summer (March and April[3]), where monthly means of around 1.75×10^{14} molec cm^{-2} and 2.4×10^{9} molec cm^{-3} are obtained for VCDs and surface concentrations, respectively. The transition towards elevated glyoxal abundances in the pre-monsoon months May and June occurs more abrupt compared to HCHO probably indicating different chemical production mechanisms from their precursors. Monthly mean VCDs and surface concentrations of ca. 2.75×10^{14} molec cm^{-2} and 3.3×10^{9} molec cm^{-3}, respectively, are obtained pre-monsoon. Similar to HCHO, the surface concentrations are stronger reduced than the VCDs during monsoon. While the latter stay above their annual minimum values, the surface concentrations drop to values close to their absolute minimum. Compared to HCHO, the glyoxal VCDs are only weakly reduced during monsoon and also the reduction of the glyoxal surface concentrations is apparently smaller. In the post-monsoon months, the surface concentrations reach their absolute maximum (approximately 5.0×10^{9} molec cm^{-3} in October), whereas the VCDs show values only slightly above their pre-monsoon maximum (approximately 3.25×10^{14} molec cm^{-2} in October). Afterwards, as temperatures decrease and crop residue fires cease, both quantities show a decrease towards their minima. Nevertheless, especially for the near-surface concentrations, in the early winter months still relatively high glyoxal abundances prevail. Similar to HCHO, this is most likely linked to private fires for domestic heating.

3 Especially the monthly medians (red lines) indicate that overall minima occur during April.

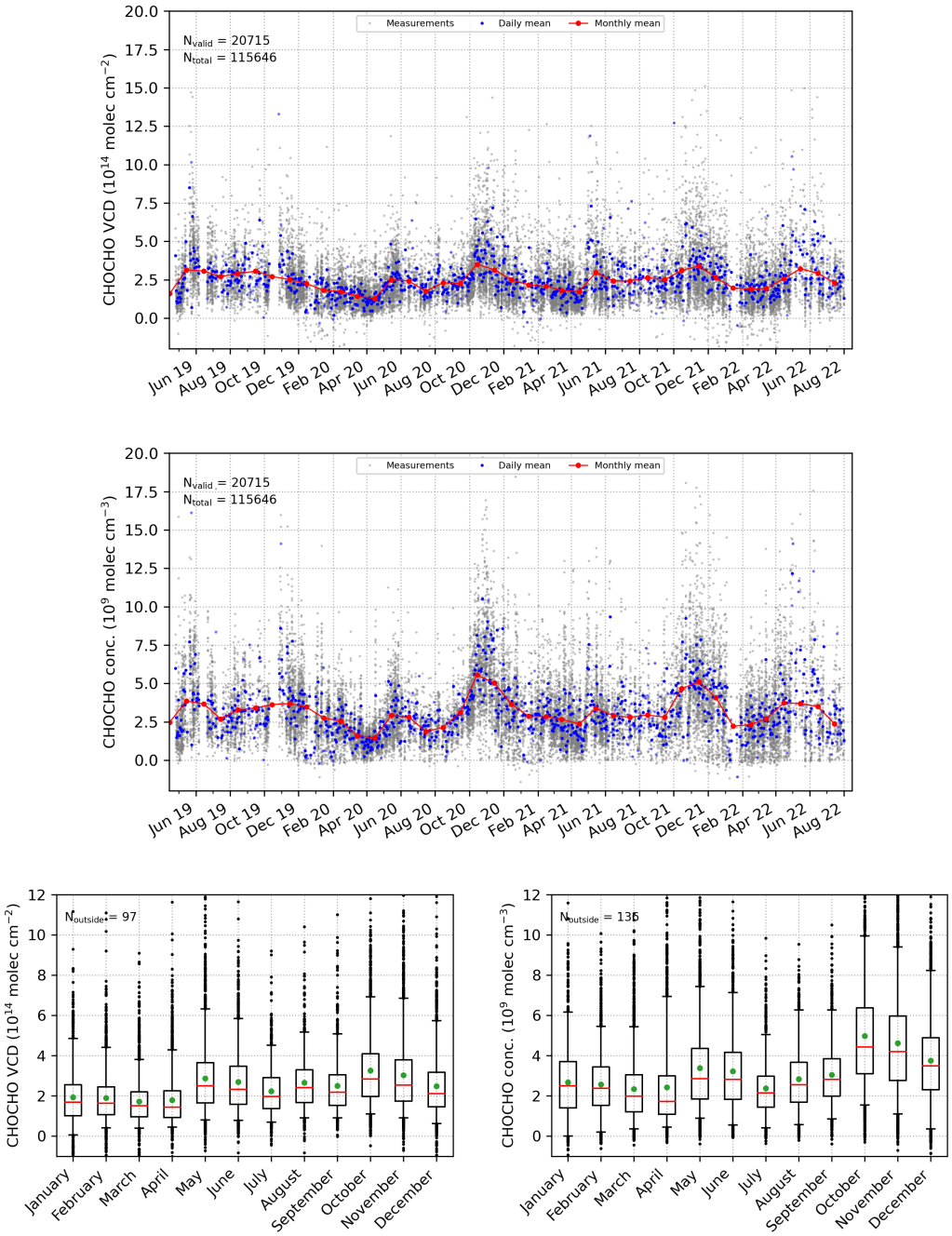


Figure 12.3: The two upper panels show time series of glyoxal VCDs and surface concentrations for Mohali. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 60 valid data points. The lower panels depict box-whisker-plots of the annual cycles of glyoxal VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

The in general similar annual variations of HCHO and glyoxal abundances indicate overall common sources of both species which were outlined in detail above. This is stressed by the fact that especially the annual cycles of the near-surface concentrations of both species are very similar. The slight differences in the detailed patterns might be explained by a combination of dynamical and chemical factors. For instance, differences in the profile shapes probably caused by different chemical source and sink mechanisms. Also the type of burned crop residue can have an influence and might lead to different types of precursor VOCs yielding different amounts of both species. Given the less good quality of the used MAX-DOAS instrument, Kumar et al. (2020) did not report on glyoxal results since no meaningful retrieval was possible.

No major differences in the annual characteristics of the glyoxal abundances are observed for the fixed O_4 SF of 0.8, where slightly larger values are obtained for both quantities. This effect is stronger pronounced for the surface concentrations and seems to have, like for HCHO, different strengths for the different seasons.

Next, the average diurnal variations of glyoxal VCDs (left) and near-surface concentrations (right) in the different seasons are presented in Fig. 12.4.

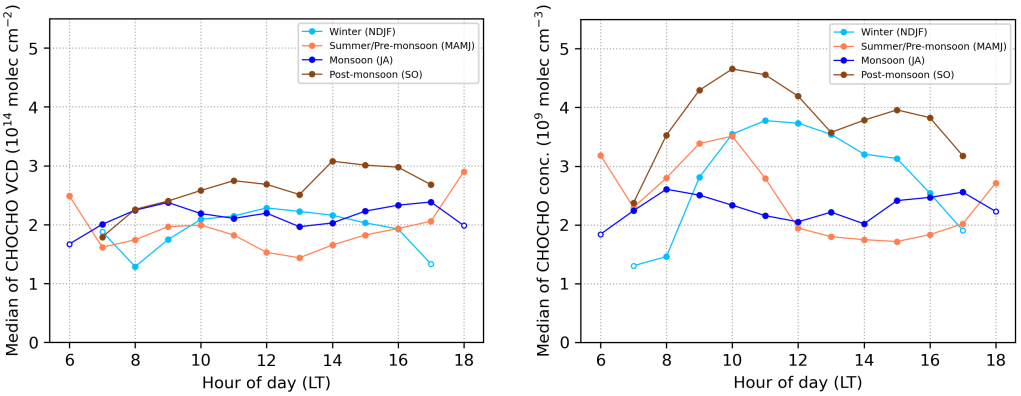


Figure 12.4: Diurnal cycles of glyoxal VCDs (left) and surface concentrations (right) for Mohali for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

While the HCHO VCDs showed almost no diurnal variations, more variability is found for the glyoxal VCDs. In particular, in the post-monsoon season an increase of the glyoxal abundance over the course of the day can be observed which was barely visible for HCHO. A possible reason for this increase could be rather constant sources throughout the day in combination with already less sunlight during this part of the year leading to an accumulation of glyoxal (and faintly also HCHO) abundances. During winter, largest glyoxal VCDs are measured around noon, where an inverse "U-shaped" diurnal pattern is found. The diurnal variation during summer/pre-monsoon is very similar to the patterns observed for the surface concentrations of both glyoxal and HCHO, but less pronounced. Despite showing some variability, no consistent diurnal cycle is found during the monsoon season. Given the rather high fluctuations in the observed diurnal patterns, it should be again stressed that time bins in the early

mornings and late evenings might be affected by statistical artefacts possibly leading to apparent diurnal cycles that are not real.

For the near-surface concentrations, in winter an inverse "U-shaped" diurnal pattern, similar to the VCDs, is found while the diurnal patterns in the remaining seasons are overall similar to the ones observed for HCHO. For instance, an increase during the early morning hours towards maximum glyoxal concentrations which are reached around ca. 8 to 10 a.m. can be observed. This is followed by a decrease of the surface concentrations which reach their minima during early afternoon. Finally, a slight increase towards evening is found. Nevertheless, more overall variability and larger differences between the seasons are visible compared to HCHO.

The rather similar patterns of HCHO and glyoxal surface concentrations again indicate overall similar sources and characteristics of both species which were already outlined above, while the differences in the diurnal variations of the VCDs suggest different profile shapes. The latter might be caused by different chemical processing and lifetimes of both species. The reason for the different pattern observed for glyoxal during winter is not clear at the moment.

Figure 12.5 shows the annual characteristics of NO_2 VCDs and near-surface concentrations.

Both quantities show very similar annual patterns with highest NO_2 abundances during winter, where daily means exceeding 1×10^{16} molec cm^{-2} and 2×10^{11} molec cm^{-3} can be observed for VCDs and near-surface concentrations, respectively. This period of enhanced NO_2 starts in the late post-monsoon season and lasts until early summer where typically lowest NO_2 is measured. Also during the monsoon season only low NO_2 abundances prevail. In the latter two seasons, daily mean NO_2 VCDs and surface concentrations are typically around 0.5×10^{16} molec cm^{-2} and 0.75×10^{11} molec cm^{-3}, respectively. The effect of the monsoon cycle on the measured NO_2 abundances is weaker as compared to HCHO and glyoxal since only minor maxima are observed pre-monsoon. The post-monsoon maximum which occurred in October for HCHO and glyoxal is probably partly superposed by the strong increase of the NO_2 abundances towards winter. This less tight connection to the monsoon cycle is caused by shorter NO_2 lifetimes during the seasons in which the crop residue fires occur compared to the other seasons (more solar radiation). Since these fires do not take place at Mohali itself but in the rural surroundings (compare Sect. 4.3), the NO_2 is probably already diluted or degraded when these air masses reach the instrument location. Thus, only slight enhancements of NO_2 are observed in the corresponding seasons. HCHO and glyoxal, on the other hand, are mainly secondarily formed from precursors emitted by the fires and can be well detected by the instrument. The fires for domestic heating mainly occur during winter, where NO_2 lifetimes are typically longer leading to the observed maxima.

Compared to Bayfordbury and Mainz, the day-to-day variability of the NO_2 is apparently smaller. Reasons for this might be more stable wind patterns (compare Sect. 12.3) and the closeness of many strong NO_2 sources at Mohali. However, the maximum NO_2 pollution levels at Mohali seem to be slightly smaller compared to the other two stations.

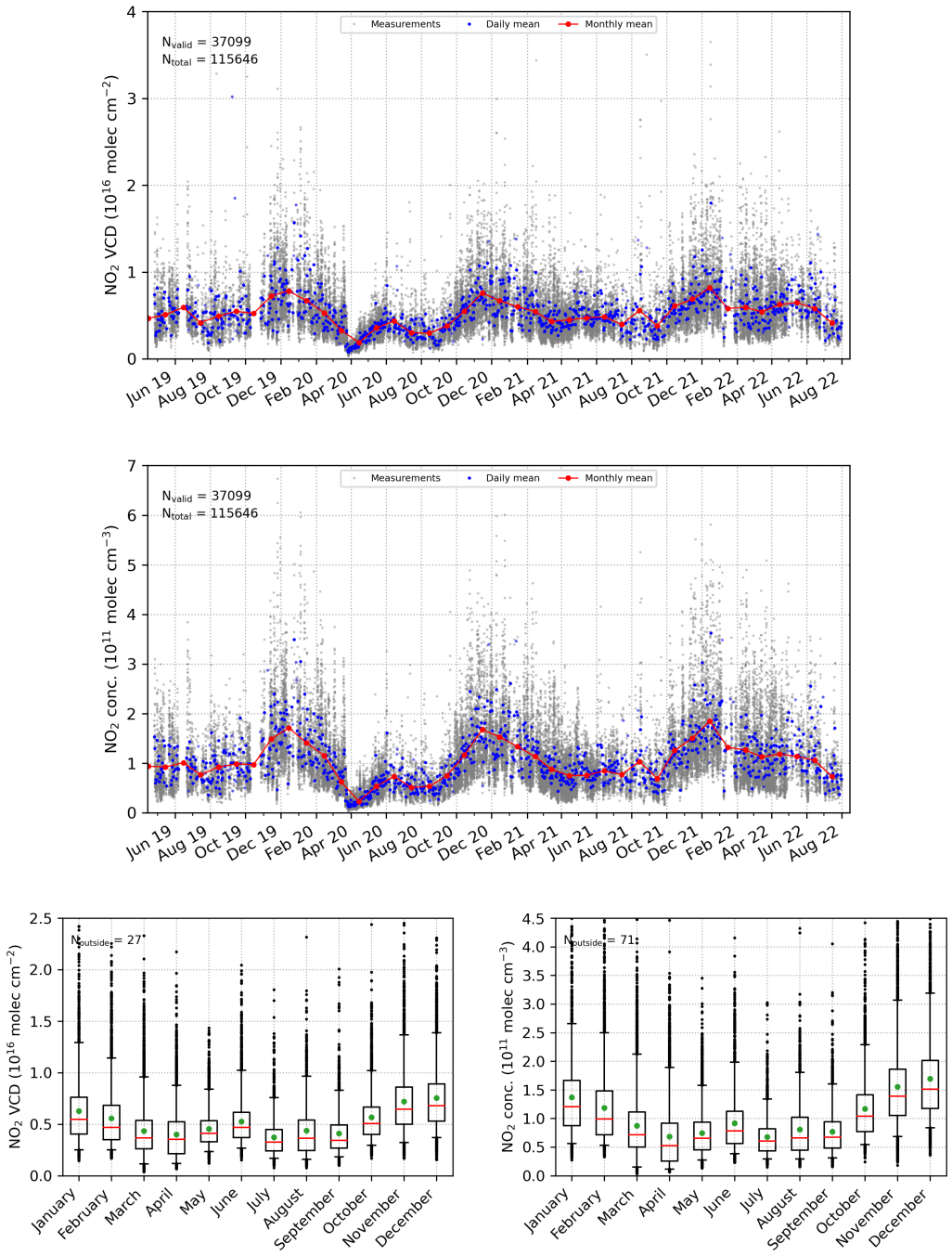


Figure 12.5: The two upper panels show time series of NO_2 VCDs and surface concentrations for Mohali. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 60 valid data points. The lower panels depict box-whisker-plots of the annual cycles of NO_2 VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

Lastly, it should be mentioned that, contrary to Bayfordbury and Mainz, a clear influence of the COVID-19 lockdown measures in March and April 2020 is found for Mohali, where daily mean NO_2 VCDs and surface concentrations as low as approximately 0.12×10^{16} molec cm^{-2} and 0.2×10^{11} molec cm^{-3} are observed, respectively. Also the NO_2 pollution levels in the pre-monsoon months of 2020 are slightly lower than in other years. A close look on the upper panels of Fig. 12.1 and 12.3 shows that also the HCHO and glyoxal abundances were slightly reduced during this period. This was not the case for Bayfordbury and Mainz suggesting that at those stations HCHO (and glyoxal) abundances are mainly of biogenic origin, while anthropogenic emissions have a major effect on the levels of both species at Mohali.

The average annual variations of NO_2 VCDs (left) and near-surface concentrations (right) are presented in the two lower panels of Fig. 12.5 in terms of monthly box-whisker-plots and basically confirm the previously outlined annual cycles with highest NO_2 abundances in winter and minimum values in early summer and during monsoon. The increase of the NO_2 pollution levels in the post-monsoon season transitions into the rise towards winter probably masking a distinct post-monsoon peak. The overall amplitude of the annual variations is stronger for the surface concentrations caused by slightly shallower profiles during winter compared to the rest of the year (see Fig. 12.21) and stronger vertical mixing during summer.

Overall the observed annual patterns are similar to the findings by Kumar et al. (2020), however, slightly higher NO_2 abundances are reported in their study. One reason for this might be successful pollution control measures leading to lower NO_2 levels. Further, Kumar et al. (2020) applied a fixed O_4 scaling factor of 0.8 which leads, as pointed out earlier, to slightly larger NO_2VCDs and near-surface concentrations, while the general patterns are not changed.

Figure 12.6 shows median diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for the different seasons.

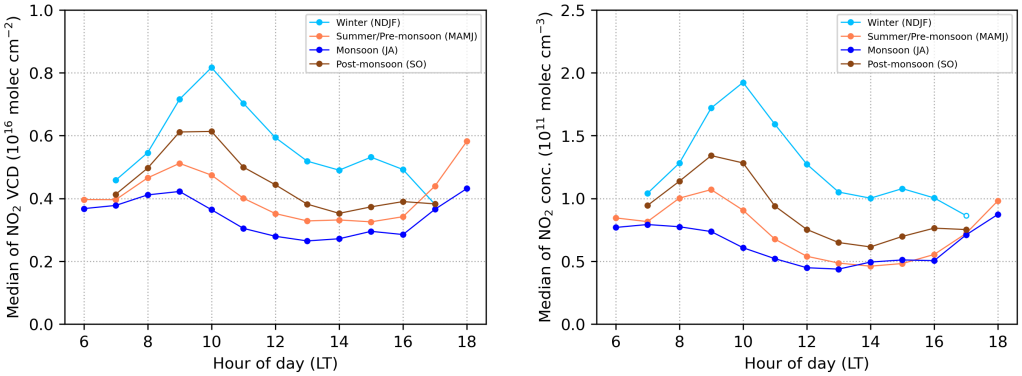


Figure 12.6: Diurnal cycles of NO_2 VCDs (left) and surface concentrations (right) for Mohali for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

For both quantities, very similar diurnal variations are found in all seasons, whereby the amplitudes of the patterns vary strongly between the individual

seasons. Further, these patterns are slightly stronger pronounced for the near-surface concentrations (larger relative amplitudes) which are strongly affected by local emissions. In all seasons, an increase of the measured NO_2 abundance is observed during the early morning which peaks in the late morning. The exact time at which the maximum is reached varies slightly between the seasons in the time range from 8 to 10 a.m., but the maximum occurs simultaneously for VCDs and near-surface concentrations within a season. The latter again stresses that the emission sources are located not too far away from the instrument. A similar behaviour was found at Mainz, while a time lag between the maxima of both quantities was observed for Bayfordbury where the major emission sources are located at ca. 30 km from the measurement site. After reaching their maximum values, VCDs and near-surface concentrations decrease significantly and show their minima in the early afternoon. Interestingly, both quantities reach their minima simultaneously at ca. 2 p.m. in all seasons. Towards sunset, a slight increase in the NO_2 abundances can be observed again. However, larger variability between the seasons and both quantities is found in the evening hours which is, at least partly, caused by statistical artefacts.

The similar patterns found for both quantities suggest that the diurnal patterns are mainly determined by the surface concentrations indicating emission sources rather close to the instrument together with a shallow vertical profile. The shape of the observed diurnal cycles and the finding that the diurnal patterns are stronger pronounced for the surface concentrations are caused by a combination of emission patterns, tropospheric chemistry and boundary layer dynamics. For instance, Kumar et al. (2020) reported that emissions from traffic and biomass burning for cooking and heating peak in the morning and evening hours. The minimum NO_2 lifetime is reached in the afternoon as OH concentrations peak which explains the decrease of NO_2 during the course of the day. Further, the rise of the boundary layer during the day leads to an extension of the layer into which the NO_2 is mixed. As a consequence, the change of the VCDs is smaller than that of the surface concentrations which show slightly larger (relative) amplitudes. Highest NO_2 abundances are observed in winter where also the largest amplitude is found. In these months, less variability in the boundary layer height and NO_2 lifetime throughout the day exists and the diurnal pattern is mainly determined by the emission patterns (Kumar et al., 2020). Largest boundary layer heights with strongest diurnal variations are expected in summer/pre-monsoon and indeed the largest difference in the relative amplitudes between VCDs and surface concentrations is found in this season. Lowest diurnal variations are found in the monsoon season where biomass burning ceases and also the other effects are only weakly pronounced. Overall these findings are consistent with the results from Kumar et al. (2020).

Finally, it should be stressed that, despite small differences in the details, the observed diurnal variations for NO_2 are similar to the diurnal cycles of the near-surface concentrations of HCHO and glyoxal. This finding indicates similar sources of all three species. On the other hand, the VCDs, especially of HCHO, show different diurnal variations probably caused by different lifetimes, different chemical processing and shallower profiles of NO_2.

Lastly, aerosol results are discussed whose annual cycles are summarised in Fig. 12.7 which also includes AERONET AODs for comparison. However, it should be mentioned that no AERONET station is located in the vicinity of the Mohali instrument and the used AERONET data were taken from the station at the Amity University Gurugram in Haryana located approximately 45 km south-west of New Delhi and around 250 km south of Mohali itself (compare Sect. 5.2). As a consequence, no good agreement can be expected.

The annual variations of the AOD (uppermost panel) are less clear compared to the species discussed before. Distinct minima occur only during early summer (March to May) in most years, while rather high AODs are observed during the rest of the year where the values scatter around values close to 0.5 and show a large variability. In some monsoon seasons, e.g. in 2020, local minima seem to exist. Slightly enhanced AODs are observed during winters. Obviously, no good agreement with the AERONET data is given due to the reasons outlined before. Nevertheless, also the AERONET AODs show enhancements during winter. Lastly, an influence of the COVID-19 lockdown measures is also visible for the AODs in March and April 2020 where values of only around 0.15 are found. The near-surface aerosol extinctions at 367 nm (middle panel), however, show clear annual cycles with highest extinctions during winter and lowest values in early summer. Local maxima are found pre-monsoon and partly also in the monsoon season. No obvious enhancements are found post-monsoon. As seen for NO_2, these might be hidden by the strong increase of the near-surface quantities towards winter.

The monthly box-whisker-plots of AODs (left) and near-surface aerosol extinctions (right) mainly confirm the previous findings. A weak average annual cycle is found for the AODs with lowest values in early summer, where the monthly mean is 0.27 in April. However, it should be noted that the mean of April is probably affected by the low values in 2020 due to lockdown measures. Higher AODs are obtained in the rest of the year (monthly means of ca. 0.45) with maximum monthly means of or slightly below 0.5 in winter (November to February). No clear maxima can be observed before and after the monsoon season. Further, no clear tendency is visible during the monsoon (July and August) where the variability is rather large as indicated by the extensions of the boxes. Similar to the annual patterns of the NO_2 near-surface concentrations, a pronounced annual cycle is found for the near-surface extinctions. Highest values are obtained during winter, whereby the maximum occurs in December with a mean extinction of ca. $1\,km^{-1}$. Again the lowest aerosol content is observed during early summer (monthly mean in April: $0.27\,km^{-1}$). Slight enhancements are found pre-monsoon while no distinct behaviour is observed during and post-monsoon. The latter again is hidden by the strong increase of the near-surface extinctions towards winter.

The observed annual patterns of the aerosol results, more particularly the near-surface extinctions, are closely related to the annual cycles of biomass burning and private domestic fires which is also indicated by the similarity of these patterns to the ones of the NO_2 surface concentrations. While biomass burning peaks in the pre- and post-monsoon periods, private fires are most important

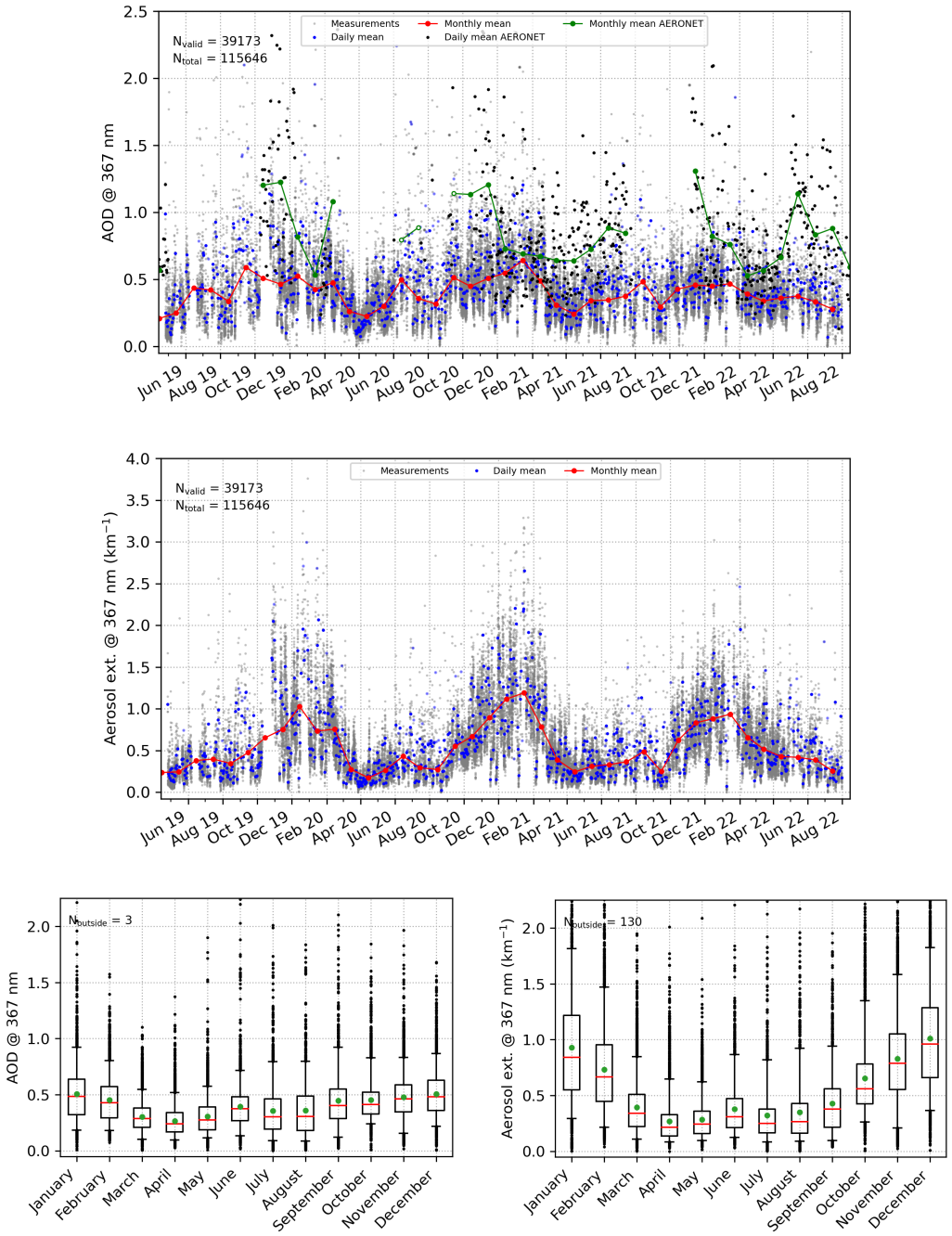


Figure 12.7: The two upper panels show time series of AOD and surface aerosol extinction at 367 nm for Mohali. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 60 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and surface aerosol extinction at 367 nm (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

during winter leading to severe fog and haze events. The latter are indicated by the largely enhanced near-surface aerosol extinctions during winter. Further, the importance of boundary layer dynamics is evident. While most of the pollution is trapped within a very shallow boundary layer during winter, the aerosol is mixed into higher layers during the warmer period of the year leading to an only weakly pronounced pre-monsoon peak in the near-surface extinctions. The fact that the annual cycle is almost not visible for the AODs indicates the strong effect of local emissions.

Compared to the other stations, two to three times higher AODs are observed at Mohali, whereas the near-surface aerosol extinctions can be up to ca. five times higher during winter compared to Bayfordbury and Mainz. This is caused by the large effect of biomass burning and the low quality of the materials burned in the domestic fires. Further, the aerosol at Bayfordbury and Mainz is mainly traffic and industry related, where nowadays typically particle filters are used.

For Mohali, the aerosol results obtained with a fixed O_4 SF of 0.8 yield slightly different annual patterns for the AODs (compare Fig. B.15 in the appendix). Their minimum is still found during early summer and also during winter enhanced AODs are observed. However, using an SF of 0.8 yields highest AODs during monsoon and in the early post-monsoon season, whereby also the scatter of the results is largely increased. It should be mentioned that the annual patterns of the near-surface extinctions are almost unchanged. For the latter, only the absolute values and the scatter are increased. Similar to the previous stations, larger values are obtained with the fixed SF for both quantities.

Kumar et al. (2020) applied an SF of 0.8 and found maximum AODs in the monsoon season with values close to 1. The values obtained in this work are significantly smaller also for the fixed SF (ca. 0.75). Kumar et al. (2020) performed several investigations whether the observed patterns are real or caused by artefacts and concluded that the maximum during the monsoon might be caused by hygroscopic growth of aerosol particles in a very humid environment. Although this conclusion is reasonable, the monsoon season is typically associated with high cloud cover leading to larger uncertainties in the MAX-DOAS results than in other seasons. Further, at Mohali no local AERONET station is available which could otherwise help to confirm these findings and the unexpected results in the diurnal patterns for the fixed SF (see below). Therefore, in this work it is left at descriptions of the observed effects and no detailed discussions are provided.

The diurnal cycles of AOD (left) and near-surface aerosol extinction (right) at 367 nm observed for the different seasons are summarised in Fig. 12.8.

Almost no diurnal variations are found for the AODs, where only marginal increases over the course of the day can be observed in summer, during the monsoon and in the post-monsoon season. No consistent pattern is found in winter. For the near-surface concentrations, however, diurnal patterns similar to the ones of the near-surface concentrations of HCHO, glyoxal (in some seasons) and NO_2 are found in all seasons, showing an increase during the morning, a peak in the late morning and a decrease afterwards. The time of the peak seems to be slightly shifted compared to the trace gases. A clear increase towards sunset is only found during summer and monsoon. Also during monsoon, the

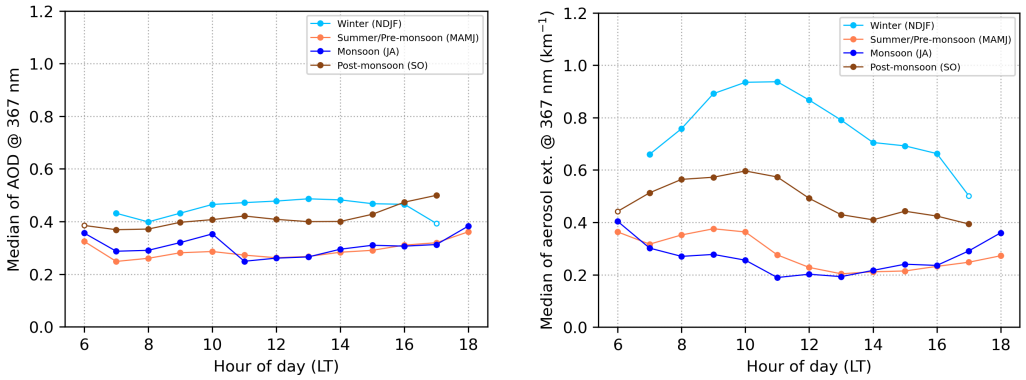


Figure 12.8: Diurnal cycles of AOD (left) and surface aerosol extinction (right) at 367 nm for Mohali for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

increase during morning is not observed, where instead a decrease is present. Note that the values in the early morning have to be considered with care. Like seen for NO_2, the largest amplitude is found during winter, whereas the smallest amplitude occurs during monsoon.

The similarity of the diurnal patterns of the near-surface extinctions to the ones of the NO_2 surface concentrations again indicates common anthropogenic sources which peak in the morning. During monsoon, temperatures are rather high (no heating needed) and generally biomass burning activities are low. As a consequence, no strong enhancements occur in the morning. The decrease of the near-surface extinctions towards afternoon together with slightly increasing AODs is caused by a rising boundary layer which distributes the pollution from the morning hours within the lower tropospheric column over the course of the day. During winter, the rise of the boundary layer is weaker and the aerosol stays in the lowermost atmospheric layers during the whole day which does not lead to an enhancement of the AODs and surface extinctions above the ones from the other seasons are found throughout the day.

Significantly different diurnal variations of the aerosol results are obtained using the fixed O_4 SF. Figure 12.9 shows the corresponding diurnal cycles of AOD (left) and near-surface extinction (right).

Especially, the AOD shows an unexpected and strongly pronounced inverse "U-shaped" diurnal cycle which peaks around noon. While this pattern is very strong during monsoon and post-monsoon, it is weaker in winter and summer, yet still visible. For the near-surface extinctions, the differences between the diurnal patterns obtained with the two SF scenarios are smaller. Nevertheless, the observed diurnal pattern for the fixed SF is still superposed by a strong increase towards noon, especially in summer and during monsoon.

The reasons for these differences in the annual and diurnal cycles are still not clear. No such systematic differences occurred for Bayfordbury and Mainz, while a similar behaviour is found for the ATTO measurements presented in Chapt. 13. Possibly, the variable SF compensates for effects that depend on SZA and RAA.

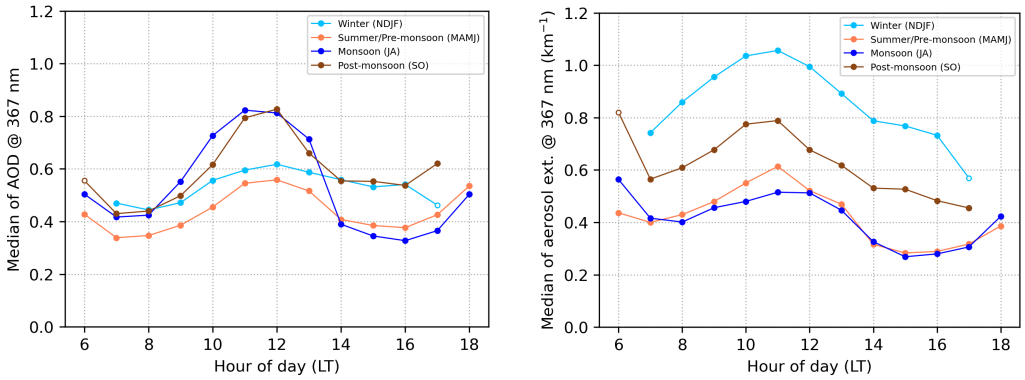


Figure 12.9: Same as Fig. 12.8 but for a fixed O_4 SF of 0.8 applied for the aerosol retrieval.

Further, inaccuracies of the LUTs used for the MAPA retrieval are possible. Also the used parametrisation of the phase-function describing aerosol scattering might contribute to the observed effects. For the calculations of the LUTs, the phase-function is parametrised using the Henyey–Greenstein parametrisation. This parametrisation does not fully capture the forward peak of aerosol scattering which becomes important at small RAAs and SZAs, i.e the instrument points close to the sun (Beirle et al., 2019). The latter might be particularly important for stations at lower latitudes (such as Mohali and ATTO), where the sun is higher during noon. In principle, MAPA includes a flag to account for this, however, as mentioned, the variable SF which is an additional degree of freedom for the retrieval might compensate for such inaccuracies of the assumed phase-function. Hence, small scattering angles in theory might be still problematic for low latitude stations like Mohali in case of a fixed O_4 scaling factor. However, the instrument at Mohali points towards north (350°) and low elevations are thus probably not affected since the sun is always south of the instrument. Further, the effect is smallest during summer where the SZA is smallest. Consequently, also the zenith measurements are most likely not affected by too small scattering angles. Kumar et al. (2020) applied a fixed SF of 0.8 and therefore obtained similar patterns as shown in the left panel of Fig. 12.9 and attributed them to the formation of SOA from VOCs and boundary layer dynamics. In summary, no conclusive explanation for the observed effect can be given here. Therefore, it is concluded that the increased values around noon are probably an artefact which occurs when a fixed SF is used for the MAPA retrieval at Mohali. This additionally justifies to present the results for a variable SF in the main part of this work (compare Sect. 9.3) since the results of the other atmospheric investigations for Mohali are not affected by the choice of the O_4 SF.

Apart from the species whose results were presented in detail in this section, also HONO and water vapour retrievals were performed. Similar to Mainz, interesting results were found for HONO which are presented in Sect. 12.7. For water vapour, expected results showing a very strong annual cycle with highest values in the monsoon season and lowest atmospheric water vapour content during winter are obtained. Lastly, almost no diurnal cycles are found for H_2O.

12.3 INFLUENCE OF METEOROLOGICAL CONDITIONS

In this section, the dependencies of the trace gas and aerosol abundances measured at Mohali on air temperature, wind speed and wind direction are studied. Again MAX-DOAS data were selected as described in Sect. 9.3 and hourly means were used to minimise the scatter of the data.

Hourly mean HCHO VCDs (left) and near-surface concentrations (right) as function of temperature are depicted in the uppermost panels of Fig. 12.10.

A weak (non-linear) temperature dependence is visible for the VCDs. Compared to Bayfordbury and Mainz, more variability is found at Mohali and values above 1×10^{16} molec cm^{-2} can occur at all prevailing temperatures. No clear dependence on temperature can be observed for the near-surface concentrations. Although the majority of the (valid) measurements lie in a temperature range from ca. 25 to 30 °C (indicated be the high point density in this temperature range), elevated near-surface concentrations can occur rather independently of the air temperature. The observed temperature dependencies of both quantities, including the differences between them, can be explained by the strong influence of anthropogenic activities, mainly crop residue fires, on the HCHO abundances at Mohali. These anthropogenic activities are closely related to the monsoon cycle with maxima of the HCHO abundances in the pre- and post-monsoon seasons (compare Fig. 12.1). However, the monsoon cycle and thus the anthropogenic activities are not synchronous with the annual variations of the air temperatures at Mohali (compare Fig. 4.14). In particular, the pre-monsoon maximum of the HCHO abundances in June is associated with the highest air temperatures while the post-monsoon maximum with similar HCHO VCDs and maximum surface concentrations in October and November occurs at already substantially lower temperatures. Further, the clearly reduced HCHO levels during the monsoon are associated with temperatures comparable to the pre-monsoon temperatures. The differences between the temperature dependencies of VCDs and near-surface concentrations originate from their slightly different annual variations (compare above). The weak non-linearity of the temperature dependence of the VCDs indicates an additional effect of biogenic sources of HCHO which seem to not affect the surface concentrations. Nevertheless, overall also the findings for the temperature dependencies of HCHO VCDs and near-surface concentrations suggest that anthropogenic activities, in particular crop residue fires, are of major importance for the HCHO abundances at Mohali.

The dependencies of hourly mean HCHO VCDs (left) and near-surface concentrations (right) measured at Mohali on wind speed and wind direction are shown in the middle and lower panels of Fig. 12.10, respectively.

Overall, the angular pattern is rather homogenous for the VCDs and does not show strong gradients. Slightly enhanced VCDs occur under north-easterly (sector from ca. 20 to 70°) and south-south-easterly to west-north-westerly (sector from ca. 150 to 290°) winds, where median values faintly above 1.25×10^{16} molec cm^{-2} are obtained. In the remaining sectors, the VCDs are slightly lower and values in the order of 1.1×10^{16} molec cm^{-2} are observed. More variability and stronger gradients can be observed for the near-surface concentrations which show their

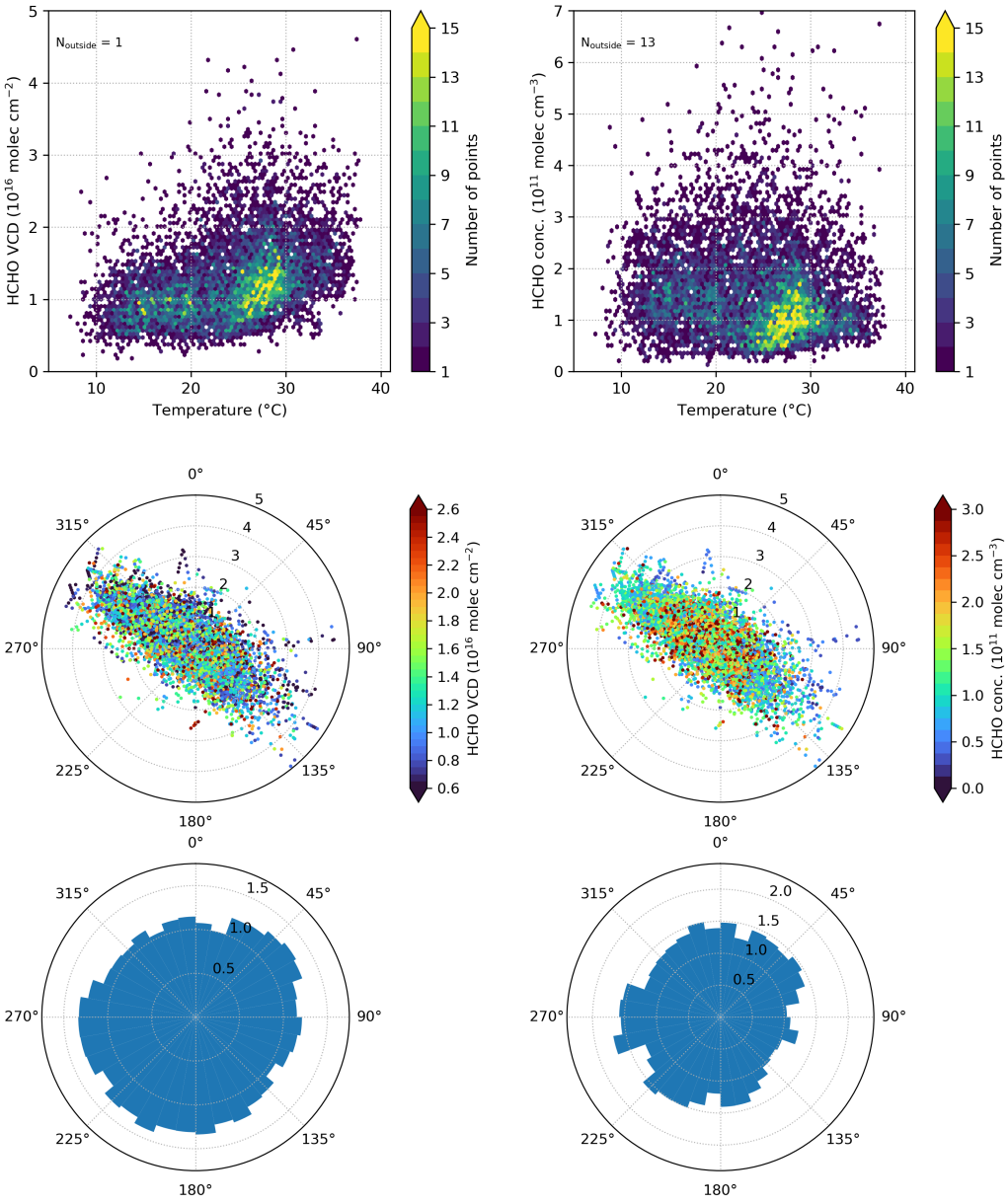


Figure 12.10: The two uppermost panels show hourly mean HCHO VCDs (left) and surface concentrations (right) measured at Mohali as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HCHO VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HCHO VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

absolute maximum under westerly (sector from ca. 250 to 290°) winds, where median values slightly above 1.5×10^{11} molec cm^{-3} are found. Less strong enhancements with median surface concentrations faintly below 1.5×10^{11} molec cm^{-3} are found under south-south-easterly to south-westerly winds (sector from ca. 150 to 230°) and, similar to the VCDs, under north-easterly winds (sector from

ca. 20 to 70°). In the remaining sectors, lower medians are obtained, whereby the absolute minimum is observed under easterly to south-easterly winds in a sector from ca. 100 to 150° with values of only approximately 1×10^{11} molec cm^{-3}. However, no such distinct minimum is found for the VCDs.

The maximum under north-easterly winds found for both quantities shows the influence of the *"Tricity"*, which surrounds the measurement site in the sector from north to east, and its emissions on the HCHO abundances measured at Mohali. More specifically, Chandigarh is located north-east of IISER (compare map in Fig. 4.13). The finding that the VCDs are stronger enhanced than the near-surface concentrations suggests the emission of mostly precursor substances producing HCHO which affect the surface concentrations less. The enhanced near-surface concentrations accompanied with elevated VCDs under westerly winds are probably caused by two factors. Firstly, in this sector mainly agricultural fields (and smaller towns) are situated. In these areas, crop residue fires and private fires take place which heavily affect the air quality at Mohali and mostly enhance the near-surface concentrations. Secondly, as supported by the middle panels of Fig. 12.10, the wind speeds under straight westerly winds are typically low leading to high surface concentrations and also elevated VCDs. Nevertheless, the effect of wind speed is apparently smaller for the VCDs. Similar conclusions hold for south-south-easterly to south-westerly winds, however, winds from these directions occur only rarely as shown in panel (b) of Fig. 4.14. Under winds from north-west, which is the primary wind direction during large parts of the year, also comparably high wind speeds can occur. These are often associated with lower HCHO near-surface concentrations caused by enhanced mixing, while the VCDs are almost not affected by this. Further, an influence of long-range transport of pollution from cities like Lahore (ca. 250 km), Pakistan, or the Asian continent in general cannot be ruled out under winds from these north-westerly sectors, which might also explain why the VCDs are not reduced under these wind directions. Although New Delhi is located at a similar distance to the south, its influence is probably less important since wind speeds are clearly lower under southerly winds and emissions of precursors at sources that far away thus have little impact on HCHO at Mohali. As described in Sect. 4.3, the main fetch region is south-east during the summer monsoon season. Although a mix of agricultural lands and smaller towns is located in this wind sector, biomass burning activity is low during this season of the year and heavy rainfall under cloudy sky conditions as well as rather high wind speeds occur frequently. As a consequence, the surface concentrations show a clear minimum under winds from these directions. Contrary to that, the VCDs do not show a distinct minimum under south-easterly winds. Similar findings were already obtained for the annual cycles, where the surface concentrations showed their absolute minimum during monsoon while the VCDs were less reduced (compare Fig. 12.1). As outlined there, the latter is probably caused by biogenic emissions as temperatures are still rather high during the monsoon season.

Figure 12.11 shows the dependence of glyoxal VCDs (left) and near-surface concentrations (right) measured at Mohali on temperature (uppermost panels) and wind speed and direction (middle and lowermost panels).

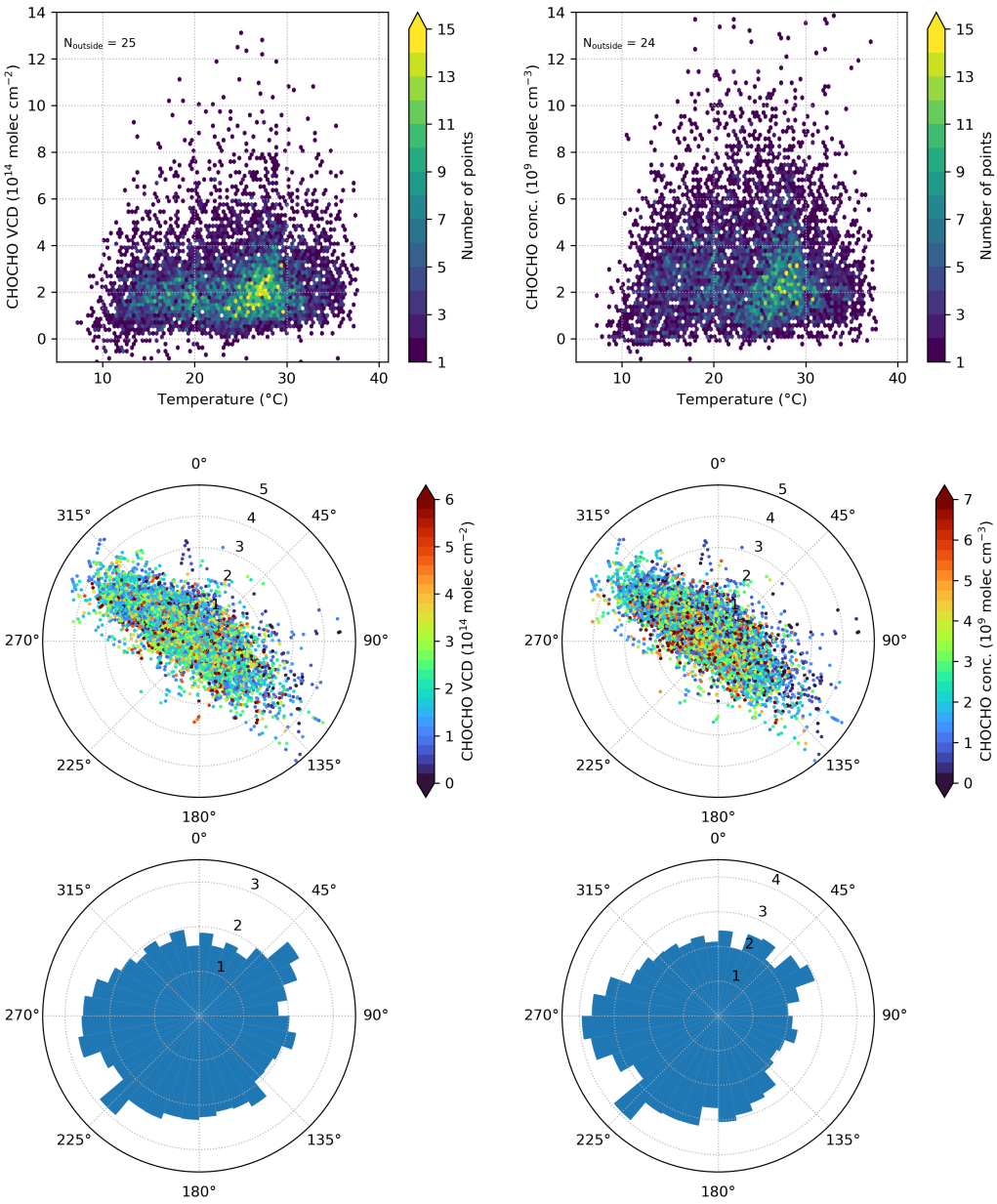


Figure 12.11: The two uppermost panels show hourly mean glyoxal VCDs (left) and surface concentrations (right) measured at Mohali as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean glyoxal VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median glyoxal VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

Only a weak temperature dependence is found for the glyoxal VCDs with a faint tendency of generally larger values at higher temperatures, while practically no dependence on temperature is visible for the near-surface concentrations. Overall, these findings are comparable to the ones obtained for HCHO indicating similar sources and atmospheric processing of both species. However, compared

to HCHO even more variability is found in the temperature dependence of the VCDs leading to a pattern that is more similar to the one found for the surface concentrations. This is consistent with the observed annual cylces which showed more similar annual variations of both quantities for glyoxal as compared to HCHO indicating different profile shapes of both trace gases. Similar to HCHO, the observed patterns in the temperature dependencies of both quantities, including the (small) differences between them, can be explained by an interplay of the annual cycles of anthropogenic activities (biomass burning and private fires) and air temperatures at Mohali. For a more detailed explanation, refer to the corresponding paragraph above.

The observed angular patterns are generally comparable to the ones found for HCHO and enhanced glyoxal abundances occur under winds from north-east and from a sector spanning from south-easterly to west-north-westerly winds, while lower abundances are observed in the remaining wind sectors (compare lowermost panel of Fig. 12.11). However, compared to HCHO the angular patterns are less homogenous and the sectors showing enhancements seem to be more confined, especially for the VCDs, whereby the angular patterns of both quantities are more similar for glyoxal. In particular, glyoxal VCDs and near-surface concentrations show highest median values under winds from more confined sectors spanning from ca. 40 to 80° and 250 to 290°. In the latter sector, median VCDs and near-surface concentrations are around 2.65×10^{14} molec cm^{-2} and almost 3×10^{9} molec cm^{-3}, respectively. Nevertheless, slightly enhanced glyoxal abundances are also observed under winds from south-south-east to south-west which was also observed for HCHO. No special pattern occurs under north-westerly winds where both quantities show medium values. The single very pronounced enhancement in the 225°-bin might indicate an effect of the Rajpura Thermal Power Plant (compare Sect. 4.3). Also the other species show single enhancements of different strengths in this sector, mainly for the respective surface quantities (compare polar plots of the respective species in this section). Lastly, a rather distinct minimum of the glyoxal surface concentrations can be observed under east-south-easterly winds originating from a sector spanning from about 90 to 140°, corresponding to the wind direction in the monsoon season. Although the overall patterns are more similar for both quantities as compared to HCHO, no such distinct minimum is observed for the VCDs which are only marginally reduced in this sector. Again this can probably be explained by biogenic sources of glyoxal which are still active during the monsoon season as temperatures are still rather high.

Like for HCHO, the observed angular patterns can be explained with the location of the different sources relative to the measurement site in combination with wind speeds and emission type. In particular, the enhancements under north-easterly winds which are stronger pronounced for the VCDs than for surface concentrations can be attributed to the city of Chandigarh, while biomass burning activities and domestic fires can be identified as sources of the enhancements under south-easterly to west-north-westerly winds. Especially for the latter, substructures exist in the angular patterns, e.g. under straight westerly winds. A possible explanation for these substructures might be in general shallower

glyoxal profiles for which the exact location of the biomass burning events is better resolved by the angular distribution while for higher reaching profiles, like e.g. for HCHO, this effect might be smeared out. The latter is supported by the fact that also the HCHO near-surface concentrations showed this effect, while no such structures were observed for the corresponding VCDs. Also the typically lower wind speeds under westerly winds affect the angular patterns. For more detailed descriptions of the different sources, it is referred to the explanations for HCHO above since both species show in general very similar dependencies.

NO_2 VCDs (left) and near-surface concentrations (right) as function of air temperature are shown in the two uppermost panels of Fig. 12.12.

Both quantities show a weak temperature dependence with generally larger values at comparably low temperatures especially for the near-surface concentrations, whereby the absolutely highest values occur at temperatures below ca. 20 °C. The temperature dependence of the VCDs is apparently less pronounced. However, also the VCDs exhibit a tendency towards slightly larger values at temperatures below and above approximately 20 and 30 °C, respectively, while the lowest VCDs are found in the temperature range in between. These findings are consistent with the annual cycles of both quantities presented in Fig. 12.5. In particular, highest NO_2 abundances originating from private fires for domestic heating are found in winter when the temperatures are at their minimum, whereas the lowest pollution levels are observed during early summer and monsoon at moderate temperatures below 30 °C. The local NO_2 maxima for both quantities pre-monsoon caused by crop residue fires coincide with the highest temperatures at Mohali leading to the slight increase, especially for the VCDs, at temperatures above 30 °C. Lastly, the enhancement of the NO_2 abundances during the winter months is stronger for the surface concentrations resulting in a slightly more pronounced increase of the latter towards lower temperatures.

The middle and lower panels of Fig. 12.12 summarise the dependence of the measured NO_2 VCDs (left) and surface concentrations (right) at Mohali on wind speed and direction.

High NO_2 VCDs are observed under winds from a wide sector spanning (clockwise) from south to east, where median values vary around 0.5×10^{16} molec cm^{-2}. Significantly lower VCDs are only observed under easterly to southerly winds (ca. 100 to 150°) being the dominant wind direction in the monsoon season which is associated with reduced NO_2 abundances. Particular enhancements of the NO_2 VCDs can be observed in three more confined wind sectors, namely 250 to 290°, 310 to 340° and 10 to 80°. In these sectors, the obtained median VCDs are mostly slightly above 0.5×10^{16} molec cm^{-2}. The angular patterns are very similar for the surface concentrations indicating sources rather close to the measurement site. However, differences in the details can be observed. More specifically, the particular enhancements in the three confined wind sectors mentioned before as well as the reduction under wind directions dominating in the monsoon season (ca. 100 to 150°) are stronger pronounced compared to the VCDs. Further, no enhancements can be observed under southerly winds for the near-surface concentrations.

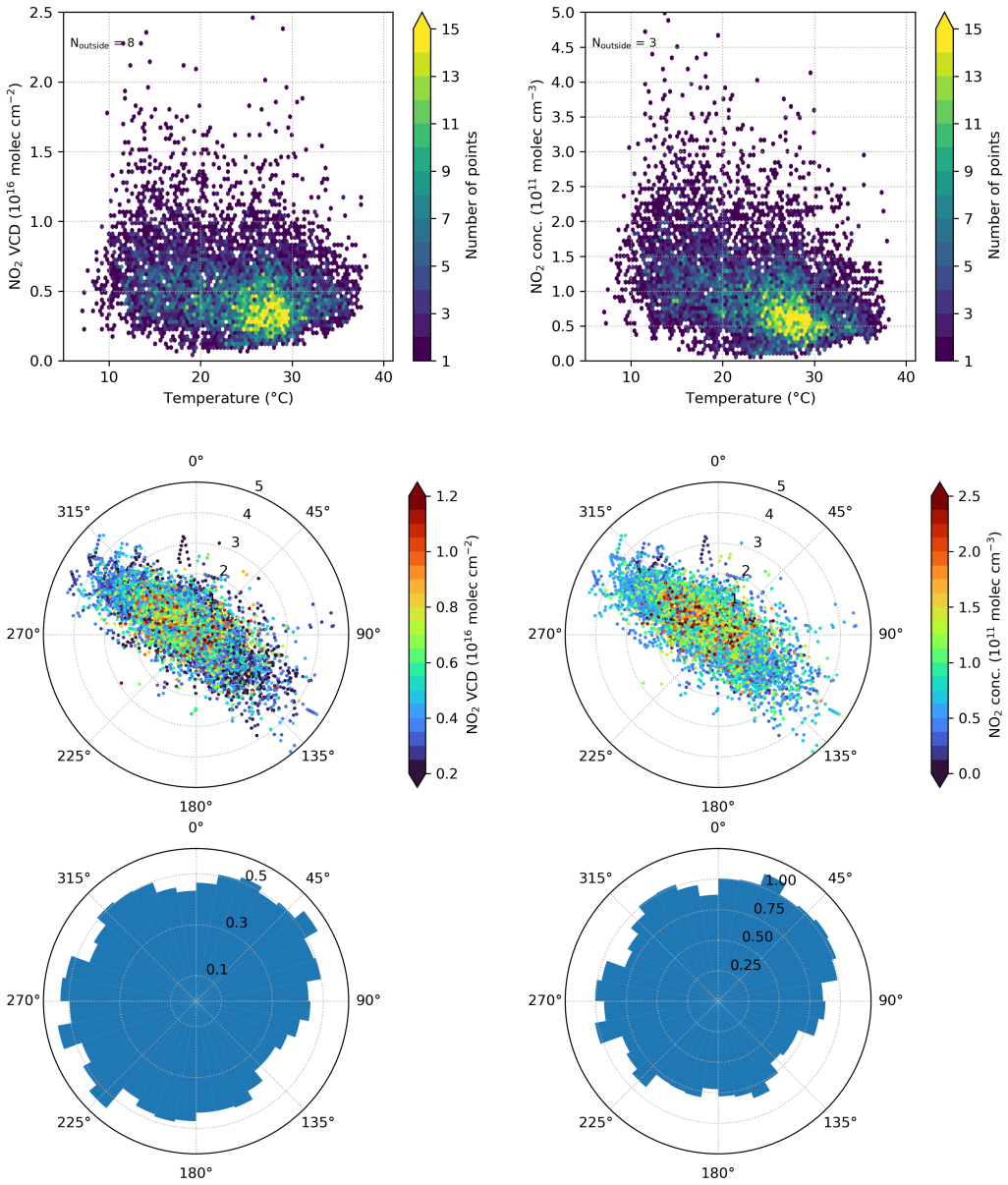


Figure 12.12: The two uppermost panels show hourly mean NO_2 VCDs (left) and surface concentrations (right) measured at Mohali as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean NO_2 VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median NO_2 VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

The location of the Mohali measurement site very close to the densely populated *"Tricity"* of Chandigarh becomes evident by the generally high NO_2 pollution levels occurring under winds from a wide sector spanning (clockwise) from south to east in which the metropolitan city surrounds the measurement site (compare Fig. 4.13). The specific enhancements under westerly winds (ca. 250 to

290°) which are found for both quantities and were also observed for HCHO and glyoxal can be mainly attributed to emissions originating from biomass burning activities and private fires for domestic heating and cooking. In addition, low wind speeds, occurring frequently under westerly winds, are typically associated with rather high NO_2 abundances (compare middle panels) which probably also explains parts of the enhancements under such westerly winds. Under north-westerly and north-easterly winds the measurement site is strongly influenced by pollution originating from the cities of Mohali and Chandigarh, respectively. The enhancements under southerly wind directions in principle also suggest an effect of biomass burning since also in this sector agricultural lands and smaller towns are located (compare Sect. 4.3), however, unlike to the westerly wind sectors, no enhancements are found for the near-surface concentrations in these sectors. This finding could indicate long-range transport of pollution from New Delhi which probably would not affect the surface concentrations strongly, while an effect might be visible in the VCDs. However, given the rather low wind speeds under southerly winds (compare also paragraph regarding HCHO) this remains speculative. Thus, the reason for the difference in the patterns between both quantities under southerly wind directions cannot be fully revealed. Despite this, the slight differences in the angular patterns obtained for both quantities can probably be explained by the specific position of the potential sources since surface concentrations are stronger affected by emissions rather close to the instrument, while VCDs are potentially also enhanced by emissions from sources farther away. The low pollution levels during monsoon are consistently reflected by the angular distributions of both quantities which show low NO_2 levels under easterly to south-easterly winds.

The annual cycles of the NO_2 abundances presented in Fig. 12.5 suggested less day-to-day variability of the NO_2 levels compared to Bayfordbury and Mainz. This is supported by the observed angular distributions of NO_2 at Mohali since similar NO_2 abundances, especially VCDs, are observed under winds from a wide sector as described above.

Apart from differences in the details, the described patterns of the angular distributions of the NO_2 abundances are in principle quite similar to the ones found for HCHO and glyoxal, stressing once more the strong influence of anthropogenic emissions on the HCHO and glyoxal levels at Mohali. However, NO_2 is emitted by almost all anthropogenic activities while HCHO and glyoxal have typically more specific sources. Consequently, high NO_2 abundances are observed in wider angular ranges compared to HCHO and glyoxal.

Lastly, the dependence of the aerosol quantities at 367 nm on meteorological conditions is investigated. The uppermost panels of Fig. 12.13 show AOD (left) and near-surface aerosol extinction (right) as function of temperature.

Reflecting the only weakly pronounced annual cycle, practically no temperature dependence is found for the AOD. The relatively large variability during monsoon is indicated by the outliers occurring at temperatures between 25 and 30 °C, while the slightly enhanced AODs during winter are represented by a few large values at temperatures close to 10 °C. For the near-surface aerosol extinctions, a rather distinct temperature dependence with highest values being observed

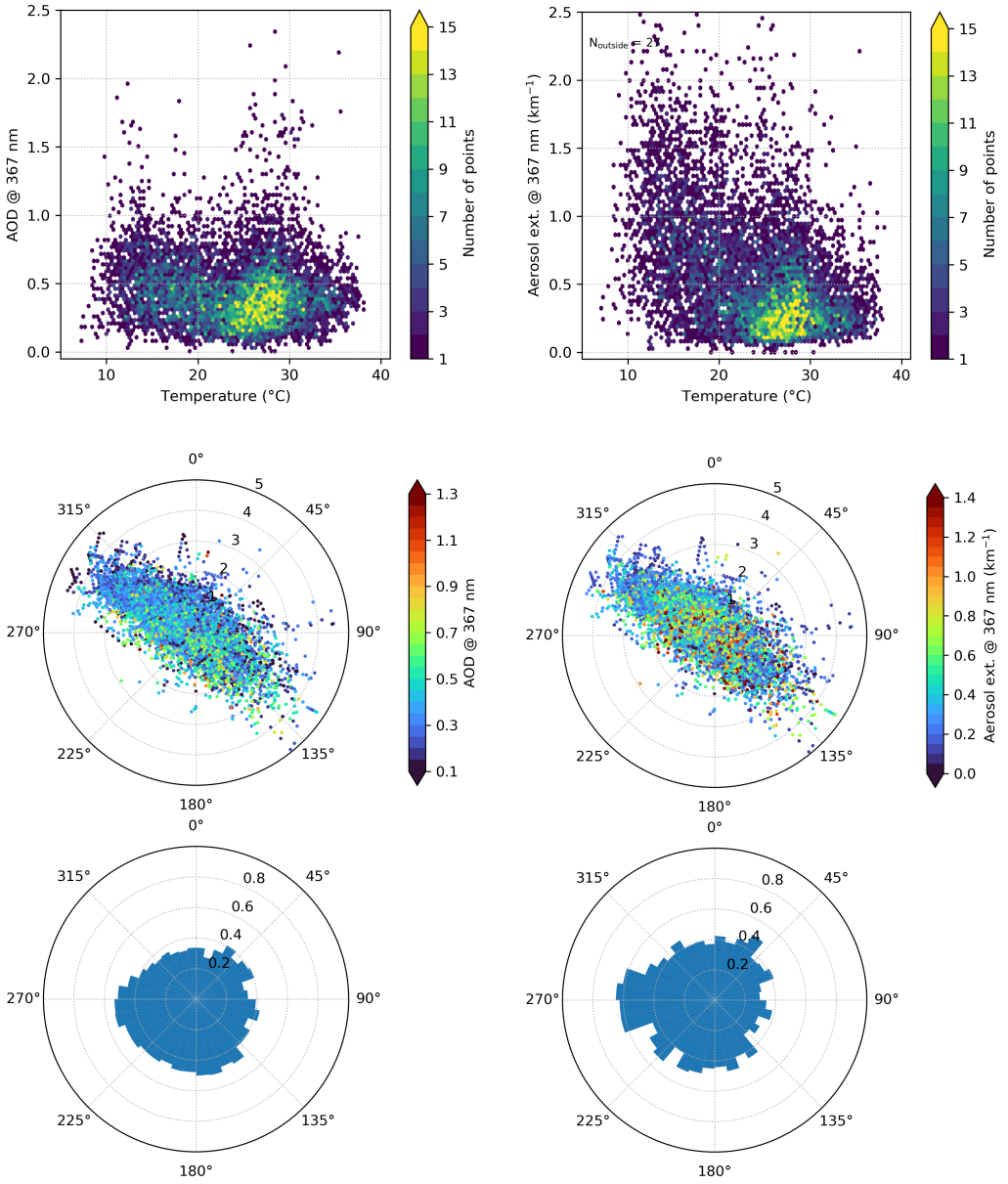


Figure 12.13: The two uppermost panels show hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) measured at Mohali as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean AODs (left) and surface aerosol extinctions at 367 nm (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median AODs (left) and surface aerosol extinctions at 367 nm (right) obtained under different wind directions in bins of 10° are indicated by the radius.

at low temperatures is found. Also this is consistent to the observed annual cycle which showed strongly enhanced near-surface extinctions during winter caused by private fires in combination with a shallow boundary layer. The observed temperature dependence is similar to the one obtained for the NO_2

surface concentrations but stronger pronounced which fits the finding that also the relative amplitude of the annual cycle is larger for the near-surface aerosol extinctions. The slightly enhanced values of the latter during the pre-monsoon season (June) at high temperatures are only weakly reflected by the temperature dependence. In particular, the minimum near-surface extinctions show a faint increase at the highest temperatures. However, also at low temperatures the minimum near-surface extinctions are marginally increased caused by the factors mentioned before. Lastly, it should be mentioned that for a fixed O_4 SF of 0.8 a slightly different pattern is found for the AODs, while the patterns are practically unchanged for the near-surface concentrations which only show an increased variability. The annual cycles of the AODs obtained using the fixed SF mainly showed a difference during monsoon, where the annual maximum AODs are observed for the SF of 0.8. Further, the scatter of the results was increased. Consistent to these findings, the temperature dependence of the AODs for a fixed O_4 SF of 0.8 shows more variability and largest values are found at temperatures between 25 and 30 °C which correspond to typical temperatures during monsoon (compare Fig. 4.14).

The dependencies of the retrieved AODs (left) and near-surface aerosol extinctions (right) at 367 nm on wind speed and direction are presented in the middle and lowermost panels of Fig. 12.13.

Elevated AODs with values above 0.4 occur under winds originating from a sector spanning from south-east to west-north-west (140 to 290°), whereby the absolutely highest AODs are observed under westerly winds in a more confined sector spanning from ca. 250 to 280° where medians slightly above 0.5 are found. In the remaining sectors, median AODs of mostly below 0.4 are obtained. The angular patterns of the near-surface aerosol extinctions show more variability, nevertheless, enhanced extinctions are also found under south-easterly to west-north-westerly winds. While the surface extinctions are above 0.4 km^{-1} in these sectors, highest values exceeding 0.6 km^{-1} are found under winds originating from the west. Smaller enhancements are also observed under north-easterly winds, whereas the AODs show only marginal enhancements there. Lowest near-surface extinctions clearly below 0.4 km^{-1} occur under easterly to south-easterly winds. Consistent to the annual cycles presented above, the latter indicates lower near-surface aerosol loads during monsoon which is almost not visible for the AODs. Lastly, low wind speeds are typically associated with higher values which can be observed especially for the near-surface extinctions.

The enhancements of both quantities under south-easterly to west-north-westerly winds can be attributed to biomass burning activities and private fires since mainly agricultural fields and smaller towns surround the measurement site in this sector (compare Sect. 4.3). The particular enhancements under straight westerly winds (ca. 250 to 280°) which are especially pronounced for the near-surface extinctions were already observed for the other species in this section. No specific source can be identified in this direction, but probably fires (biomass burning and domestic) occur closer to the measurement site in this more confined sector. Further, wind speeds are typically lower under westerly winds (compare middle panels) leading to enhanced near-surface quantities as outlined earlier.

Interestingly, the city emissions have no strong effect on the observed aerosol abundances, whereas the NO_2 abundances were strongly affected. Only the near-surface extinctions show small enhancements under north-easterly winds, when air masses influenced by Chandigarh are probed by the instrument. The AODs, however, are only marginally affected by this. These findings can probably be explained by the specific sources which are fires under westerly winds while the city emissions are more traffic related. Also domestic heating and cooking are probably "cleaner" in the city compared to the rural regions.

Finally, consistent to the observations made above, the angular patterns of the AODs change when the results for the fixed O_4 SF are considered. Especially, the difference between AODs observed under south-westerly and north-westerly winds is largely enhanced. Further, the median AOD of the 135°-bin (monsoon) is enhanced for the fixed scaling factor. Lastly, the angular patterns of the near-surface extinctions are not significantly changed which is also consistent to the previous findings.

For the dependencies of the measured atmospheric water vapour content on the meteorological conditions, expected results are obtained. In particular, an increase of the retrieved H_2O VCDs and surface concentrations with increasing temperatures is found which is superposed by a strong maximum between 25 and 30 °C corresponding to typical temperatures during monsoon. Consistent to the wind patterns at Mohali which mainly follow the annual monsoon cycle, strongly enhanced H_2O abundances are observed under easterly to south-easterly winds, while lower water vapour contents are found in the remaining wind sectors. The results for HONO are briefly described in Sect. 12.7.

12.4 TRACE GAS CORRELATIONS

In order to gain more insights into the relations between the different species, their VCDs (or AODs) and the corresponding near-surface quantities are correlated amongst each other. For that, hourly means were calculated to reduce the scatter of the data and obvious outliers were excluded from the correlation plots. A more detailed explanation of how the data were treated is provided in Sect. 10.4.

Figure 12.14 shows the correlation of glyoxal and HCHO VCDs and near-surface concentrations in the upper and lower parts, respectively. A simple linear regression whose results are shown in red was performed as an orientation.

For both quantities, reasonable correlations between glyoxal and HCHO can be observed with R^2-values of 0.40 and 0.57 for VCDs and near-surface concentrations, respectively. Also the results of the linear regressions are similar for both quantities yielding slopes of 2.0 % and 1.9 %. The rather good correlation reflects the very similar annual variations of both species, as described in Sect. 12.2, indicating similar or even common sources. As already observed for the Bayfordbury measurements, better correlation is found for the near-surface concentrations which also show slightly more similar annual cycles compared to the VCDs. This finding can be explained by on average different profile shapes with the majority of glyoxal closer to the ground, while HCHO profiles typically reach higher (compare Sect. 12.5). While the surface concentrations correlate reasonably well

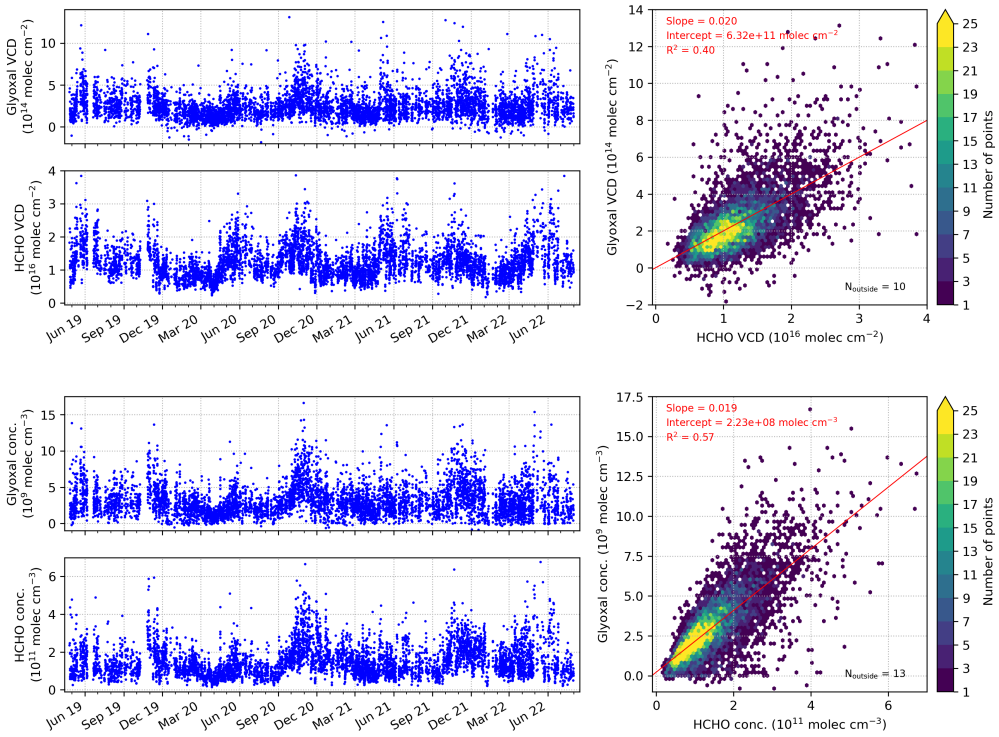


Figure 12.14: Comparison of glyoxal and formaldehyde VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The results of a simple linear regression are shown in red, while the number of excluded data points is indicated in the lower right corners.

in all seasons[4], in the monsoon season clearly worse correlation between both trace gases is found for the VCDs. The latter is caused by the fact that the HCHO VCDs are clearly reduced during monsoon, whereas the glyoxal VCDs show only slightly lower values as compared to the pre-monsoon values (compare time series). Particularly good correlation of the surface concentrations of both species can be observed in summer/pre-monsoon and post-monsoon with R^2-values of 0.64 and 0.62, respectively, strongly suggesting common sources most likely being crop residue fires. Nevertheless, also biogenic sources might contribute in these warm seasons.

The findings so far suggested a close relation between anthropogenic activities and HCHO abundances at Mohali. To further investigate this, Fig. 12.15 summarises the correlations of HCHO and NO_2 VCDs (upper part) and near-surface concentrations (lower part). Since no simple linear correlation between both trace gases is found, no linear regression was performed here.

Compared to the rather clear correlations between glyoxal and HCHO, more complex (diverging) patterns are found for the correlations between HCHO and NO_2. Although these diverging patterns are comparable to the ones found for the

4 No seasonal correlation plots are shown in the main part. For the resulting R^2-values of the correlations on seasonal basis, refer to Table A.2 in the appendix.

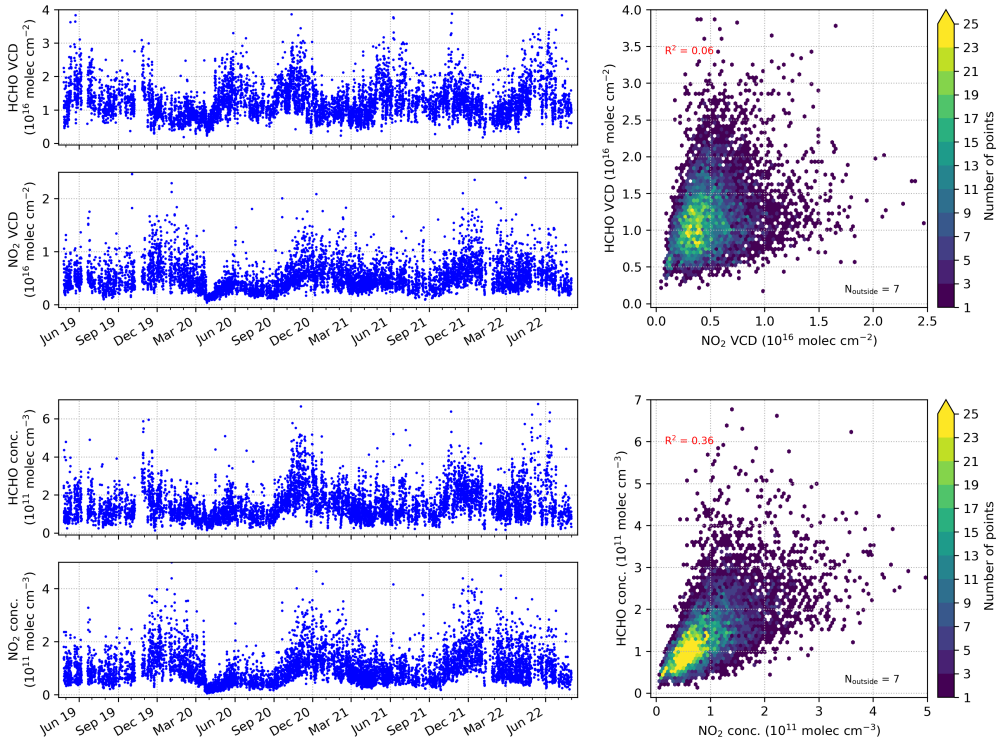


Figure 12.15: Comparison of formaldehyde and NO_2 VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Bayfordbury and Mainz measurements, they show apparently less spread. For the near-surface concentrations, this spread is particularly small and reasonable correlation between both species is found yielding an R^2 of 0.36. Consequently, a clear separation into branches, like done for Bayfordbury and Mainz (compare Sect. 10.4 and 11.4), which can be attributed to individual seasons is not possible anymore. Altogether, these findings suggest again that at Mohali the HCHO abundances are tightly connected with anthropogenic activities, in particular biomass burning events. This is supported by the reasonable correlation of the near-surface concentrations of both trace gases which is particularly high in the seasons in which biomass burning fires are active. The large spread which is rather prominent for the VCDs shows that, despite having similar sources, the atmospheric processing, the lifetimes and the vertical profiles of NO_2 and HCHO differ.

The correlations between glyoxal and NO_2 VCDs (upper part) and near-surface concentrations (lower part) are presented in Fig. 12.16.

Similar to the correlation of HCHO and NO_2, the correlation between glyoxal and NO_2 shows patterns that are rather complex but more compact than observed for Bayfordbury which does not allow a simple separation into branches corresponding to different seasons. However, compared to Fig. 12.15, more variability is found for the correlation of glyoxal with NO_2. In particular, for the

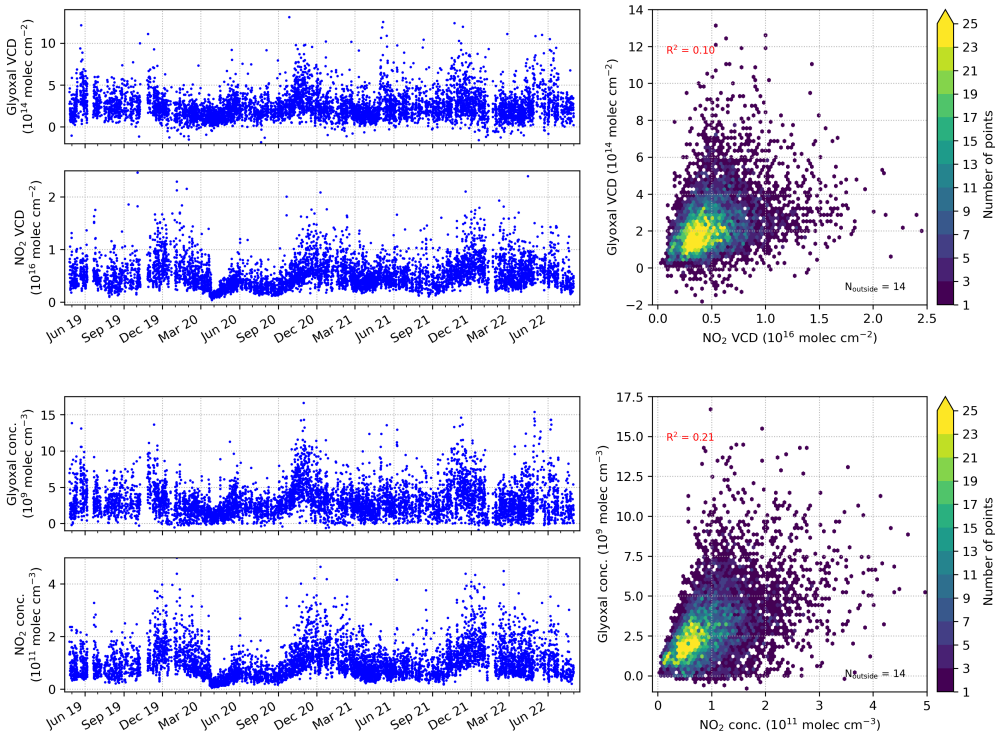


Figure 12.16: Comparison of glyoxal and NO_2 VCDs (upper part) and surface concentrations (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

near-surface concentrations a large spread is visible leading to similar patterns of both quantities. The latter is consistent to the previous findings showing that for glyoxal VCDs and surface concentrations are more closely connected to each other than for HCHO. Nevertheless, also the generally larger variability of the glyoxal results contributes to the differences in the patterns. For instance, for the glyoxal surface concentrations many values close to 0 occur (compare lower part of Fig. 12.16). This larger variability of the glyoxal results is most likely caused by the rather challenging glyoxal retrieval, i.e. weaker absorption of glyoxal compared to HCHO and NO_2 as well as the strong effect of spectral interferences as outlined earlier (compare Sect. 7.2). If the surface concentrations close to 0 are ignored, the patterns of the correlations of HCHO and glyoxal with NO_2 surface concentrations are comparable. In this case, reasonable correlation of the near-surface concentrations of glyoxal and NO_2 is found which is particularly high in the biomass burning seasons, e.g. an R^2 of 0.33 is obtained for the pre-monsoon season. These findings stress the close relation between HCHO and glyoxal also under environmental conditions that are strongly influenced by biomass burning.

Next, correlations between trace gas and aerosol results are investigated. Figure 12.17 shows the correlation between HCHO and aerosol column (upper part) and near-surface (lower part) quantities.

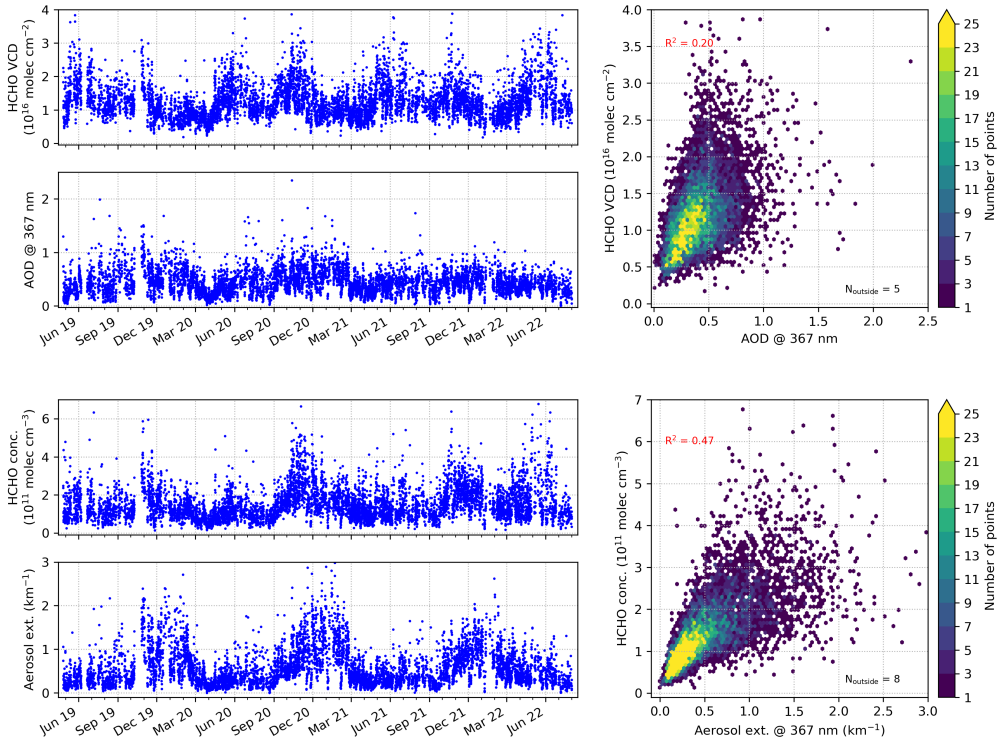


Figure 12.17: Comparison of formaldehyde and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

For both quantities, increasing aerosol loads are generally associated with increases of the HCHO abundances, whereby notably better correlation between both species is observed for the near-surface quantities ($R^2 = 0.47$). A separation of the data into the individual seasons reflects the annual cycles of both species. Despite the rather weak annual cycle of the AODs, reasonable correlation is found for the column quantities of both species during summer/pre-monsoon and post-monsoon ($R^2 = 0.37$) as both HCHO VCDs and AODs show similar annual patterns in these two seasons. For the near-surface quantities, in these seasons particularly good correlations are found with R^2-values of 0.61 and 0.62 for pre- and post-monsoon, respectively. During winter, the correlation reflects the different variations of both near-surface quantities throughout the season (compare lower left panels in Fig. 12.1 and 12.7) leading to a diverging pattern. The good correlation found for the near-surface quantities of both species, especially in the seasons in which biomass burning fires are active, confirms the previous finding that at Mohali the HCHO abundances are strongly affected by anthropogenic activities. Consistent to the observation from above that a slightly different annual cycle is obtained for the AODs using the fixed SF, a slightly different pattern of the correlation between HCHO VCDs and AODs is found in that case. In particular, a branch is visible with moderate HCHO VCDs at rather high AODs which corresponds to the monsoon season.

The correlations between glyoxal and aerosol results are summarised in Fig. 12.18.

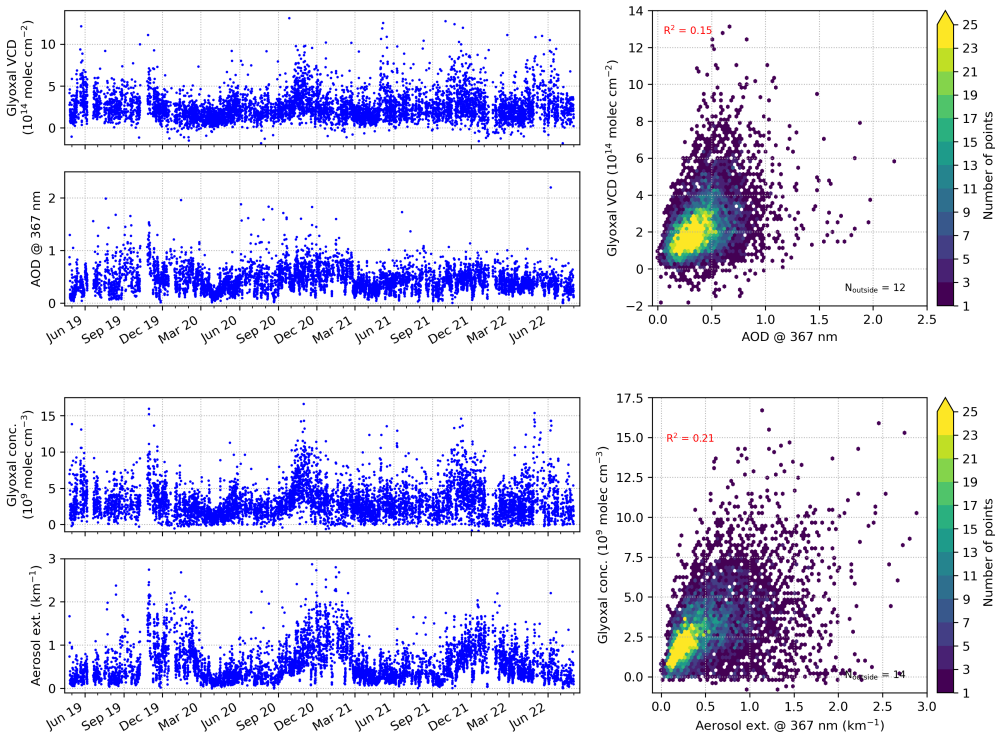


Figure 12.18: Comparison of glyoxal and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Given the similar annual cycles of glyoxal and HCHO as described in Sect. 12.2 and consistent to the previous findings which suggested similar sources of both species, also the general patterns of the correlations between glyoxal and aerosol VCDs/AODs and near-surface concentrations/extinctions are comparable to the ones obtained for the correlations of HCHO and aerosol results. The slight differences in the details of the annual variations of glyoxal and HCHO are not fully reflected by this simple correlation study. Nevertheless, apparently the patterns in the correlation plots of column and near-surface quantities are more similar for glyoxal compared to HCHO. Further, larger variability is found for the correlations of glyoxal and aerosol results, e.g. many values close to 0 exist for the surface concentrations of glyoxal (lower part in Fig. 12.18). The latter can be mainly attributed to measurements during winter. Like for the correlations between HCHO and aerosol results, the glyoxal and aerosol results correlate best in the pre- and post-monsoon seasons. Particularly good correlations are again observed for the near-surface quantities of both species yielding R^2-values of 0.42 and 0.38 for pre- and post-monsoon, respectively. In view of the larger variability of the glyoxal results, these R^2-values are lower compared to the correlations of HCHO and aerosol results.

Finally, the NO_2 VCDs (upper part) and near-surface concentrations (lower part) are compared to the corresponding quantities of the aerosol results in Fig. 12.19.

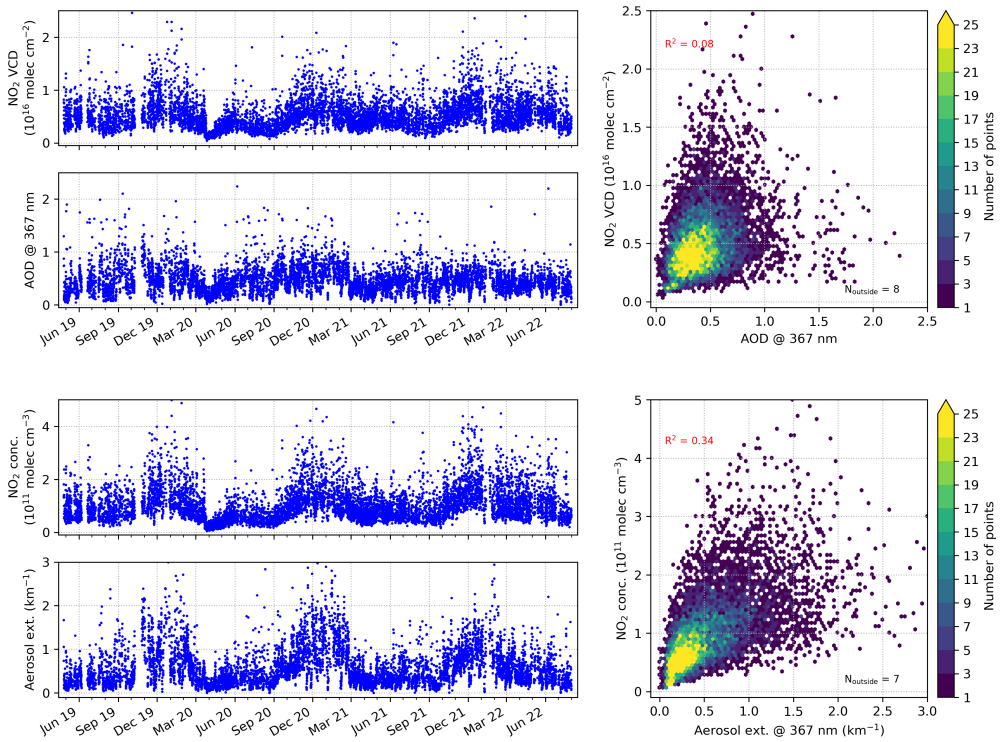


Figure 12.19: Comparison of NO_2 and aerosol results for VCDs and AODs (upper part) and near-surface concentrations and extinctions (lower part). Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

Compared to the correlations of HCHO and glyoxal with aerosol, more complex patterns are found for the correlations of NO_2 with the aerosol results. For the column quantities of both species, two branches seem to exist in the correlation plot. One is associated with a strong increase of the NO_2 VCDs with increasing AODs, while the other branch shows only a weak increase of the NO_2 VCDs when the AODs are increased. This observation fits the findings of the investigations of the dependencies on the wind direction from above (compare Fig. 12.13) which showed that elevated NO_2 VCDs are not always associated with enhanced AODs depending on the source region of the probed air masses (rural vs. city areas). Together, these branches lead to a rather bad correlation ($R^2 = 0.08$) for the column quantities of NO_2 and aerosol. Better correlation is found for the near-surface quantities of both species for which an R^2 of 0.34 is obtained. Also this is consistent to the angular wind patterns which showed less reduced near-surface extinctions compared to the AODs under north-easterly winds (city influence). Nevertheless, apparently also here two branches exist probably corresponding to different source regions of the probed air masses. The small differences between the patterns of the correlations of column and near-surface quantities show that generally both species have similar sources but are differently processed in the

atmosphere. Further, the sources of NO_2 are apparently less specific than for the other species. For the column quantities, the separation of the data into the individual seasons mainly reflects the annual cycle of NO_2 since no strong annual cycle exists for the AODs, whereby both branches are basically visible in all seasons resulting in similarly bad correlations throughout the year. In principle, the same is valid for the near-surface quantities of NO_2 and aerosol, however, best correlation is clearly found during pre-monsoon with an R^2 of 0.37, while the values are notably smaller in the other seasons.

Despite showing overall similar angular patterns as NO_2 in the dependencies on the wind direction above, the HCHO abundances seem to be closer connected to aerosol (compare Fig. 12.17). As mentioned before, this might be explained by several factors. In particular, NO_2 is formed rapidly from the emitted NO, while the majority of the atmospheric HCHO content is most likely formed by co-emitted precursor substances which happens comparably slower. Further, NO_2 is emitted by almost all anthropogenic activities, whereas the main sources of HCHO at Mohali are apparently fires which are typically associated with high aerosol loads. In addition, both species underlie different atmospheric processing (chemically and dynamically) also leading to different atmospheric lifetimes. Another effect might be the formation of SOA from HCHO leading to an intrinsically closer relation between HCHO and aerosol. The latter, however, is a complex process which cannot be easily resolved by simple correlation studies as presented in this work.

Lastly, annual cycles of the R_{GF} (ratio glyoxal to formaldehyde as described in Sect. 10.4) are investigated to further examine the relation between the glyoxal and HCHO abundances. Similar to Bayfordbury, only a small fraction, namely ca. 12 %, of the results yields overlapping valid HCHO and glyoxal results, whereby the fraction is approximately three times larger than at Bayfordbury consistent to the generally larger fraction of valid MAPA retrievals at Mohali (compare Table 9.8). Average annual cycles in terms of monthly box-whisker-plots of the R_{GF} obtained using VCDs (left) and near-surface concentrations (right) are presented in Fig. 12.20.

Given the large scatter of the results and consistent to the finding that the annual cycles of the surface concentrations of HCHO and glyoxal are very similar, no clear annual pattern is found for the R_{GF} obtained from the near-surface concentrations. The corresponding monthly means vary between roughly 2 and 2.7 %, whereby the minimum is found in early summer (April) and the maximum during monsoon (August). Contrary to that, a distinct annual cycle can be observed for the R_{GF} calculated from the VCDs. Highest monthly means are observed during winter (monthly means of about 2.65 % in January and December), while lowest monthly means are obtained in early summer where a value of approximately 1.65 % is found in April. This annual cycle indicates that during winter the glyoxal VCDs are (relatively seen) less reduced than the HCHO VCDs. Further, especially when medians[5] are considered (red lines in the box-whisker-plots), the R_{GF}-values obtained with both quantities are similar in winter, while the surface concentrations yield higher R_{GF}-values in the pre- and post-monsoon

5 This effect is weaker for the means which are strongly affected by outliers.

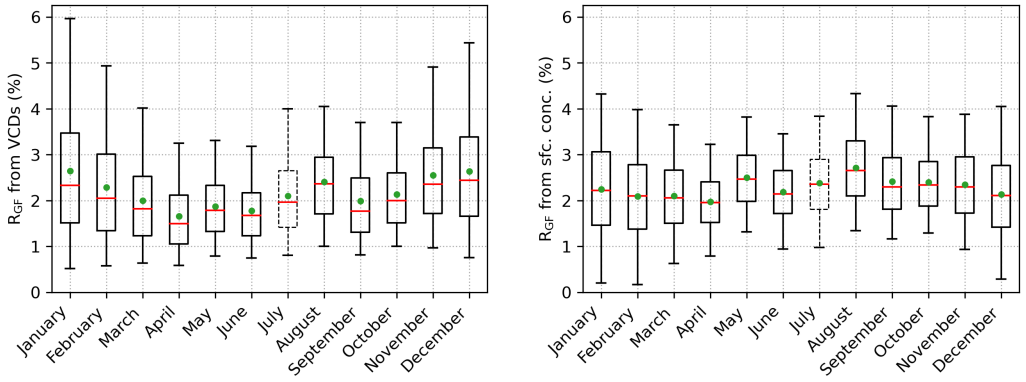


Figure 12.20: Monthly box-whisker-plots of the annual cycles of the ratio R_{GF} obtained from VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Dashed boxes indicate months with less than 5 % remaining data.

seasons. These findings suggest that the bulk of glyoxal is typically located closer to the ground than the bulk of HCHO which is particularly important in the warm periods of the year. Here, the boundary layer reaches higher and a smaller (relative) fraction of the total HCHO abundance is concentrated in the lowest atmospheric layers since HCHO is vertically well-mixed into the boundary layer. As the glyoxal profiles are shallower, a large (relative) fraction is still rather close to the ground. Consequently, the R_{GF} calculated from the near-surface concentrations is larger than the one obtained from the VCDs. The enhanced R_{GF}-values for both quantities during monsoon reflect the finding that apparently the glyoxal abundances are less reduced in that season as already described in Sect. 12.2. Overall, median R_{GF}-values are determined to be 1.9 and 2.2 % for VCDs and surface concentrations, respectively, which is about twice as high as observed at Bayfordbury indicating different chemical regimes at both stations which lead to different annual variations. In conclusion, the findings for the R_{GF} together with the results of the correlations between both species (see Fig. 12.14) suggest again similar sources, while the vertical distributions of both trace gases are apparently different (compare next section).

12.5 VERTICAL PROFILES

While the investigations of the results of the Mohali measurements so far were based on VCDs/AODs and near-surface concentrations/extinctions, the vertical distributions of HCHO, glyoxal and NO_2 concentrations as well as aerosol extinctions at 367 nm are now investigated. In order to do so, seasonal median vertical profiles of these four species were calculated analogously to Bayfordbury and Mainz and displayed in Fig. 12.21. Again, the profiles of all species show an artefact in the lowermost layers as described in detail in Sect. 9.2 and 10.5.

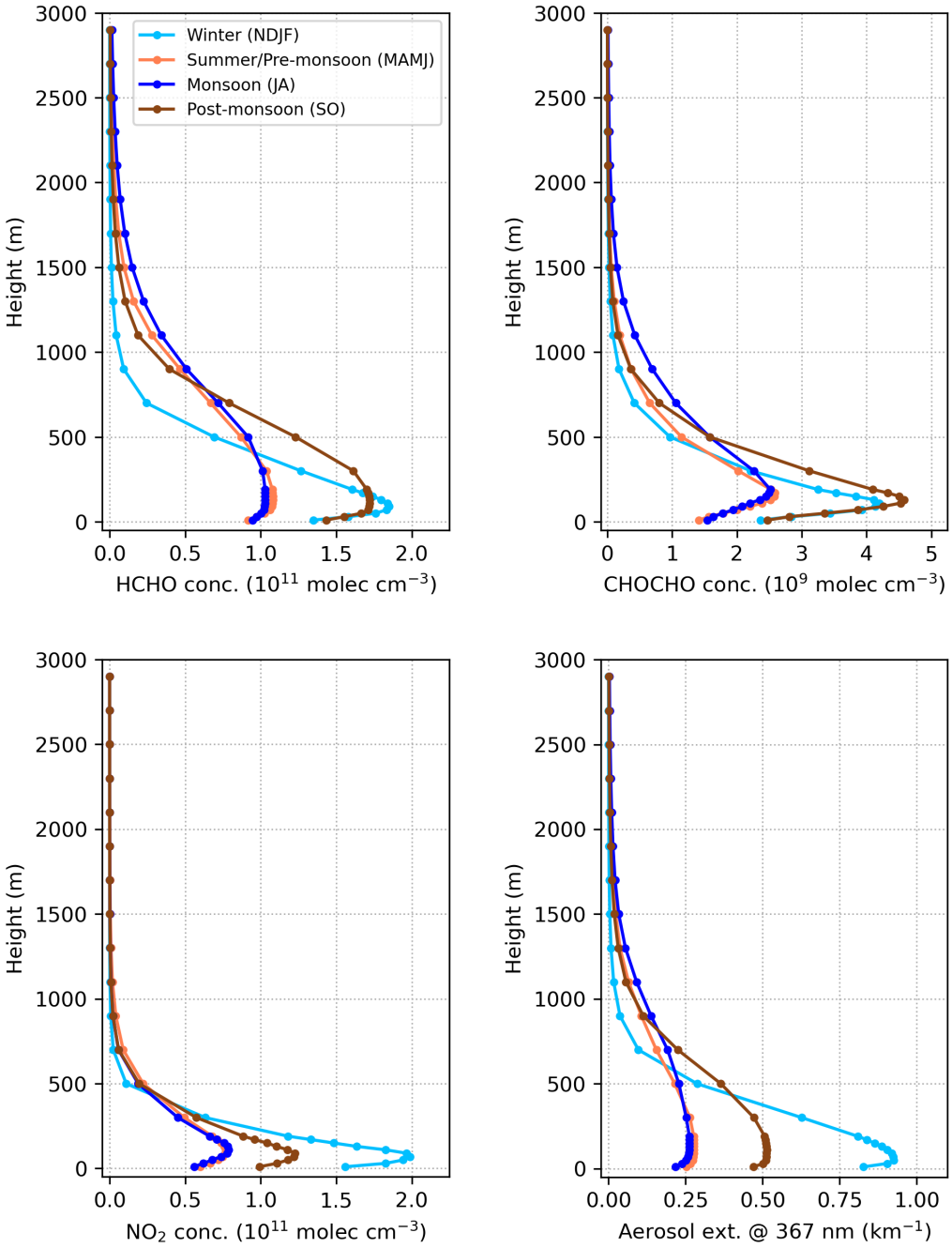


Figure 12.21: Seasonal median profiles of formaldehyde (upper left), glyoxal (upper right) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the Mohali measurements. Profiles for the different seasons are indicated by different colours.

Compared to Bayfordbury and Mainz, the seasonal variations of the relative profile shape of the HCHO concentration profiles are larger at Mohali. While the maximum concentrations are found up to a height of approximately 300 m in most parts of the year, the maximum concentrations reach only to a height of ca.

130 m in winter. Above these heights, the concentrations smoothly decrease with increasing altitude, whereby the majority of the HCHO is distributed up to about 1.25 km in summer, during monsoon and in the post-monsoon season and below only 1 km in winter. Besides variations in the profile height, also the absolute concentrations and the layer thickness in which the bulk of HCHO is concentrated vary strongly between the seasons. In particular, a very confined layer contains the bulk of HCHO in winter, while large fractions of the HCHO content are distributed in broader layers in summer and during monsoon. Consistent to the observed annual cycles of VCDs and near-surface concentrations (presented in Fig. 12.1), the concentrations in the lowest layers are highest in post-monsoon and winter, while comparably low near-surface concentrations are found in summer/pre-monsoon[6] and during monsoon. Since the profiles reach highest in the latter two seasons, maximum VCDs are found in these warmer parts of the year, whereas the very shallow profiles in winter lead to lower VCDs despite high near-surface concentrations. As pointed out earlier, the observed seasonal variations of the profile shape can be explained by a combination of the annual variations of sources, boundary layer dynamics and atmospheric chemistry. While at Mohali relative profile shapes comparable to Bayfordbury and Mainz are found in large parts of the year, the profiles reach generally less high. Further, especially in winter, the HCHO concentration profiles are significantly shallower in Mohali compared to Bayfordbury and Mainz and the HCHO content is concentrated very close to the surface in a rather confined layer. These generally shallower profiles at Mohali can probably be explained by different sources located not too far from the instrument. Lastly, the observed near-surface concentrations of HCHO are roughly two to three times higher at Mohali compared to the other two stations.

Also the glyoxal concentration profiles show their highest concentrations up to an altitude of ca. 300 m in summer/pre-monsoon and during monsoon, whereas the maximum concentrations occur below roughly 130 m in winter. An altitude of approximately 170 m is found post-monsoon. Despite these similar heights below which the maximum concentrations occur, the glyoxal profiles are shallower than the concentration profiles of HCHO throughout the year except for winter. This is reflected by the finding that the majority of glyoxal is distributed below a height of about 1 km in all seasons caused by a steeper decrease of the concentrations with increasing height after reaching their maxima. Interestingly, the relative shapes of the HCHO and glyoxal profiles are comparable in winter. Further, the thickness of the layer in which the bulk of glyoxal is concentrated is more confined compared to HCHO. The annual variations of the profile shape are similar for both species, e.g. thicker layers and slightly higher reaching profiles are found in the warmer seasons. The latter effect, however, is only weakly pronounced for glyoxal. The observations made for the profiles of HCHO and glyoxal are consistent to the previous findings which suggested similar behaviours for both species with glyoxal typically being concentrated closer to the ground. For instance, slightly better correlation was found for the near-surface quantities of both species which also yielded slightly higher ratios between both species compared to the VCDs. In

6 The used definition of the seasons allocates March and April to summer, where lowest VCDs and surface concentrations are found which affects the median profile of this season.

general, at Bayfordbury the differences between the HCHO and glyoxal profiles are similar, however, there the differences are apparently stronger pronounced, e.g. the heights of the maximum concentrations differ between both species. This might be related to the fact that at Bayfordbury the sources are located farther away from the instrument putting a stronger weight on possibly different atmospheric processing of both trace gases.

Similar to Bayfordbury and Mainz, NO_2 exhibits the shallowest vertical profiles of the four main species, whereby the profiles are particularly shallow at Mohali since basically the whole column is concentrated in the lowest 500 m of the atmosphere. These very shallow profiles are observed in all seasons which is consistent with the finding that VCDs and near-surface concentrations of NO_2 show very similar annual cycles as presented in Fig. 12.5. The slight seasonal variations of the NO_2 profiles mainly reflect the annual variability of the atmospheric NO_2 content, while the relative profiles shapes are similar in all seasons. Only the thickness of the layer containing the bulk of NO_2 is slightly increased during summer/pre-monsoon and monsoon which is also indicated by the observation that highest NO_2 concentrations are found below approximately 110 m in the post-monsoon season and winter, while the maximum concentrations occur below ca. 170 m during summer/pre-monsoon and monsoon. Also at Bayfordbury and Mainz, only small differences in the relative profile shape of NO_2 were observed with marginal (Mainz) to no (Bayfordbury) increases of the thickness of the layer in which the bulk of NO_2 is concentrated in the warmer seasons of the year. Like for these two stations, this rather unexpected finding (no significant increase of the profile height during summer despite a higher reaching boundary layer) can probably be explained by a considerably shorter atmospheric lifetime in the warmer parts of the year which compensates the effect of stronger vertical mixing. The particularly shallow profiles at Mohali can probably be explained by a remarkably short lifetime of NO_2 in combination with sources that are situated rather close to the measurement site.

Similar to the HCHO profiles, comparably strong seasonal variations of the shape of the aerosol extinction profiles can be observed. Highest extinctions occur below about 300 m in summer/pre-monsoon, during monsoon and in the post-monsoon season, while maximum extinctions are found below ca. 130 m during winter. Above, the extinctions smoothly decrease with increasing altitude. Following the general tendency that the profiles are shallower at Mohali compared to Bayfordbury and Mainz, the majority of aerosol is distributed up to an altitude of roughly 1.25 km in summer/pre-monsoon, during monsoon and in the post-monsoon season and notably below 1 km in winter. Also the extinctions in the lowest layers vary strongly between the four seasons reflecting the annual cycle of the near-surface aerosol extinctions as discussed in Sect. 12.2. Similar to Mainz, the strong variations of the near-surface extinctions are counteracted by changes in the profile height leading to only weak annual variations of the AODs. Lastly, also the thickness of the layer in which the bulk of aerosol is concentrated varies strongly between the seasons, whereby rather confined layers are found in winter while large fractions of the aerosol content are distributed in broader layers in summer and during monsoon. Like for HCHO, the observed seasonal variations

of the profile shape can be explained by annual variations of pollution sources in combination with boundary layer dynamics.

Similar to the two previous stations, the usage of a fixed O_4 scaling factor of 0.8 for the MAPA inversions does almost not change the general findings for the trace gases, whereas the absolute values of VCDs and near-surface concentrations are slightly changed. While this effect is only weak for the VCDs, the near-surface concentrations are stronger affected, whereby larger values are obtained for both quantities using the fixed SF. As explained in detail for Bayfordbury (compare Sect. 10.5), this is caused by a slightly different vertical distribution within the total column. For Mohali, however, the effect is apparently differently strong for the different seasons with the largest differences (for the trace gases) between both SF scenarios occurring pre- and post-monsoon. No clear explanation for the latter can be given so far, however, an effect caused by an unperfect description of the aerosol properties in the radiative transfer simulations is very likely since the aerosol loads are markedly higher at Mohali compared to Bayfordbury and Mainz. For the aerosol retrieval, a different annual cycle is observed for the AODs obtained with the fixed SF, while the annual cycle of the near-surface extinctions is almost unchanged (compare Sect. 12.2). Similar to Bayfordbury and Mainz, both quantities are comparably enhanced for the aerosol retrieval. For details, refer to Sect. 10.5. Further, different diurnal cycles are obtained with the different SF scenarios as outlined in detail in Sect. 12.2.

For water vapour, high reaching profiles are obtained. However, compared to Bayfordbury and Mainz also for water vapour the profiles are shallower at Mohali during winter, summer and post-monsoon with the majority of H_2O being distributed below ca. 1.5 to 2 km depending on the season. During monsoon, the profiles even reach above 3 km. Also the strong annual variation of the atmospheric water vapour content is reflected by the profiles.

Overall, the findings for the observed seasonal variations of the profile shapes and heights are comparable to the study of Kumar et al. (2020). For instance, also Kumar et al. (2020) found no major differences of the relative profile shape of NO_2 between the different seasons. Further, also higher reaching HCHO and aerosol profiles in summer/pre-monsoon compared to winter were found in their study. A more detailed comparison is beyond the scope of this work and remains difficult since Kumar et al. (2020) used a coarser MAPA output grid with a resolution of 200 m over the whole vertical profile which hides some of the finer details.

12.6 COMPARISON TO TROPOMI

Following the approach outlined in detail for Bayfordbury in Sect. 10.6, Fig. 12.22 summarises the comparison of formaldehyde VCDs obtained by MAX-DOAS and the satellite instrument TROPOMI on daily (upper part) and monthly (lower part) basis at Mohali. Again only MAX-DOAS data recorded ± 30 min around the local equator crossing time of TROPOMI (13:30 LT) were considered for the comparison.

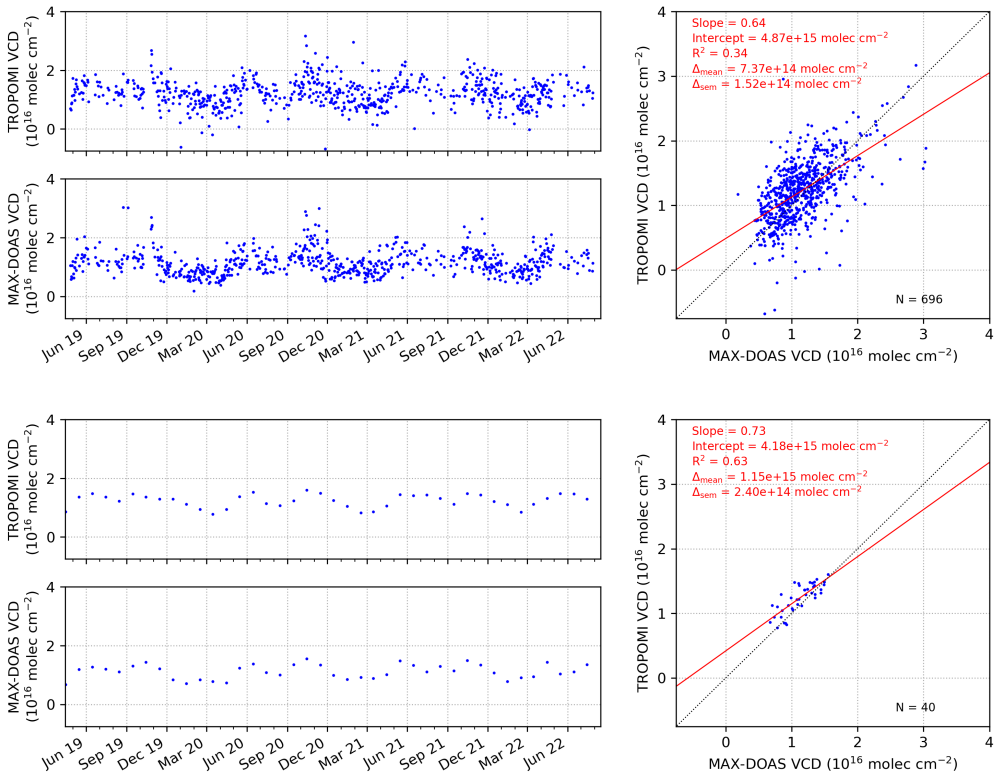


Figure 12.22: Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Mohali on daily (upper part) and monthly basis (lower part). Each part consists of time series of TROPOMI (upper left panels) and MAX-DOAS (lower left panels) VCDs as well as the corresponding correlation plots (right panels). The results of a simple linear regression are shown in red and the number of data points used in the comparisons is indicated in the lower right corners.

Generally, both data sets agree reasonably well, whereby an R^2 of 0.34 is found. The bulk of the daily mean VCDs lies rather close to the 1-to-1 line of the correlation plot, where no over- or underestimation of particularly low or high VCDs is found which was the case for the highest HCHO abundances at Bayfordbury and Mainz. A large fraction of the data points indicates higher VCDs for the TROPOMI measurements which is surprising since typically satellite retrievals underestimate the abundances of trace gases located close to the ground which was also observed at Bayfordbury and Mainz. This tendency is also reflected by the mean difference between both data sets which is $(7.37 \pm 1.52) \times 10^{14}$ molec cm^{-2} and corresponds to approximately 7 % of the mean MAX-DOAS HCHO VCD around the satellite overpass (1.09×10^{16} molec cm^{-2}). So far, no final explanation for the observed overestimation by TROPOMI was found. However, a possible explanation could be that the high aerosol loads at Mohali are interpreted as clouds by the TROPOMI algorithm. Hence, the supposed cloud effect (shielding effect) is corrected with a comparably small clear-sky AMF (compare Sect. 5.3) resulting in a too large TROPOMI VCD. However, aerosol actually tends to increase the sensitivity of satellite measurements, i.e. the AMF should be larger and thus the calculated VCD smaller. Another factor possibly contributing to the

overestimation of the HCHO VCDs by TROPOMI could be an incorrect HCHO profile assumed in the TROPOMI algorithm as the HCHO profiles at Mohali were found to be rather shallow (compare Sect. 12.5). Also a specific combination of the known effects that can cause deviations between MAX-DOAS and satellite data might contribute to this observation. These typical reasons for deviations are different sensitivities of both measurement principles to near-surface absorbers, the size of the satellite ground pixels and the fact that satellite data were averaged over 20 km around the measurement site. For further details and explanations, refer to Sect. 10.6.

The overestimation of the HCHO VCDs by TROPOMI becomes even more evident when monthly medians are considered (lower part of Fig. 12.22). The R^2 is increased to 0.63 and also the slope of the linear regression is improved. However, a larger mean difference between both data sets is calculated, namely $(1.15 \pm 0.24) \times 10^{15}$ molec cm^{-2}, which corresponds to ca. 11 % of the mean MAX-DOAS HCHO VCD. This unexpectedly increased mean difference can be attributed to a statistical artefact caused by a few daily mean TROPOMI VCDs of supposedly bad quality that show large negative deviations from the 1-to-1 line (compare upper correlation plot). These largely negative values strongly affect the mean difference and lead to a lower than expected mean difference for the daily comparison, while this effect vanishes when monthly medians are considered. Contrary to Bayfordbury and Mainz, the annual variations of the HCHO abundances are well captured by TROPOMI. One possible explanation for this finding might be comparably high HCHO VCDs throughout the year. Moreover, the fraction of valid MAPA retrievals is notably larger at Mohali compared to Bayfordbury and Mainz (compare Table 9.8).

When R^2-values are considered, the correlation of both data sets at Mohali performs worse than for the previously discussed stations, whereas the slopes of the linear regressions are larger for Mohali which can probably be explained by the shorter integration time of the Mohali measurements. This leads, as outlined previously, to more valid MAPA retrievals but also larger scatter of the results is expected for shorter integration times. Further, the rather large aerosol load at Mohali might contribute. Nevertheless, the (absolute) mean difference between the VCDs of both data sets is smaller at Mohali indicating overall a rather good agreement between TROPOMI and MAX-DOAS measurements.

Like for the previous stations, Table 12.1 summarises the (daily) comparison results using, besides the standard scenario, the MAX-DOAS retrieval results for a fixed O_4 SF of 0.8 or applying different selection criteria, namely the inclusion of MAX-DOAS measurements taken ± 60 min around the local equator crossing time and the additional usage of MAPA retrievals flagged with warnings. For details, refer to Sect. 10.6.

No significant changes in the comparison results between the different scenarios are found. While R^2-values, slopes and intercepts show only small differences, larger variations are found for the mean differences. In particular, notably smaller mean differences are obtained for the results using a fixed SF and including results flagged with a warning in the comparison. However, this can be attributed to the same statistical artefact described above which is stronger pronounced in

Table 12.1: Comparison results between daily TROPOMI and MAX-DOAS data at Mohali for different selection criteria.

SCENARIO	QUANTITY				
	R^2	Slope	Intercept (molec cm^{-2})	Δ_{mean} (molec cm^{-2})	Δ_{sem} (molec cm^{-2})
Standard	0.34	0.64	4.87×10^{15}	7.37×10^{14}	1.52×10^{14}
SF 0.8	0.34	0.57	5.59×10^{15}	3.02×10^{14}	1.74×10^{14}
$\pm$60 min	0.33	0.60	5.35×10^{15}	6.49×10^{14}	1.47×10^{14}
Warnings incl.	0.33	0.57	5.76×10^{15}	4.89×10^{14}	1.47×10^{14}

these cases leading to even smaller mean differences. Thus, similar comparison results are obtained for all scenarios.

Also Kumar et al. (2020) compared MAX-DOAS HCHO VCDs to satellite data. Different OMI HCHO products were used in their study and a difference of 3.6×10^{15} molec cm^{-2} between MAX-DOAS and satellite VCDs was found indicating an overestimation by the satellite instrument consistent to this work. Kumar et al. (2020) attributed this difference to the limited sensitivity of MAX-DOAS instruments to absorbers in the free troposphere where HCHO produced by methane oxidation might be present contributing with roughly 1×10^{15} molec cm^{-2} to the total column. However, this effect should also apply to the other stations and no such overestimation of TROPOMI was found at Bayfordbury nor Mainz. Thus, a reason specific for Mohali is probably the cause for the observed overestimation (see above).

12.7 FURTHER RESULTS: HONO

Similar to Mainz, interesting results were found for HONO at Mohali which are briefly discussed in this section.

Average annual cycles of HONO VCDs (left) and near-surface concentrations (right) represented by monthly box-whisker-plots are shown in Fig. 12.23.

The HONO results show a large variability which becomes evident by many outliers and long whiskers. Nevertheless, the monthly means (green dots) indicate a distinct annual cycle of the HONO abundances which is very similar to the annual cycle of NO_2 with maximum and minimum abundances during winter and early summer, respectively. Weaker enhancements occur pre-monsoon (June) while only low HONO abundances are measured during monsoon. As seen for Mainz, the large scatter affects the monthly means rather strongly, however, also the more robust medians show a clear annual cycle. The increase of the HONO abundances starts in the late post-monsoon season (October). After that, the monthly means of both quantities peak in December, where values of 2.0×10^{14} molec cm^{-2} and 4.0×10^{9} molec cm^{-3} are obtained for VCDs and near-surface concentrations, respectively. Towards early summer the HONO abundances decline reaching their

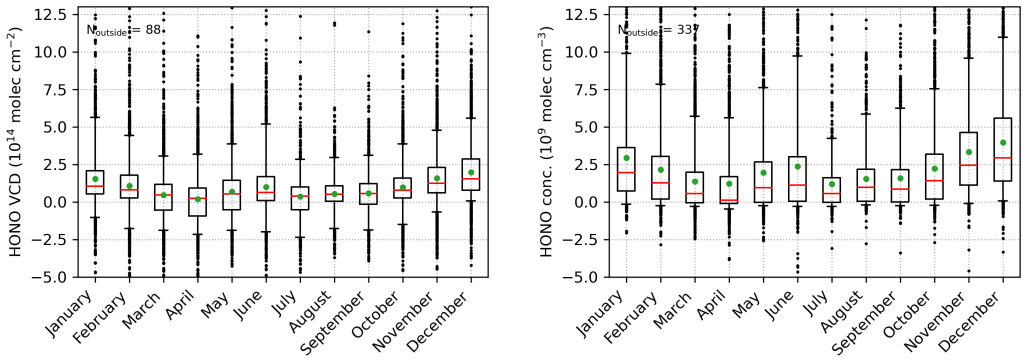


Figure 12.23: Box-whisker-plots of the annual cycles of HONO VCDs (left) and surface concentrations (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

minimum in April, where a monthly mean VCD of ca. 0.2×10^{14} molec cm^{-2} and a mean surface concentration of about 1.2×10^{9} molec cm^{-3} are found. After a minor maximum in June, during monsoon the abundances drop to values close to the minimum values found in early summer. Similar to Mainz, the annual variations of the HONO abundances can be mainly explained by the fact that HONO is closely related to NO_2 which shows basically the same annual cycle. Thus, the HONO sources at Mohali are also closely linked to anthropogenic activities. Together with the fact that HONO is quickly photolysed, this leads to minima of both quantities in early summer, where anthropogenic activities are low and solar radiation is highest, and maximum HONO abundances in winter, where the opposite is true. Despite higher NO_2 pollution levels at Mainz, the HONO abundances at Mohali are in general comparable to the ones at Mainz. However, the near-surface concentrations during winter are apparently larger at Mohali, while the VCDs are comparable for both stations. This finding indicates very shallow HONO profiles at Mohali, particularly during winter (compare Fig. 12.26).

For both quantities, similar results are obtained for the fixed O_4 SF which mainly leads to slightly larger HONO abundances, whereby the surface concentrations are stronger affected. Also the scatter is further enhanced compared to the variable SF.

The average diurnal cycles for the different seasons indicated by different colours are shown in Fig. 12.24.

Both VCDs and near-surface concentrations show clear diurnal cycles which are similar in all seasons and for both quantities. These diurnal variations are generally very similar to the ones observed at Mainz (compare Fig. 11.18) and again stronger pronounced for the near-surface concentrations suggesting very shallow HONO profiles. However, differences in the details can be observed. In particular, at Mohali both quantities reach their minimum values quasi simultaneously around noon in all seasons except for winter, where the minimum VCDs

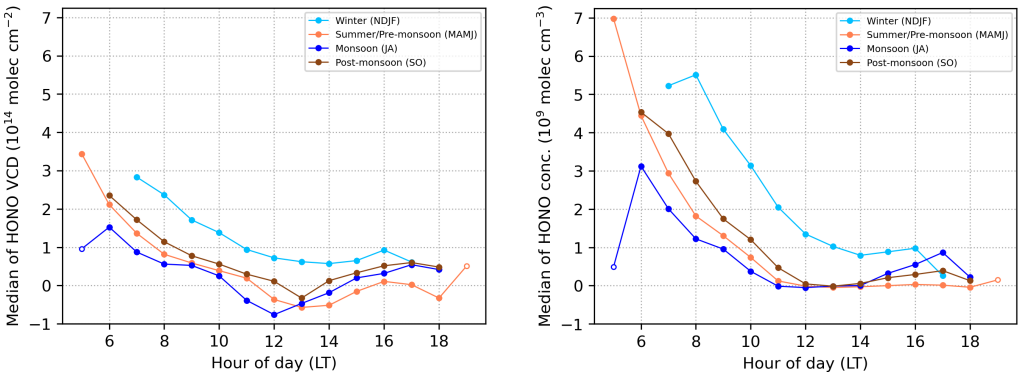


Figure 12.24: Diurnal cycles of HONO VCDs (left) and surface concentrations (right) for Mohali for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

and surface concentrations are obtained approximately two hours later (but also simultaneously). This is different to Mainz where a time lag of a few hours was observed between the minima of both quantities. This difference might be explained by shallower profiles and shorter HONO lifetimes at Mohali. In addition, less variability of the times at which the minima occur is found at Mohali, probably caused by less variability of the day lengths between the seasons as Mohali is located at a lower latitude than Mainz. Besides these differences, the diurnal patterns can again be explained by rapid photolysis of HONO under the influence of solar radiation. This leads to HONO abundances of practically 0 during the afternoon in all seasons[7] except winter, where still median HONO VCDs of ca. 0.6×10^{14} molec cm^{-2} and surface concentrations of around 1×10^{9} molec cm^{-3} occur in the afternoon. The latter leads to higher HONO abundances during winter compared to the other seasons despite similar values in the early morning. Similar to Mainz, no significant increase of the HONO abundances towards sunset is observed and the small increases that are visible especially for the VCDs can most likely be attributed to statistical artefacts. The same is valid for the particularly low surface concentration at 5 a.m. during monsoon which is also flagged for having less than 5 % valid MAPA retrievals.

The dependencies of the HONO abundances on temperature (uppermost panels) as well as wind speed and direction (middle and lowermost panels) are summarised in Fig. 12.25.

Despite the large scatter of the HONO results, a tendency of decreasing HONO abundances with increasing temperature is visible especially for the VCDs. For the near-surface concentrations, the scatter is even larger and the observed decreasing trend is less clear. Nevertheless, also the majority of the highest near-surface concentrations are found at comparably low temperatures below ca. 20 °C. These observations reflect the findings that elevated HONO abundances occur mainly in the early morning hours and that highest monthly means are found during

7 The VCDs are even slightly negative during afternoon in the warmer parts of the year, i.e. summer and monsoon.

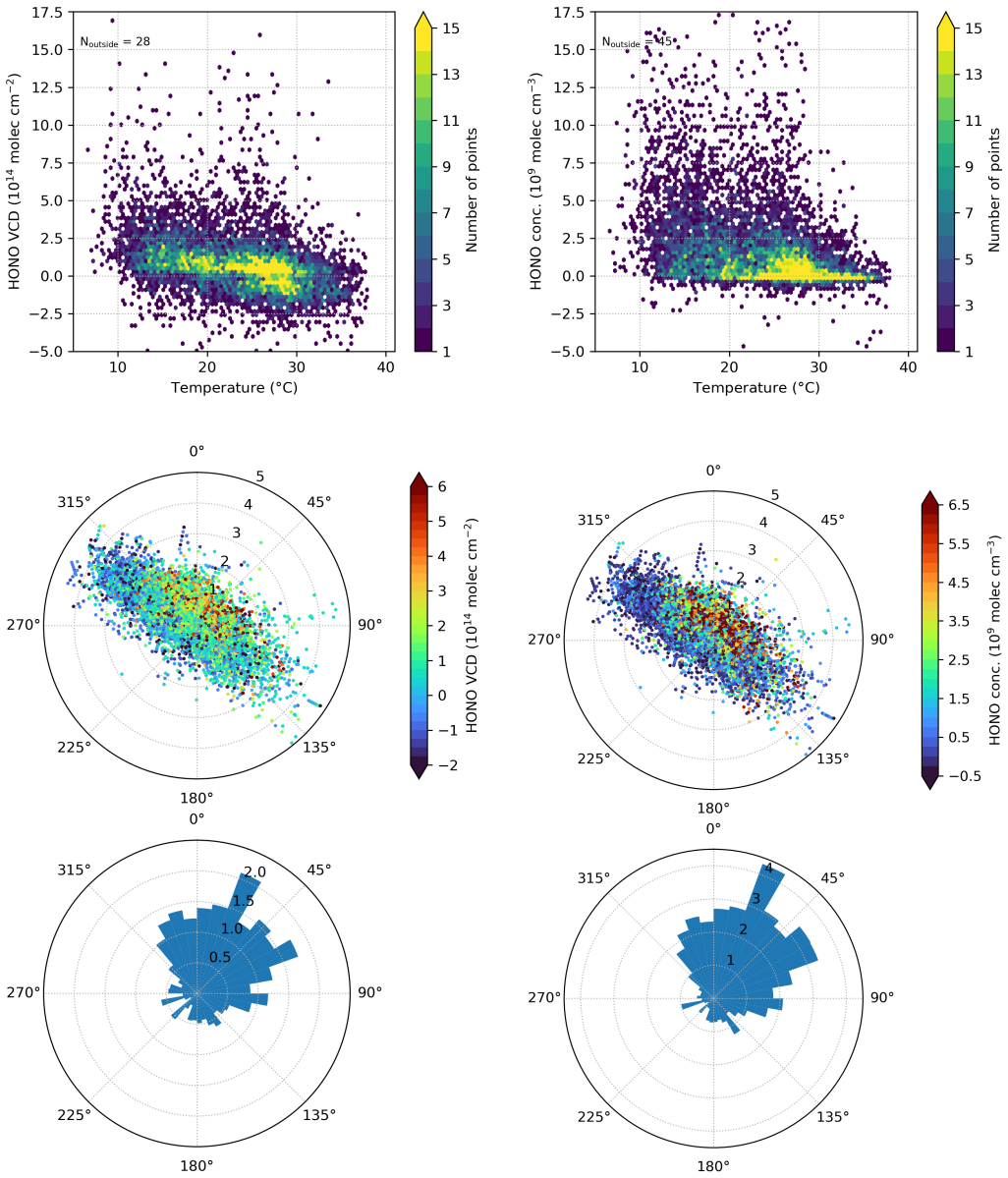


Figure 12.25: The two uppermost panels show hourly mean HONO VCDs (left) and surface concentrations (right) measured at Mohali as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HONO VCDs (left) and surface concentrations (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HONO VCDs (left) and surface concentrations (right) obtained under different wind directions in bins of 10° are indicated by the radius.

winter where temperatures are lowest. Overall, these observations are comparable to the Mainz results.

Despite the finding from above that the HONO and NO_2 abundances are obviously closely related (compare annual cycles) and the fact that NO_2 pollution levels are elevated under winds from a wide sector (compare Fig. 12.12), enhanced HONO abundances occur only under winds originating from a sector spanning

(clockwise) from north-west to east. In the remaining sectors, the HONO abundances are much smaller (mostly by at least a factor of three) and only marginal enhancements occur in single angular bins. This is valid for both quantities which show practically the same angular patterns. One possible explanation for the differences in the angular distributions of HONO and NO_2 is the distance of the respective NO_2 source from the measurement site. In particular, the emissions under north-westerly to easterly winds mostly originate from the *"Tricity"* of Chandigarh, which is very close to the measurement site, and lead to rather fresh pollution that is probed under such wind directions. Together with the comparably low wind speeds from these directions (compare middle panels of Fig. 12.25) hindering strong mixing, this can lead to strongly enhanced HONO abundances at the measurement site. On the other hand, NO_2 measured under southerly to westerly winds mostly originates from biomass burning activities and domestic fires (compare Sect. 12.3) which are located farther away from the measurements site. Given its short atmospheric lifetime under the influence of solar radiation, under such wind conditions the possibly existing HONO might already be degraded when the affected air masses reach the measurement site. Nevertheless, the marginal enhancements found in single angular bins under westerly winds could indicate a minor effect of biomass burning activities. A possible second effect leading to different angular distributions of NO_2 and HONO might be different chemical compositions of the pollution originating from city related emissions and biomass burning or domestic fires.

Another finding is that both quantities show their absolute maxima under winds originating from ca. 25°, however, no distinct emission source could be identified in view of the simple approach applied in this work. Lastly, no clear conclusion can be drawn on possible wind speed dependencies since only comparably low wind speeds prevail in the wind sector in which the enhancements are found. As outlined, the latter might be one factor why enhanced HONO abundances occur in this specific sector.

The observation that the angular patterns of both quantities are quasi identical suggests very shallow HONO profiles which is also supported by the investigations of annual and diurnal cycles. To confirm this, the vertical distributions of the HONO concentrations in the different seasons are summarised in Fig. 12.26.

Following the general tendency that the vertical profiles are comparably shallow at Mohali, also the HONO profiles are clearly shallower compared to Mainz. As already observed there, the median HONO profiles are even shallower than the NO_2 profiles (compare Fig. 12.21) and the maximum concentrations reach only to a height of ca. 70 m, while the majority of HONO is distributed below approximately 300 m. After reaching their maximum concentrations, the HONO profiles show a very steep decrease of the concentrations with increasing altitude which is also steeper than for the NO_2 profiles. No large differences of the relative profile shape between the seasons are visible. The varying concentrations in the lowest layers mainly reflect the annual variations of the mean HONO abundances, while the layer containing the bulk of HONO is only marginally thicker in the warmer seasons (summer and monsoon). Like at Mainz, the lower median concentrations during the warmer seasons can be explained by the

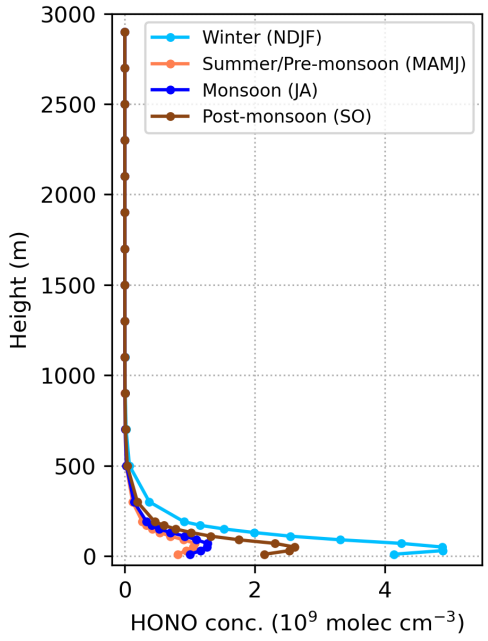


Figure 12.26: Seasonal median profiles of HONO concentrations obtained for the Mohali measurements. Profiles for the different seasons are indicated by different colours.

fact that, despite enhanced morning values (compare Fig. 12.24), the HONO abundances are close to or equal 0 during large fractions of the day in these seasons. Further, the very shallow HONO profiles and the finding that no large variations of the relative profile shape between the seasons occur can be explained by boundary layer dynamics and the short atmospheric lifetime of HONO as outlined in more detail for Mainz (refer to Sect. 11.7).

Like at Mainz, HONO can be successfully retrieved from the Mohali measurements. However, HONO is not the main focus of this work and thus no further interpretation of the results is given here. Hence, future studies could focus more on the HONO results at Mohali.

13

ATTO, BRAZIL

This chapter discusses the results of the measurements performed at ATTO, Brazil, focussing on the results for HCHO, glyoxal, NO_2 and aerosol of the two operated instruments. First, an overview of the measurements and the particularities of the analysis is provided in Sect. 13.1, followed by a discussion of annual and diurnal cycles of the four main species in Sect. 13.2. Dependencies of the measured trace gas and aerosol abundances on wind characteristics and air temperature are presented in Sect. 13.3, while the relation amongst the different species themselves is analysed in Sect. 13.4. The investigation of seasonal vertical profiles in Sect. 13.5 is followed by a brief comparison of the MAX-DOAS formaldehyde VCDs to the ones measured by TROPOMI in Sect. 13.6. Interesting findings for the annual cycle of water vapour are presented in Sect. 13.7. Since the data from both operated instruments lead overall to similar results, they are presented together in the following sections and subtle differences (if existing) are discussed. These small but systematic differences between both instruments (caused by the different altitudes) are further discussed in Sect. 13.8.

13.1 OVERVIEW AND PARTICULARITIES OF THE ANALYSIS

The first TubeMAX-DOAS instrument at ATTO, installed 80 m above ground (compare Sect. 4.4), got operational on 12 October 2017. A second MAX-DOAS instrument of the same type was installed at a height of 298 m on 29 March 2019. As of October 2023, both instruments are still running, however, in this work only data until 30 July 2022 and 31 August 2022 are considered for the 80 m and 298 m instrument, respectively.

Given the hot and humid climate at ATTO, both instruments suffered from instrumental issues, e.g. defective electronic components. These led to several rather large data gaps especially in the respective initial phases of both instruments, which are also prominently visible in the annual overview plots presented below (e.g. Fig. 13.1 and 13.2). For the 298 m instrument, the largest data gap spans over roughly 2 years caused by the fact that due to the COVID-19 pandemic no instrument repair on-site was possible. Despite these large gaps, approximately 4 and 1.5 years are covered by the 80 m and 298 m instrument, respectively.

The instrumental problems with the 80 m instrument could be solved by replacing some electronic components of the instrument, while the spectrometer itself and the optical coupling (on the spectrometer site) were unchanged. Further, the spectrometer temperature was constant over the whole measurement period with only small exceptions. Hence, one set of DC and offset spectra was used for the correction of the measured spectra and the same spectral calibration and SLF parameters could be used for all DOAS retrievals for the 80 m instrument. Contrary to that, the failures of the 298 m instrument required the exchange of

large parts of the instrument and therefore the whole instrument was replaced twice after the initial installation in March 2019. Further, a different spectrometer temperature is applied since April 2022 as the instrumental issues could be attributed to condensation of water inside the instrument housing. Obviously these factors lead to three calibration regimes and three different sets of DC and offset spectra. An overview of the wavelength ranges and spectral resolutions of the three instruments installed at 298 m was already provided in Table 4.1.

For all species, the standard fit settings summarised in Tables 9.1 to 9.5 could be applied for both instruments. However, since no 90° measurements are possible due to limitations of the instrumental setup on the tower, spectra recorded at 88° and 87° were used as FRS for the 80 m and 298 m instrument, respectively. Further, minor spectral gaps to deal with hot pixels had to be introduced in several fits for both instruments. A brief overview of these gaps is provided in Table 13.1. Also the MAPA inversions were performed using the standard settings as outlined in Sect. 9.2.1, whereby the measurements at 45° (and above) were excluded from the analysis for consistency reasons since these are only included in the elevation sequences of the two ATTO instruments. The applied elevation biases, determined as described in Sect. 9.2.2, were 1° and 0.5° for the 80 m and 298 m instrument, respectively. For both instruments, in principle data from the whole measurement period can be used for the atmospheric investigations presented in the following after applying the selection criteria as described in Sect. 9.3. However, during these studies it became evident that the glyoxal results for the last measurement period of the 298 m instrument (04/2022 – 09/2022) are of rather poor quality. Since no clear reason for this observation was found, only glyoxal results of the first two measurement periods (until 21 April 2020) are included in the studies below. For all other species and the 80 m instrument, the full data sets were used.

Table 13.1: Overview of the spectral gaps introduced to the DOAS fits of both ATTO instruments and affected species.

INSTRUMENT		GAP RANGE	AFFECTED SPECIES
80 m	10/2017 – 07/2022	415.1 – 415.4 nm	glyoxal, NO_2
		442.0 – 442.4 nm	glyoxal, NO_2
		454.9 – 455.1 nm	glyoxal, NO_2
298 m	03/2019 – 09/2019	341.4 – 341.8 nm	HCHO, HONO, O_4
		359.8 – 360.2 nm	HONO, O_4
		437.2 – 437.5 nm	glyoxal, NO_2
		450.8 – 451.2 nm	glyoxal, NO_2
	09/2019 – 04/2020	None	–
	04/2022 – 09/2022	365.4 – 365.8 nm	HONO, O_4

Given the low NO_2 abundances at ATTO which are close to the detection limit of the utilised MAX-DOAS instruments and due to the previously described strong water vapour interferences together with rapid changes in cloud coverage

(compare Sect. 4.4), only small fractions of valid NO_2 retrievals (less than ca. 5 %) are obtained for both ATTO instruments (compare Table 9.8). Although no major effects on the results of this work are expected, these low NO_2 values should not be filtered out in future, even if the relative errors of such measurements are expected to be large. Hence, an update of MAPA and the way the flags are determined is suggested here.

A last aspect that should be considered before the presentation of the results of the ATTO measurements is that MAPA yields vertical profiles relative to the altitude of the instrument. Since the two ATTO instruments are installed at altitudes of 80 m and 298 m above ground, the concentrations averaged over the lowest 400 m of the concentration profiles (compare Sect. 9.2.3) do not correspond to near-surface concentrations here. Actually, the same applies to the aerosol extinctions. Therefore, in this chapter these concentrations and extinctions are referred to as "concentrations at instrument altitude" and "(aerosol) extinctions at instrument altitude", respectively. Further, the lowermost part of the tropospheric column of a given trace gas might not be detectable by the ATTO instruments. As shown below, the latter is of major importance for the 298 m instrument (e.g. Sect. 13.5 and 13.8).

13.2 ANNUAL AND DIURNAL CYCLES

In this section, annual and diurnal cycles of HCHO, glyoxal, NO_2 and aerosol are discussed. For that, the MAPA results were filtered according to Sect. 9.3 and visualised as described in Sect. 10.2. Since both deployed instruments (compare Sect. 4.4 and 13.1) yield overall similar results, the results of both instruments are presented together in the following. An investigation of (small) systematic differences between the results obtained with both instruments (related to the difference in altitude) is performed in Sect. 13.8.

Figures 13.1 and 13.2 show the annual cycles of HCHO VCDs and concentrations at instrument altitude obtained with the 80 m and 298 m instrument, respectively. In both figures, the two uppers panels show full time series, while the two lower panels depict monthly box-whisker-plots. The same sampling criteria as for Bayfordbury were applied (compare Sect. 10.2).

For both instruments, VCDs and concentrations at instrument altitude show pronounced annual cycles with highest HCHO abundances during the dry season, while lowest HCHO abundances are observed in the wet season. For the definition of the four seasons at ATTO, refer to Sect. 4.4. The maximum daily mean VCDs (blue dots) are in the order of 2×10^{16} molec cm^{-2} and comparable to the ones obtained at Mainz and Mohali (compare Fig. 11.1 and 12.1). For both instruments, large data gaps as described above are visible in the time series.

The statistical representations of the annual cycles of HCHO VCDs (left) and concentrations at instrument altitude (right) depicted in the lowermost panels of Fig. 13.1 and 13.2 confirm that lowest HCHO is measured in the wet season. Minimum abundances are found from February to April, where monthly means of ca. 0.65×10^{16} molec cm^{-2} (both instruments) and 0.30×10^{11} molec cm^{-3} (80 m)/0.35×10^{11} molec cm^{-3} (298 m) are obtained for VCDs

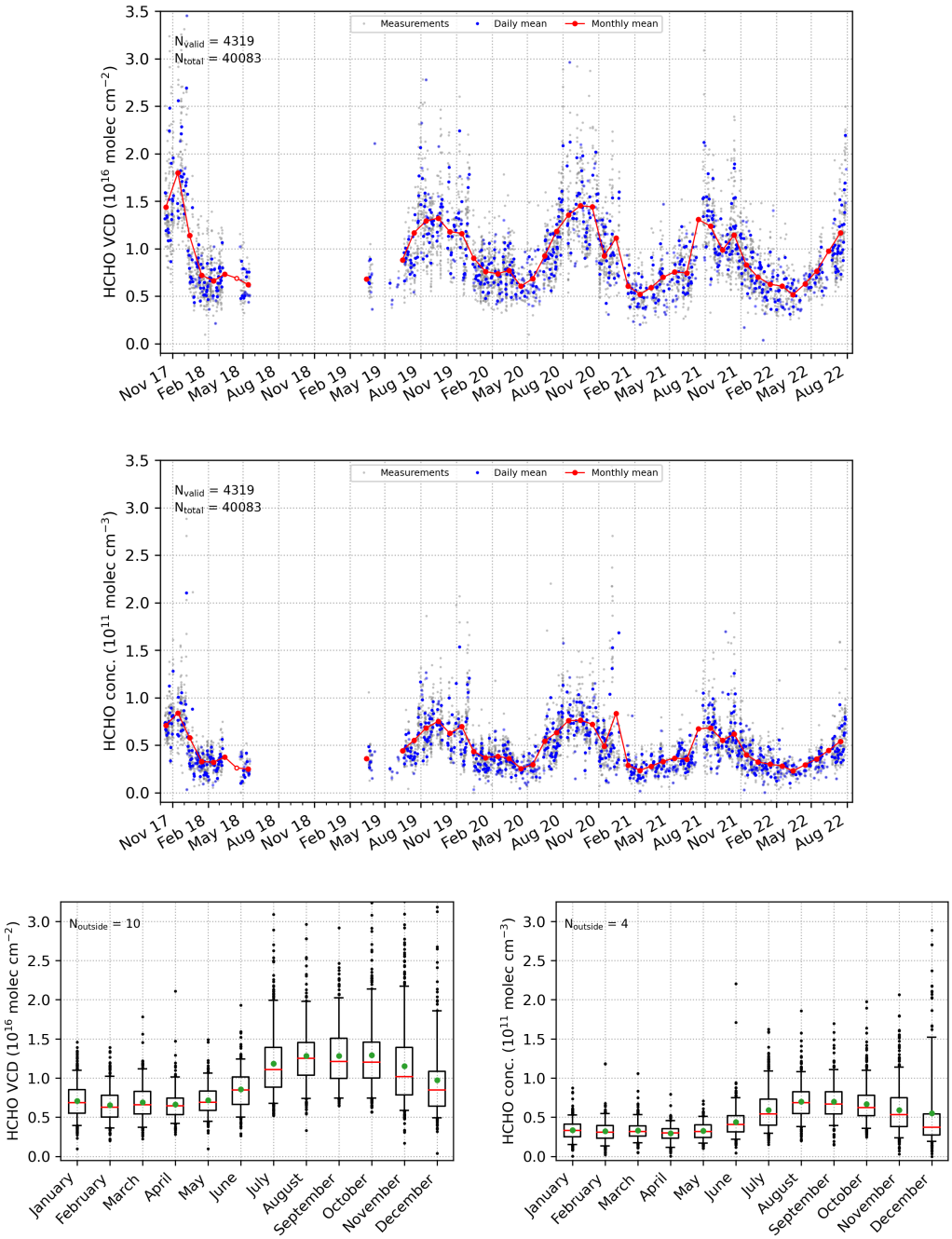


Figure 13.1: The two upper panels show time series of HCHO VCDs and concentrations at instrument altitude for ATTO (80 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of HCHO VCDs (left) and concentrations at instrument altitude (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

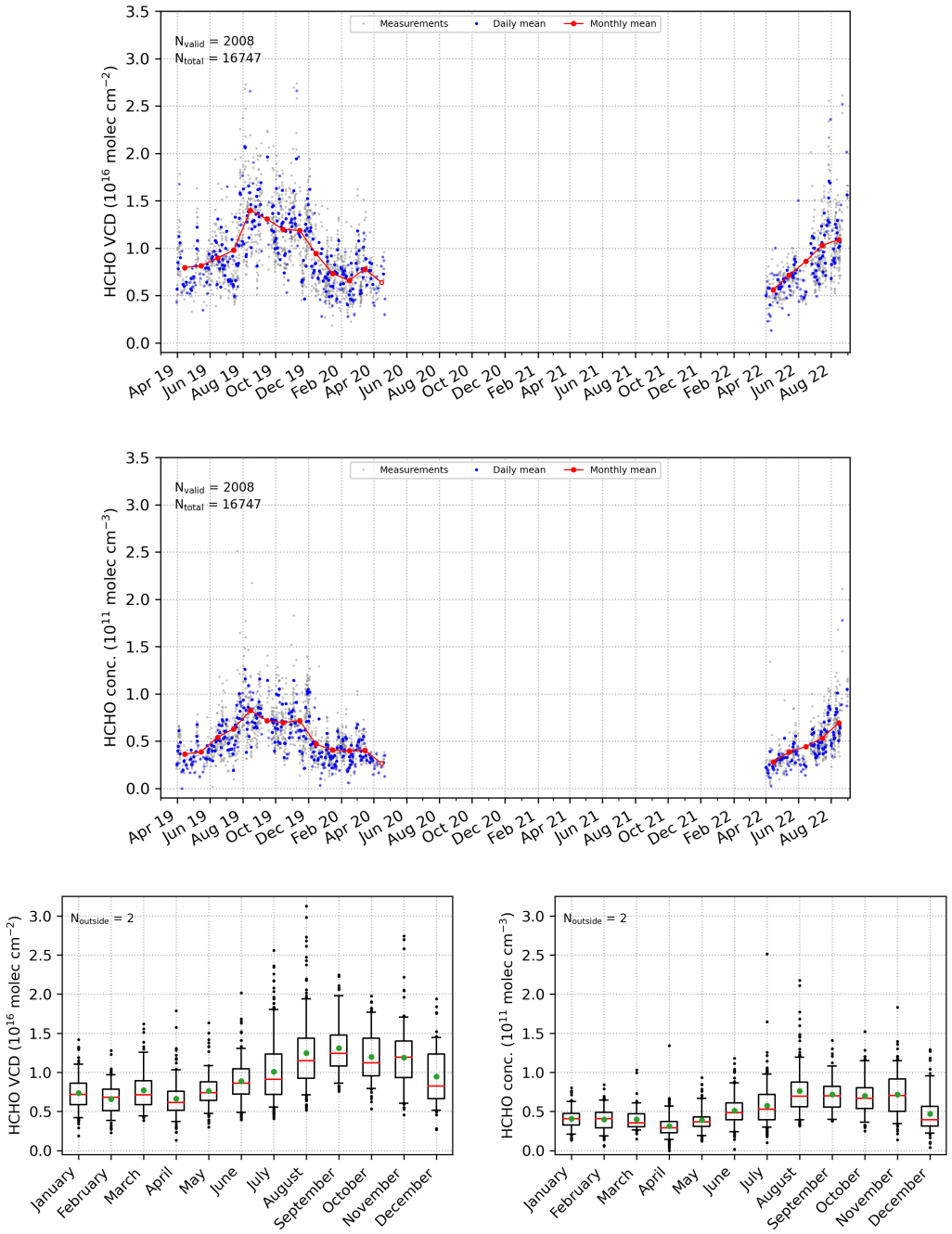


Figure 13.2: The two upper panels show time series of HCHO VCDs and concentrations at instrument altitude for ATTO (298 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of HCHO VCDs (left) and concentrations at instrument altitude (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

and concentrations at instrument altitude, respectively. Interestingly, these minimum values of both quantities are notably smaller than the corresponding minima at Mohali indicating the strong effect that anthropogenic emissions can have on the HCHO abundances. Starting at the end of the wet season (May), during the transition period towards the dry season both quantities show a continuous increase which peaks in the dry season, where from August to October monthly mean VCDs of around 1.28×10^{16} molec cm^{-2} (both instruments) and monthly mean concentrations at instrument altitude of about 0.70×10^{11} molec cm^{-3} (80 m)/0.72×10^{11} molec cm^{-3} (298 m) are found. Again these maxima are smaller than the maximum monthly means at Mohali, whereby the maximum monthly mean concentrations at instrument altitude at ATTO are even lower than the minimum monthly means at Mohali. Similar to observations made during summer at Bayfordbury and Mainz, the range of HCHO VCDs and concentrations at instrument altitude is increased in the dry season (compare whiskers in the lower panels of Fig. 13.1 and 13.2), probably because the conversion from precursors to HCHO strongly depends on the actual meteorological conditions. Further, in the dry season deforestation fires are active. Therefore, the HCHO abundances at ATTO might also vary depending on whether the site is affected by pollution originating from such fires or not resulting in an overall larger range of measured HCHO abundances in this season. At the end of the dry season and during the transition period towards the wet season, the HCHO abundances continuously decrease until January reaching values comparable to their respective minima. It should be stressed that both quantities show very similar annual variations indicating a very homogenous HCHO source close to the instrument being the rainforest itself. This also explains the observed annual cycle since the plants experience dry and heat stress during the dry season leading to enhanced emissions of biogenic VOCs acting as precursors of HCHO. Nevertheless, biogenic activity is high throughout the entire year leading to rather high HCHO VCDs, compared to Bayfordbury and Mainz, also in the wet season, while the minimum concentrations at instrument altitude are comparable to the minima at these stations. Further, it should be mentioned that more variability of the monthly box-whisker-plots is visible for the 298 m instrument caused by its rather low amount of available data resulting in worse statistics. A last finding is that the results are very similar for both instruments. However, slightly higher VCDs are obtained for the 80 m instrument, while slightly higher concentrations are found at 298 m. This indicates that the HCHO concentrations increase within the lowest 200 m above the canopy. Details are discussed in Sect. 13.8.

Similar to the previous stations, almost no differences in the annual characteristics of the HCHO results are observed for the fixed SF of 0.8. Only slightly larger values are obtained for both quantities (especially in the dry season) using the fixed scaling factor, whereby the effect is slightly stronger for the concentrations at instrument altitude again indicating small differences in the profile shapes obtained with the two different SF (compare Sect. 13.5).

Diurnal cycles of HCHO VCDs (left) and concentrations at instrument altitude (right) for both instruments are presented in Fig. 13.3, where different colours indicate results obtained for the four seasons.

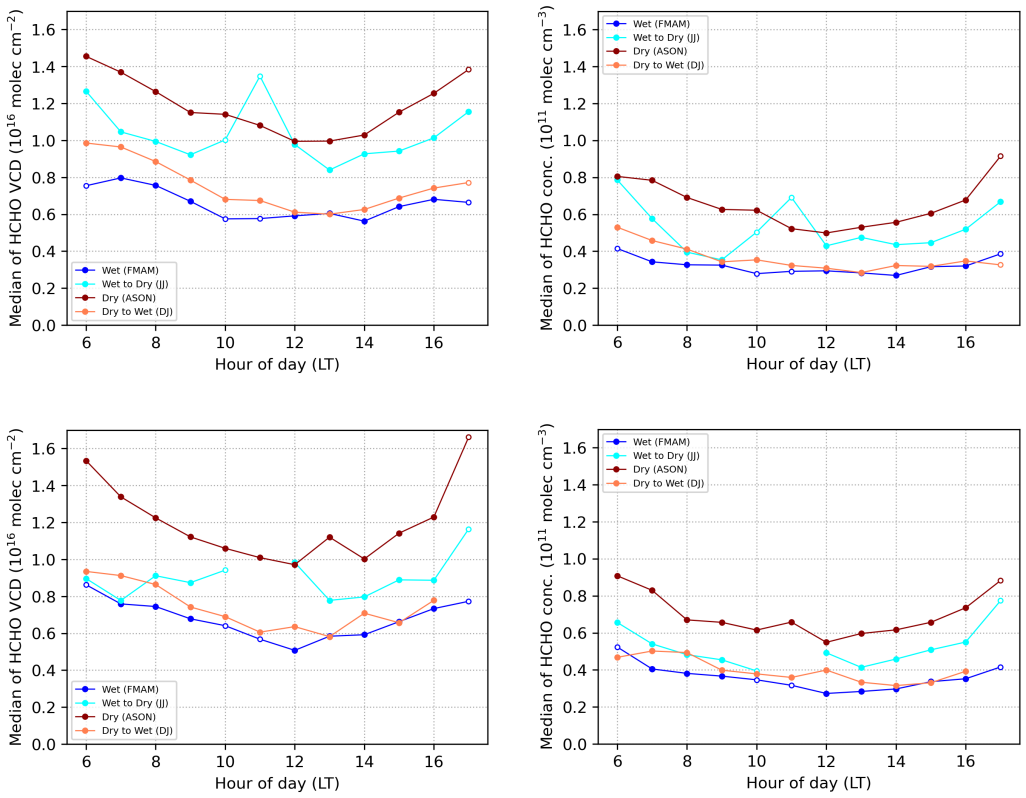


Figure 13.3: Diurnal cycles of HCHO VCDs (left) and concentrations at instrument altitude (right) for the ATTO 80 m (upper part) and 298 m (lower part) instruments for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

In all seasons, both quantities show a "U-shaped" diurnal cycle with highest HCHO abundances in the morning and evening, while the daily minimum occurs around noon/early afternoon (mostly 11 a.m. to 1 p.m.). This holds for both instruments. Statistical artefacts in the early mornings and late evenings are not probable at ATTO since all days have similar sunshine durations throughout the entire year. During the transition period from wet to dry season (June and July), a local maximum is found for VCDs and concentrations at instrument altitude around 11 a.m. for the 80 m instrument (cyan lines in the upper panels). Also in the 1-hour bins before and after slightly enhanced HCHO abundances can be observed. However, most of these bins are marked with open circles indicating that less than 5 % valid MAPA retrievals are obtained in these time bins. Given the short measurement period for the 298 m instrument, the 1-hour bin at 11 a.m. is not even displayed (lower panels) showing that less than 10 valid individual MAPA results are obtained. The 10 a.m. bin is also flagged. Therefore, it can be concluded that the observed local maximum in this season is an artefact probably caused by unfavourable measurement conditions and is thus likely not representative for the average conditions. In particular, the instruments point most likely too close to the sun (low RAA) leading to geometries that are not well captured by the LUTs used for the MAPA retrievals. As presented below,

also the aerosol retrievals show difficulties around noon, while no such effects are found for most of the other trace gases. The observed diurnal cycles are, at a first view, surprising since at ATTO the dominant HCHO source is expected to be secondary formation from biogenic VOCs which typically peaks around noon (lowest SZA). Consequently, also the maximum HCHO abundances are expected around noon or in the afternoon. Although an accumulation of VOCs in the boundary layer over the course of the day or the increase of the temperatures towards afternoon, both retarding the maximum of the HCHO abundance, may contribute to this unexpected diurnal variations, the main factor is probably that the two major loss processes, namely photolysis and oxidation by OH, also peak around noon. Nussbaumer et al. (2021) showed that already at mid-latitudes the dominant loss process of HCHO is photolysis and thus it can be expected that at ATTO this effect is even stronger. Further, also the OH concentrations are usually rather high in tropical regions (compare Sect. 2.2). Both loss processes together likely lead to particularly short lifetimes of HCHO around noon at ATTO. Altogether, despite also enhanced production from precursor VOCs caused by the high OH levels, apparently the loss processes exceed the production leading to a net loss of HCHO during the day yielding the observed diurnal variations. This is also supported by the finding that the diurnal variations are rather strongly pronounced in the dry season (dark red lines) and the transition period towards the dry season, where typically less clouds are present leading to enhanced photochemistry. The observation that very similar diurnal patterns are found for both quantities again suggests that the HCHO sources are rather homogenous at ATTO.

The annual cycles of glyoxal VCDs and concentrations at instrument altitude are presented in Fig. 13.4 and 13.5 for the 80 m and 298 m instrument, respectively.

Compared to HCHO, more variability is found for the glyoxal results of both instruments, e.g. large scatter of the daily means (blue dots in both figures) is obtained. This is caused by a generally very challenging glyoxal retrieval as described earlier in this work (i.e. strong water vapour interferences, low optical depths, etc.) leading to very noisy dSCDs and thus only little valid MAPA retrievals. As a consequence, no obvious annual cycle can be observed for the glyoxal VCDs. However, for the 80 m instrument (Fig. 13.4) at least in some years (e.g. 2020) a tendency towards slightly enhanced VCDs during the early dry season is visible, while in some wet seasons reduced glyoxal VCDs are measured (e.g. 2022). The latter is interesting since the wet season of 2022 was a particularly pronounced wet season with an above-average amount of precipitation. No such effect was observed for HCHO. Contrary to the VCDs, distinct annual variations of the concentrations at instrument altitude are found for both instruments (compare red lines in the middle panels of Fig. 13.4 and 13.5) with enhanced values during the dry seasons and lower values during the wet seasons. Also the scatter is apparently lower for the concentrations at instrument altitude. Similar to the VCDs, particularly low glyoxal concentrations are found at instrument altitude during the wet season of 2022.

As described above, only a few valid retrievals are obtained for the third measurement period of the 298 m instrument which started in April 2022. In

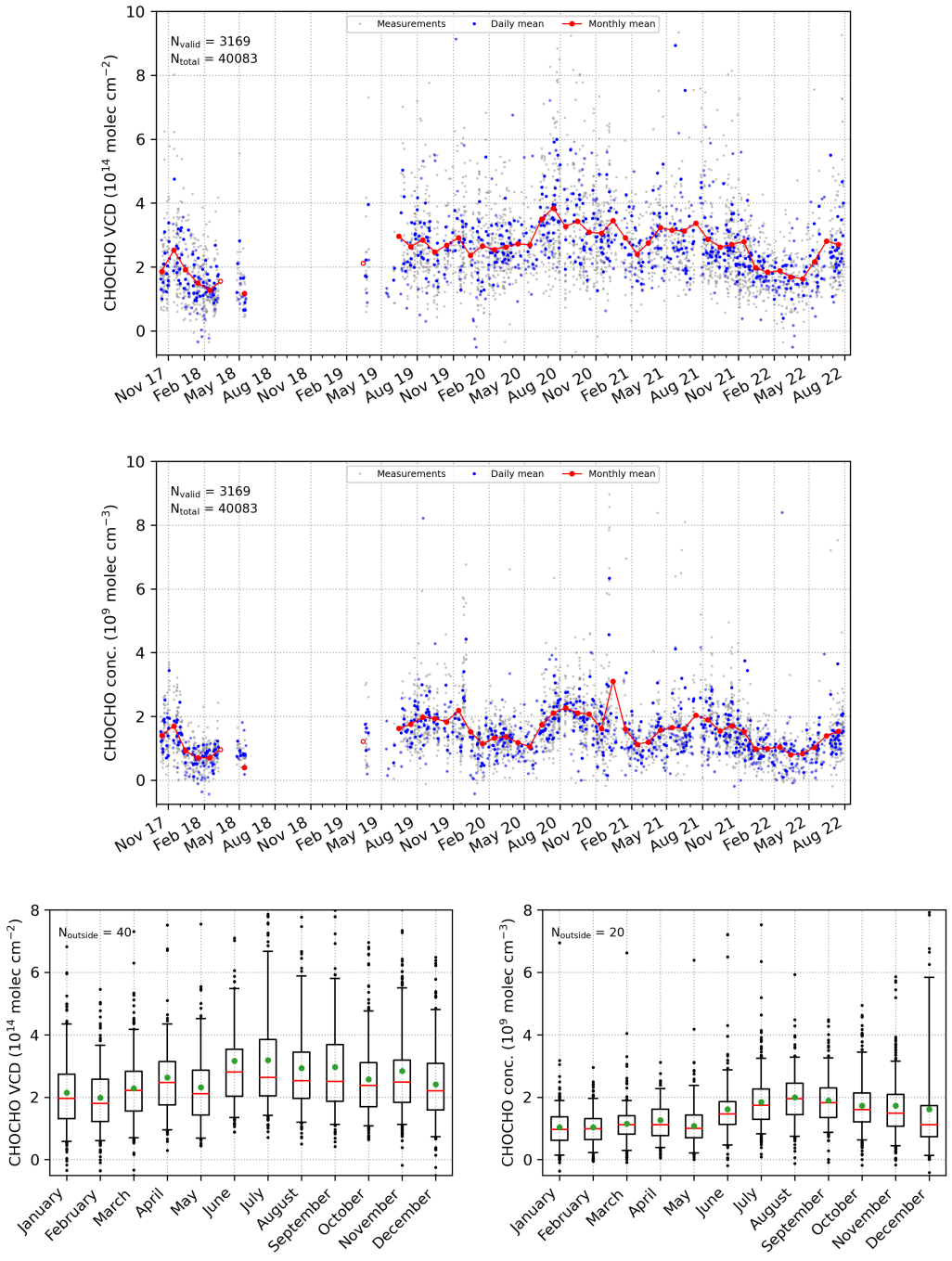


Figure 13.4: The two upper panels show time series of glyoxal VCDs and concentrations at instrument altitude for ATTO (80 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of glyoxal VCDs (left) and concentrations at instrument altitude (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

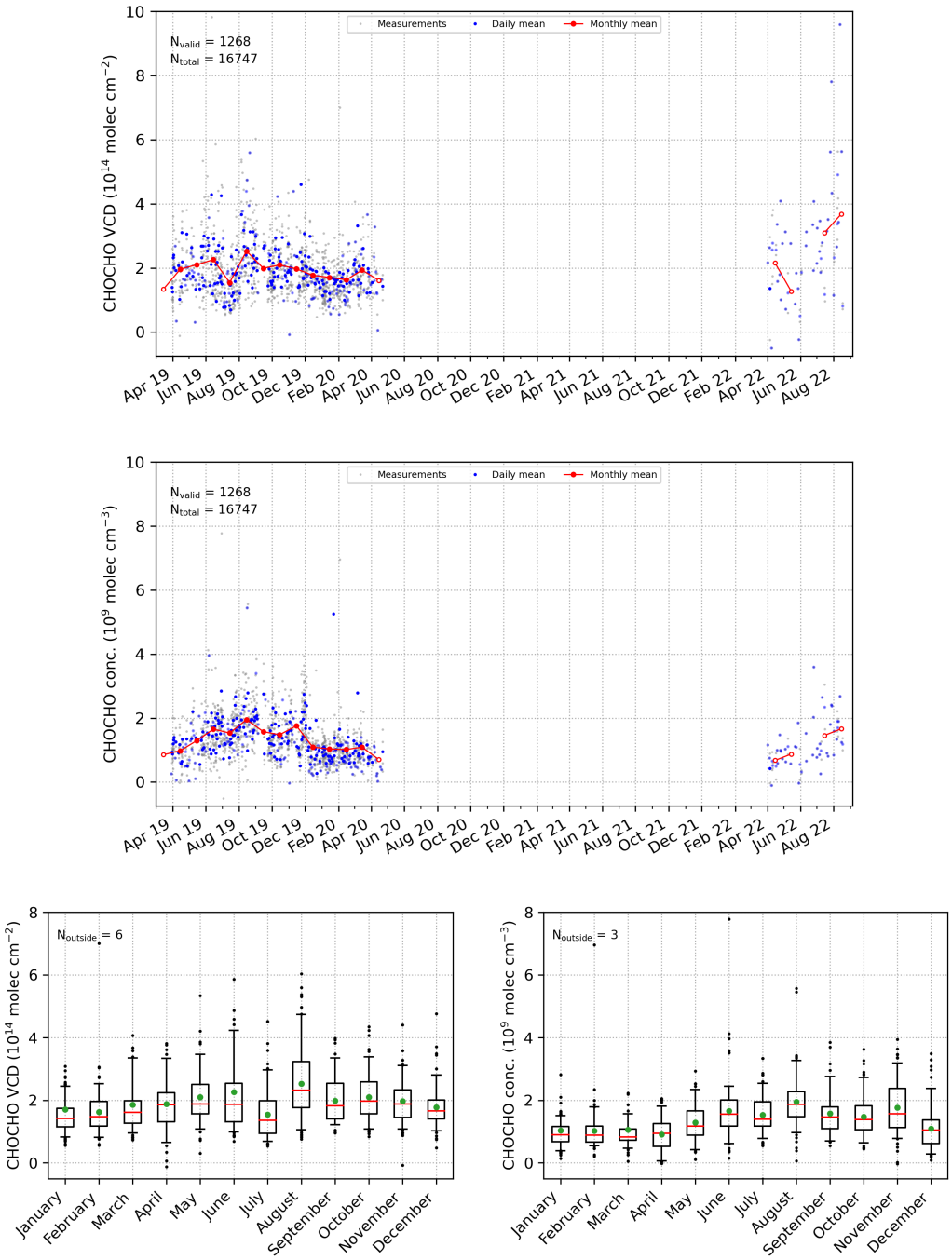


Figure 13.5: The two upper panels show time series of glyoxal VCDs and concentrations at instrument altitude for ATTO (298 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of glyoxal VCDs (left) and concentrations at instrument altitude (right) (data until 21 April 2020). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

particular, all daily means in this period, except one, are flagged for having less than 3 valid MAPA results. Thus, only data for the two first periods until 21 April 2020 are considered in the following.

Despite the large variability of the glyoxal results, the monthly box-whisker-plots which are shown in the two lowermost panels of Fig. 13.4 and 13.5 reveal systematic annual cycles for both quantities. For the VCDs (left lowermost panels), this annual cycle is only weakly pronounced, whereby the large variability is reflected by large boxes and long whiskers. Nevertheless, for the 80 m instrument systematically enhanced monthly mean glyoxal VCDs are found from June to September which corresponds to the transition period towards the dry season and the early dry season itself. In these months, the monthly means fluctuate around ca. 3×10^{14} molec cm^{-2}, while the monthly means are lower in the early wet season (February and March) with values of around and slightly above 2.0×10^{14} molec cm^{-2}. Only one full annual cycle (March 2019 to April 2020) is captured by the 298 m instrument. A comparison with the 80 m instrument shows that in this period no strong annual variations occurred (uppermost panels of corresponding figures). Hence, no distinct annual pattern emerges from the box-whisker-plot for the VCDs of the 298 m instrument and large variability is found. In addition, apparently smaller VCDs are found for the 298 m instrument. Contrary to the VCDs, the concentrations at the instrument altitudes of both instruments show rather clear annual cycles which are very similar to the ones of HCHO. Again highest concentrations at instrument altitude are found during the dry season where monthly means of approximately 2×10^9 molec cm^{-3} are obtained in August for both instruments. Lowest concentrations at instrument altitude are obtained in the wet season with monthly means of approximately 1.1×10^9 molec cm^{-3} for both instruments. Although similar values are stated here, a detailed look reveals that generally lower glyoxal concentrations are found at 298 m (for details refer to Sect. 13.8). Together with the observation that also lower VCDs are measured with the 298 m instrument, this suggests that also at ATTO the glyoxal profiles are quite shallow. In particular, already in the lowest 200 metres (height difference between both instruments) the glyoxal concentrations are declining. This is also supported by the finding that only the concentrations at instrument altitude show a clear annual cycle. Apparently, the increased concentrations at instrument altitude during the dry season are balanced by shorter atmospheric lifetimes in this season leading to only small enhancements of the VCDs. Further, the similar patterns of the glyoxal and HCHO concentrations at instrument altitude suggest similar sources of both species (mostly biogenic VOCs), but different processing in the atmosphere since only the HCHO VCDs are significantly increased during the dry season. The decrease of the glyoxal concentrations at instrument altitude starts directly after reaching the maximum in August, whereas the HCHO concentrations at instrument altitude stay at similar levels until October[1]. This finding probably indicates different chemical regimes within the dry season yielding different amounts of glyoxal but similar abundances of HCHO.

1 In principle, the same holds for the glyoxal VCDs. However, these are very noisy and thus have to be considered with care.

Given the large variability of the VCDs, no consistent differences between the VCDs obtained with either SF are found for both instruments. For the concentrations at instrument altitude, the fixed SF of 0.8 leads to slightly larger values, whereby the general patterns are not affected.

Figure 13.6 summarises the average diurnal variations of glyoxal VCDs (left) and concentrations at instrument altitude (right) for both ATTO instruments in the different seasons.

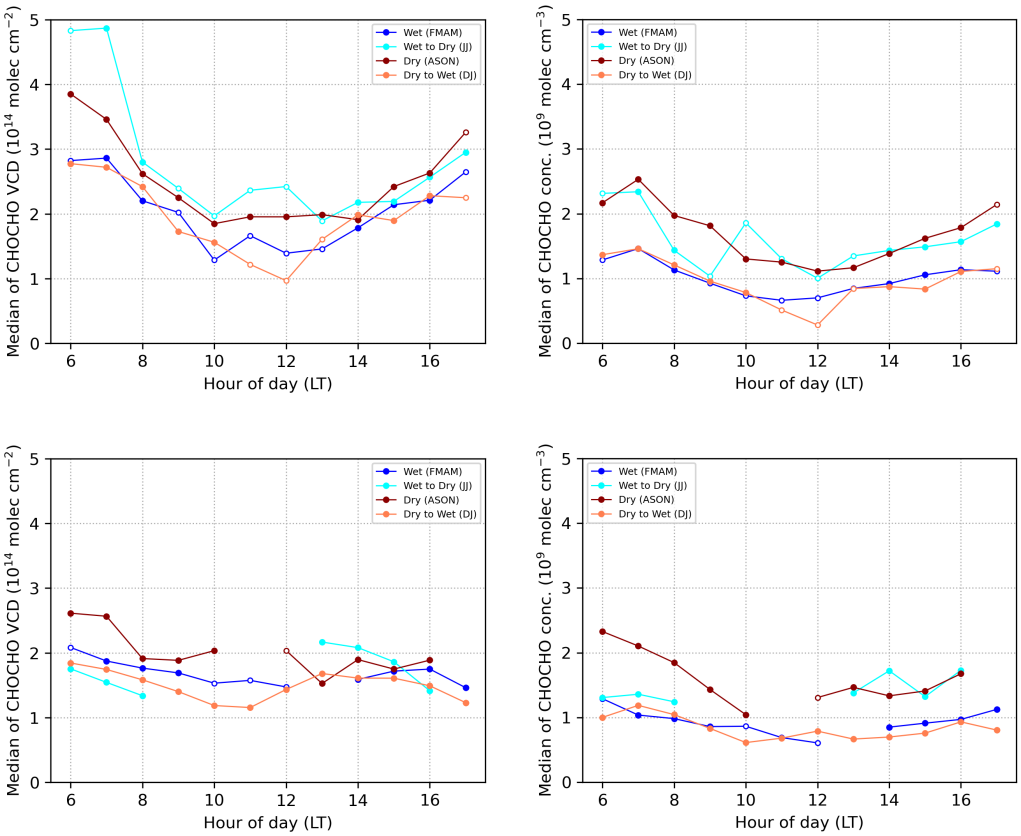


Figure 13.6: Diurnal cycles of glyoxal VCDs (left) and concentrations at instrument altitude (right) for the ATTO 80 m (upper part) and 298 m (lower part) instruments for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season. For the 298 m instrument, only data until 21 April 2020 are included.

The diurnal variations of the glyoxal abundances are very similar to the diurnal cycles of HCHO but show more variability. For both quantities and all seasons, highest values are observed in the mornings and evenings, whereas minima of both quantities are found around noon (mostly 11 a.m. to 1 p.m.). Compared to HCHO, the observed diurnal patterns of glyoxal for the 80 m instrument are (relatively seen) stronger pronounced. The diurnal variations of the glyoxal abundances measured with the 298 m instrument, however, are weaker and the (relative) amplitudes of the variations are comparable to the ones of HCHO. This finding again indicates that the glyoxal abundances decrease significantly between the altitudes of the two instruments. As a consequence of the latter, the

diurnal variations of the glyoxal abundances are only weakly visible at 298 m. The pronounced artefact around noon in the transition period between wet and dry season, which was very prominent for HCHO, is barely visible for the glyoxal VCDs and only slightly enhanced values are found in this season around noon (compare cyan line in the upper left panel). Similar to HCHO, also for glyoxal only little valid MAPA retrievals are obtained in the affected 1-hour bins (open circles). In particular, many 1-hour bins around noon are not displayed at all for the 298 m instrument since less than 10 valid MAPA retrievals are obtained. This stresses systematic problems with the measurements at ATTO in some periods of the year. The in general very similar diurnal patterns of formaldehyde and glyoxal stress the close relation of both species which is apparently stronger at ATTO compared to the other stations. More precisely, at ATTO both trace gases originate almost solely from biogenic sources, whereas at the other stations more or less strong anthropogenic influence leads to comparatively larger differences between both species. Nevertheless, also at ATTO small differences in the annual (compare above) and diurnal patterns are observed. In particular, the finding that the diurnal cycle of glyoxal for the 80 m instrument is stronger pronounced than for HCHO indicates a stronger diurnal variation of the loss processes of glyoxal which suggests different atmospheric processing of both species.

The annual patterns of the NO_2 abundances at ATTO are presented in Fig. 13.7 and 13.8 for the 80 m and 298 m instrument, respectively. Note that the y-axis scales in these figures are one order of magnitude smaller compared to the other stations.

Given its remote location far from major anthropogenic emission sources (compare Sect. 4.4), the ATTO site experiences in general very low NO_2 pollution levels throughout the year which are close to the detection limit of the used MAX-DOAS instruments. Together with high water vapour abundances leading to spectral interferences in the DOAS fit and frequently occurring rapidly changing cloud cover, this leads to only few valid MAPA retrievals (less than 5 %) and a large variability of the results. The latter is particularly the case for the 80 m instrument (compare large scatter in Fig. 13.7). Nevertheless, systematic variations of the NO_2 abundances are found for both instruments. Especially, the VCDs are significantly enhanced during the dry season while they scatter around values close to 0 in the wet season. For the concentrations at instrument altitude, overall the same annual cycles as for the VCDs are obtained, however, even larger variability is observed especially for the 80 m instrument with many values of practically 0. During the enhancements in the dry season months, daily mean NO_2 VCDs of slightly below 0.4×10^{15} molec cm^{-2} occur regularly, while single daily means can reach up to ca. 1.0×10^{15} molec cm^{-2} for the 80 m instrument. The concentrations at instrument altitude are typically in the order of 0.25×10^{10} molec cm^{-3} during these periods. The described annual cycle can be explained by two main reasons. Firstly, the transitions between the two major seasons (dry and wet) are associated with a change of the dominating wind direction and thus the main fetch region (compare Sect. 4.4). While north-easterly winds prevail during the wet season which bring pristine and clean air masses towards ATTO, more easterly winds dominate during the dry season which are influenced by anthropogenic pollution

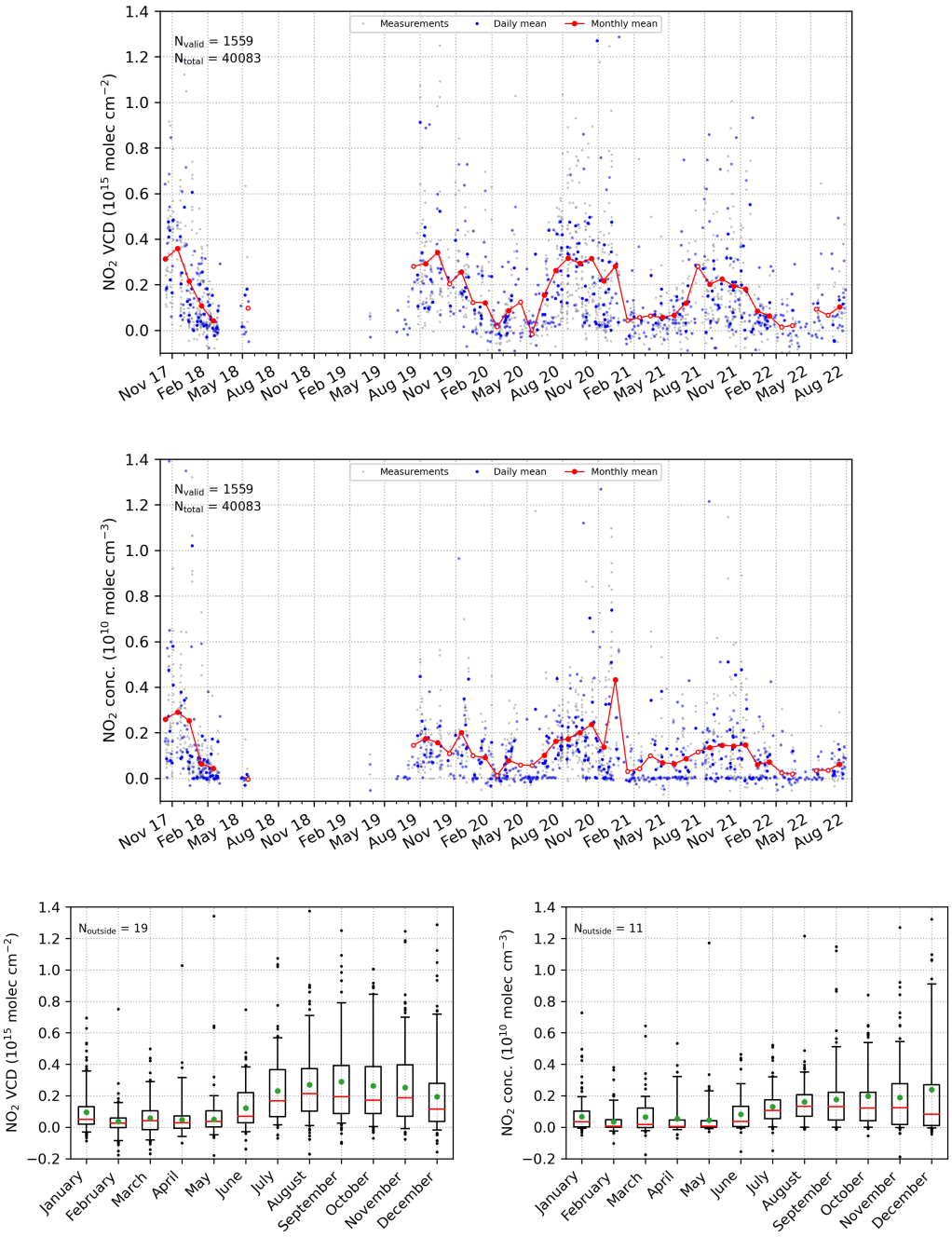


Figure 13.7: The two upper panels show time series of NO_2 VCDs and concentrations at instrument altitude for ATTO (80 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of NO_2 VCDs (left) and concentrations at instrument altitude (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

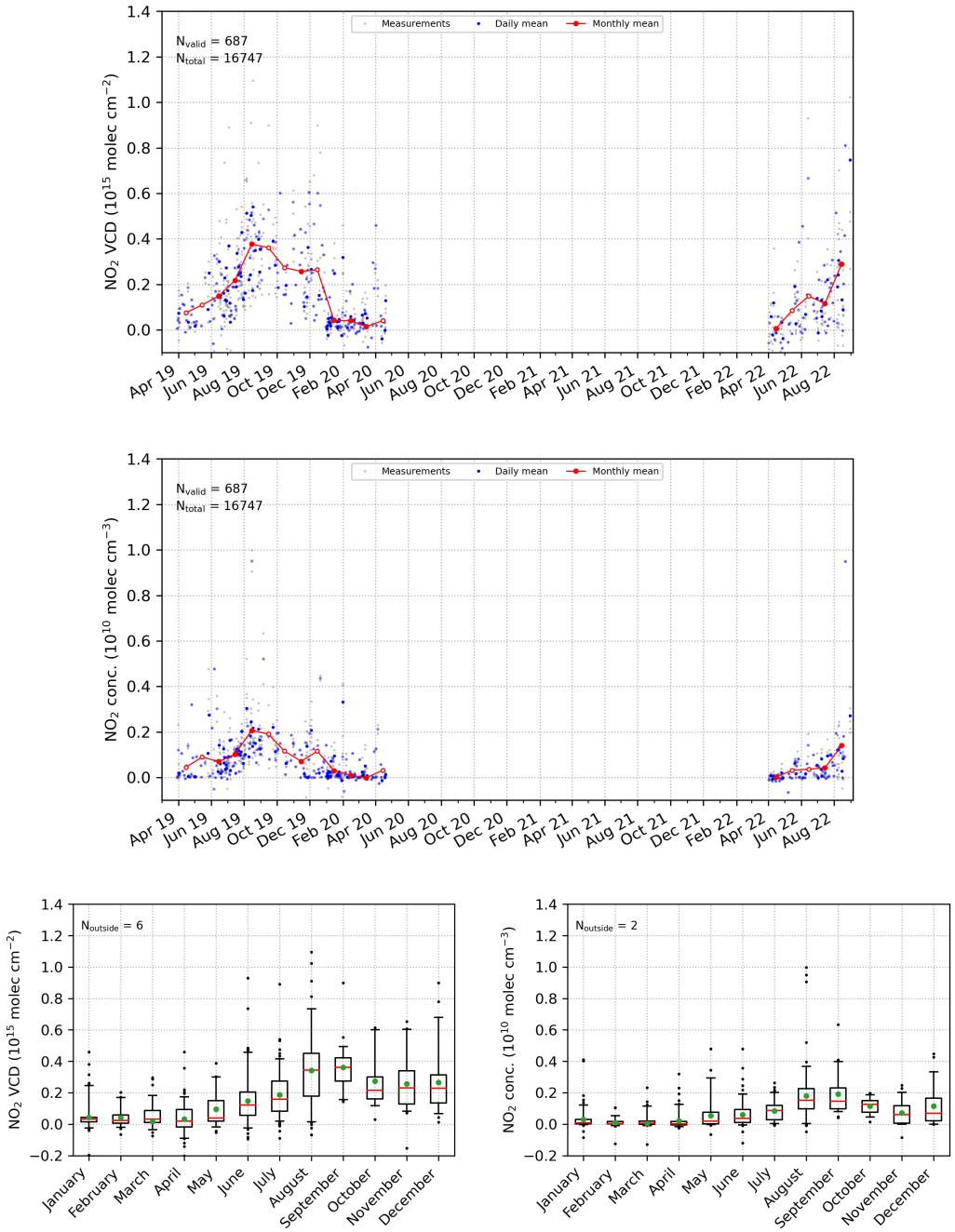


Figure 13.8: The two upper panels show time series of NO_2 VCDs and concentrations at instrument altitude for ATTO (298 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. Open red circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of NO_2 VCDs (left) and concentrations at instrument altitude (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

from more densely populated regions and forest fires. Secondly, during the dry season both natural forest fires and man-made deforestation fires occur frequently. Since the ATTO site is located in a protected area, the largest of these fires typically do not occur in the immediate vicinity of ATTO but more to the south and south-east. However, given the dominating wind direction in this season, pollution is regularly transported to the ATTO site. Details of the dependence of the observed NO_2 abundances on the wind direction are discussed in Sect. 13.3. The finding that larger variability is found for the 80 m instrument probably suggests a weak effect of small local fires. It also cannot be excluded that in some instances emissions caused by the ATTO site itself, e.g. from Diesel generators needed for power supply, might contribute to the latter.

The monthly box-whisker-plots of NO_2 VCDs (left) and concentrations at instrument altitude (right) presented in the two lower panels of Fig. 13.7 and 13.8 confirm the described annual cycles with highest NO_2 abundances during the dry season and values close to 0 in the wet season. In particular, monthly mean VCDs of approximately 0.27×10^{15} molec cm^{-2} and 0.05×10^{15} molec cm^{-2} are found in the dry and wet season, respectively, for the 80 m instrument. For the concentrations at instrument altitude, only the medians (red lines in the box-whisker-plots) show a smooth annual variation, while the monthly means are strongly affected by the large variability where single largely enhanced values lead to rather high monthly means in December. Monthly medians of marginally above 0 and around 0.13×10^{10} molec cm^{-3} are obtained for the two main seasons. The maximum monthly means/medians stated here are roughly one to one and a half orders of magnitude lower than the minima at the other three stations stressing the pristine character of the ATTO site. Only one dry season is covered by the 298 m instrument. In this particular year (2019), the enhancement of the NO_2 abundances started comparably late. For instance, in 2020 already July showed enhanced NO_2 abundances, while in 2019 the enhancements started only in August (compare monthly means in the uppermost panel of Fig. 13.7). This is also reflected in the monthly box-whisker-plots. A detailed comparison of the results of both instruments is provided in Sect. 13.8. Apart from biomass burning events which occasionally affect the site, at ATTO NO_2 has different sources than HCHO and glyoxal. Nevertheless, the annual cycle of NO_2 is similar to the ones of HCHO and glyoxal showing that the alternation between wet and dry season is the major driver of both biogenic and anthropogenic activities and thus determines the trace gas abundances at ATTO.

For both instruments, the analysis with a fixed O_4 scaling factor of 0.8 yields the same annual patterns. While for the VCDs mainly the variability is enhanced for the fixed SF, larger variability and higher values are obtained for the concentrations at instrument altitude. The latter is strongest pronounced in the dry season.

Median diurnal cycles of NO_2 VCDs (left) and concentrations at instrument altitude (right) in the four seasons for both instruments are shown in Fig. 13.9.

Given the absence of significant local anthropogenic emissions, the diurnal variations of the NO_2 abundances at ATTO show fundamentally different patterns compared to the other stations. In particular, the diurnal cycles of the NO_2 VCDs

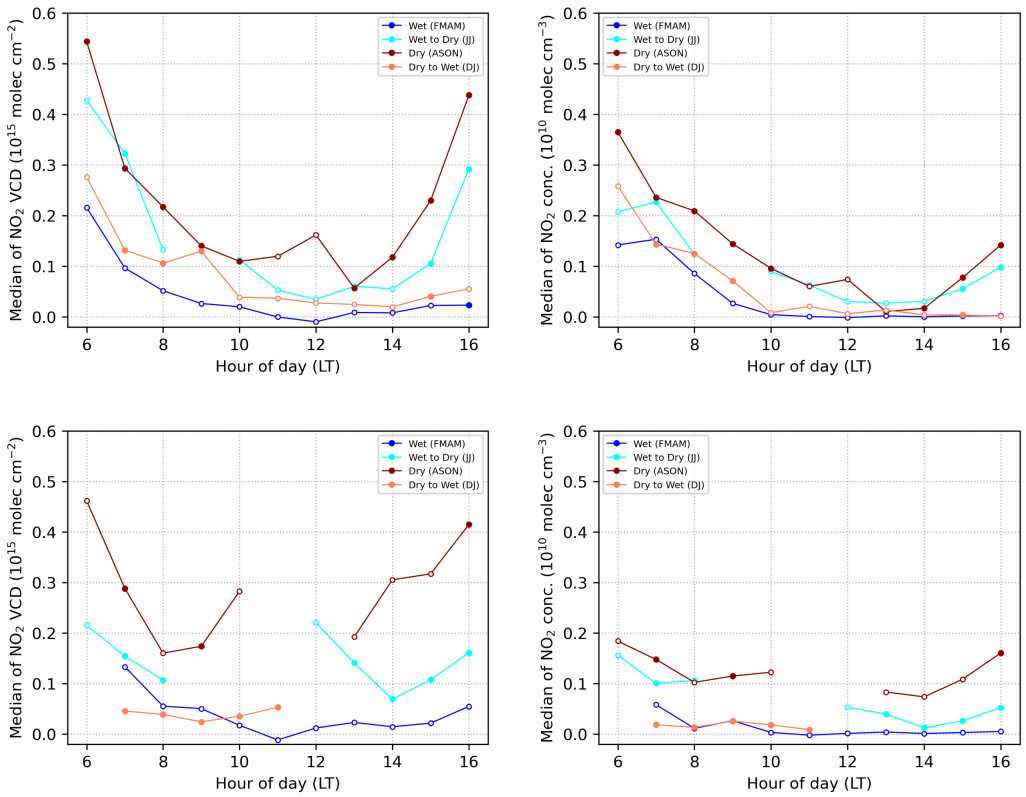


Figure 13.9: Diurnal cycles of NO_2 VCDs (left) and concentrations at instrument altitude (right) for the ATTO 80 m (upper part) and 298 m (lower part) instruments for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

are comparable to the ones of HCHO and "U-shaped" cycles with elevated abundances in the morning and evening hours and minimum NO_2 around noon are in principle found for all seasons. However, in the transition period from dry to wet season and in the wet season itself, only marginal increases of the VCDs towards sunset are visible. For the concentrations at instrument altitude, also "U-shaped" diurnal variations are found in the transition period from wet to dry season and in the dry season itself, whereby the increase of the concentrations at instrument altitude towards sunset is less pronounced than the decrease during the morning hours. No increases of the concentrations at instrument altitude in the late afternoon are found at all during the transition from dry to wet season and in the wet season itself, where the NO_2 concentrations at instrument altitude are very close to 0 around noon and during the afternoon. It should be noted that, due to only little valid MAPA retrievals for NO_2, many 1-hour bins are flagged (open circles) for having less than 5 % valid data. In principle, similar diurnal cycles and differences between the two quantities can be observed for the 298 m instrument but many 1-hour bins are flagged or not even shown at all. Since no significant local emissions of NO_2 exist, the observed diurnal cycles can mainly be explained by chemical processes. Similar to HCHO and glyoxal, the "U-shape" which is observed for the VCDs is mainly caused by the dominance of loss

processes (photolysis and oxidation by OH) during day time which peak around noon. Compared to HCHO, the diurnal variations are stronger for NO_2 probably caused by the lack of significant local NO_2 sources which leads to loss processes dominating during the day. The oxidation of NO_2 by OH yields nitric acid (HNO_3) which is highly water-soluble and can thus lead to a removal of reactive nitrogen from the atmosphere during rain events. This effect might be particularly important in the wetter part of the year where typically more precipitation occurs (transition dry to wet and wet season), hindering a strong increase of the NO_2 abundances during late afternoon in these seasons as described above. This effect might also contribute to the finding that the concentrations at instrument altitude do not increase at all towards sunset in the respective seasons. However, also the rise of the boundary layer during the course of the day might contribute to the finding that the "U" is asymmetric for the concentrations at instrument altitude in all seasons. Future studies should also focus on the influence of emissions of reactive nitrogen from biogenic sources, such as soils, on the diurnal cycles of NO_2.

Lastly, annual cycles of AOD and aerosol extinction at instrument altitude at 367 nm are presented in Fig. 13.10 and 13.11.

For both instruments, both aerosol quantities show distinct annual cycles. Highest values are typically found during the dry season, while lower AODs and extinctions at instrument altitude occur in the wet season. The annual variations are rather clear for the AODs, whereas more variability is found for the extinctions at instrument altitude. The exceptionally high monthly mean extinction at instrument altitude in May 2019 observed for the 80 m instrument is an artefact caused by only a few individual measurements since the instrument was only restarted after a maintenance break in this month. However, this month is flagged anyhow for having less than 30 valid MAPA retrievals. In general, reasonable agreement between MAX-DOAS and AERONET AODs is found. However, similar to the observations at Bayfordbury and Mainz, the highest and lowest values are not well captured by the MAX-DOAS data set. Especially when the ATTO site is affected by polluted air masses originating from forest fire regions, individual[2] AERONET AODs clearly exceed 1 which is rarely the case for the MAX-DOAS AODs since most values above 0.5 are flagged by MAPA. This is clearly below 2 which is the threshold in the standard MAPA configuration for raising a warning and means that not the AOD flag but other flags are raised. This is a possible starting point to further improve the MAPA algorithm and its agreement with AERONET aerosol results in future.

The described annual cycles are also represented by the statistical analysis of both quantities shown in the lowermost panels of Fig. 13.10 and 13.11. Highest AODs are observed from August to October corresponding to the largest part of the dry season, where monthly means of around 0.25 are obtained for the 80 m instrument. Starting already in November, the AODs drop during the transition towards the wet season, where their minimum values of approximately 0.12 are observed. In principle, the extinctions at instrument altitude show the same annual

2 Note that only daily means based on the overlap between both data sets are shown in the presented figures.

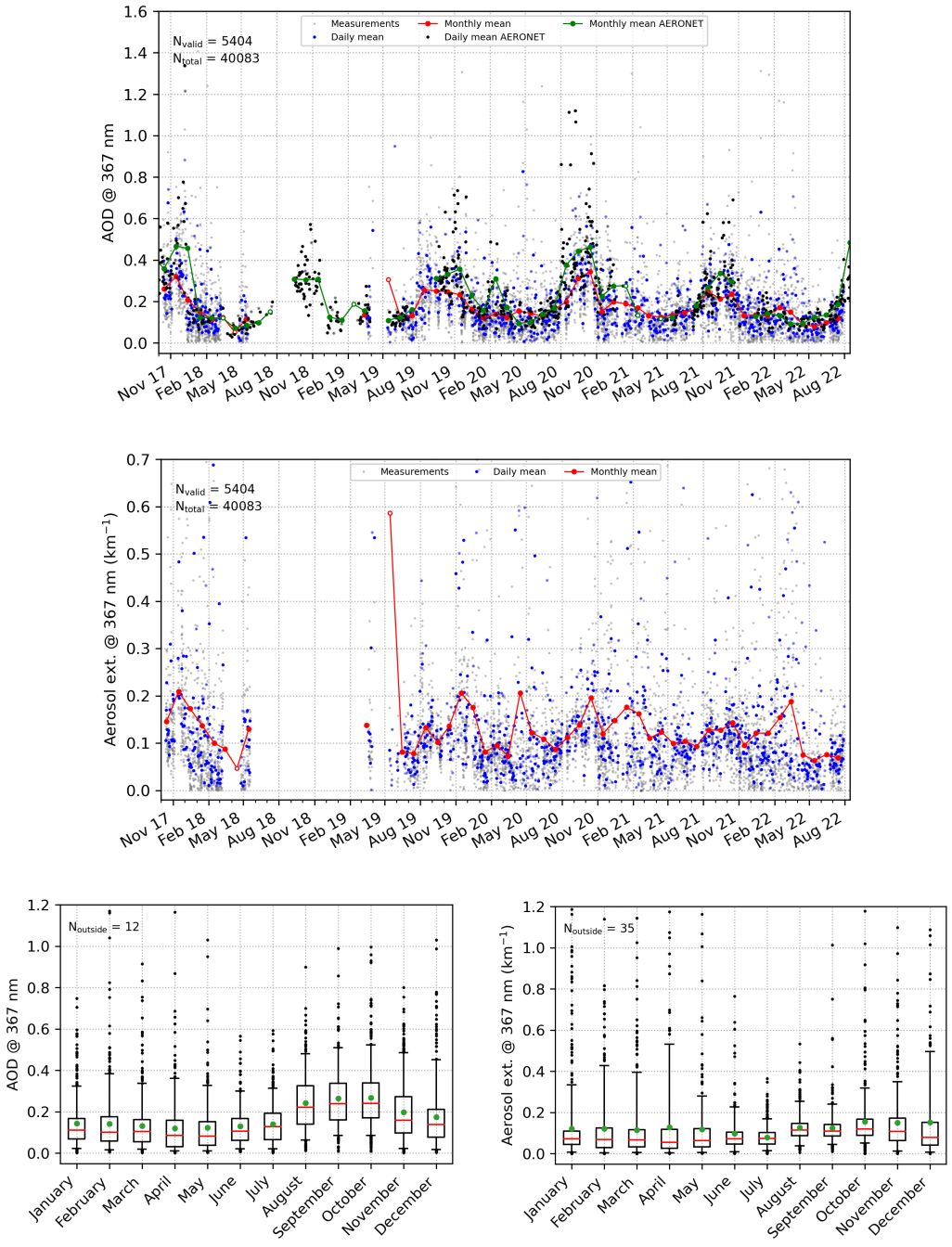


Figure 13.10: The two upper panels show time series of AOD and aerosol extinction at instrument altitude at 367 nm for ATTO (80 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and aerosol extinction at instrument altitude (right) at 367 nm. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

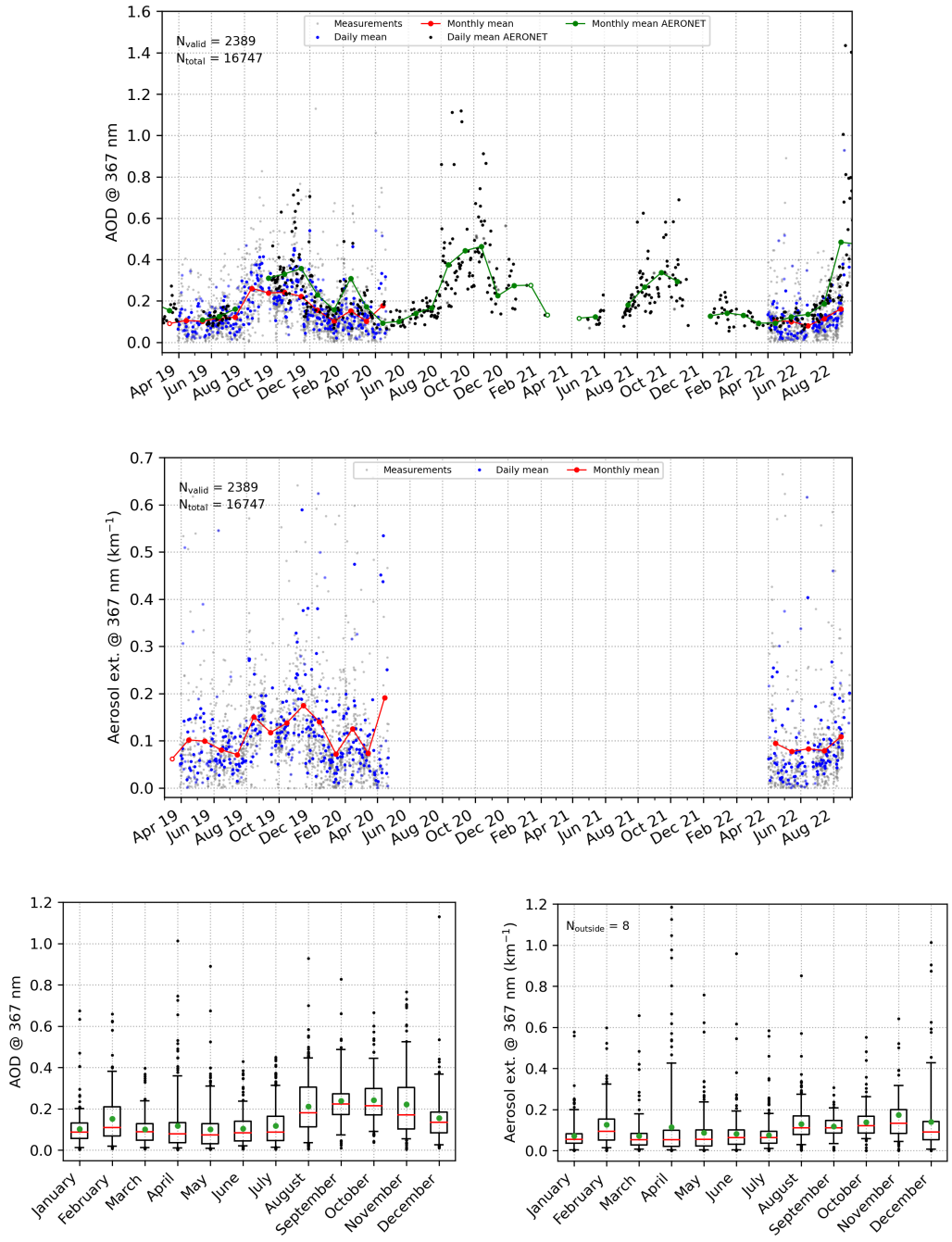


Figure 13.11: The two upper panels show time series of AOD and aerosol extinction at instrument altitude at 367 nm for ATTO (298 m). Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 30 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and aerosol extinction at instrument altitude (right) at 367 nm. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

cycle, however, very large variability is found outside the dry season with many outliers towards very high values (compare whiskers in these seasons) which is probably caused by the frequent occurrence of rapidly changing cloud coverage and precipitation. These outliers have a strong influence on the monthly means and no clear annual cycle can be observed for these. Nevertheless, a rather smooth annual cycle is found for the monthly medians (red lines in the box-whisker-plots) with maximum and minimum extinctions at instrument altitude of ca. $0.12\,km^{-1}$ and $0.06\,km^{-1}$ during the dry and wet season, respectively. The same findings are valid for the 298 m instrument but slightly lower absolute values are found here. Similar to the species discussed before, this might indicate small systematic differences between both instruments which are further assessed in Sect. 13.8. Like for NO_2 which shows overall similar annual variations, the annual cycle of the aerosol results can be explained by the dominant wind direction at ATTO which changes characteristically between the two major seasons (compare Sect. 4.4). Associated with this change, also the characteristics of the probed air masses change from very pristine air masses during the wet season to potentially polluted air masses in the dry season. For details, refer to the descriptions regarding NO_2. Interestingly, June and July are not associated with clearly enhanced aerosol abundances which is the case for NO_2 (lower left panel of Fig. 13.7). Further, NO_2 remains elevated also in November contrary to aerosol. Since air masses probed at ATTO under more easterly winds (dry season) are not only influenced by forest fires which lead to both high aerosol and NO_2 content but also by pollution from more densely populated regions which primarily leads to enhanced NO_2, these differences between both species indicate that the changes of the dominant wind direction take place in June/July and again in November/December. The fires then take place when the dry season is at its climax leading to enhanced aerosol abundances in a more confined time period only. Further, the formation of SOA might play a role in the dry season as HCHO abundances are also elevated in this season. While the AODs observed during the wet season are amongst the lowest of all stations, the AODs in the dry season are slightly above the annual maxima observed at Bayfordbury and Mainz. At Mohali, however, the minimum AODs are larger than the maximum values at ATTO.

Consistent to the findings at the other stations, the analysis with the fixed O_4 SF of 0.8 yields larger values for both aerosol quantities and the scatter of the results is largely increased. No major differences in the annual cycles obtained with either scaling factor are found. Nevertheless, some differences in the annual cycle of the AODs can be identified for the 80 m instrument most likely caused by the strongly enhanced variability which almost hides the annual cycle.

The diurnal cycles of AOD (left) and extinction at instrument altitude (right) at 367 nm found in the different seasons for both instruments are depicted in Fig. 13.12.

The diurnal variations of the AODs show generally large variability and contrary to the other species almost no clear diurnal cycles can be observed for large parts of the year. Only in the dry season (dark red curves) a tendency of increasing AODs in the afternoon is visible. These findings apply to both instruments, but the observed increase is apparently stronger pronounced for the 298 m instrument

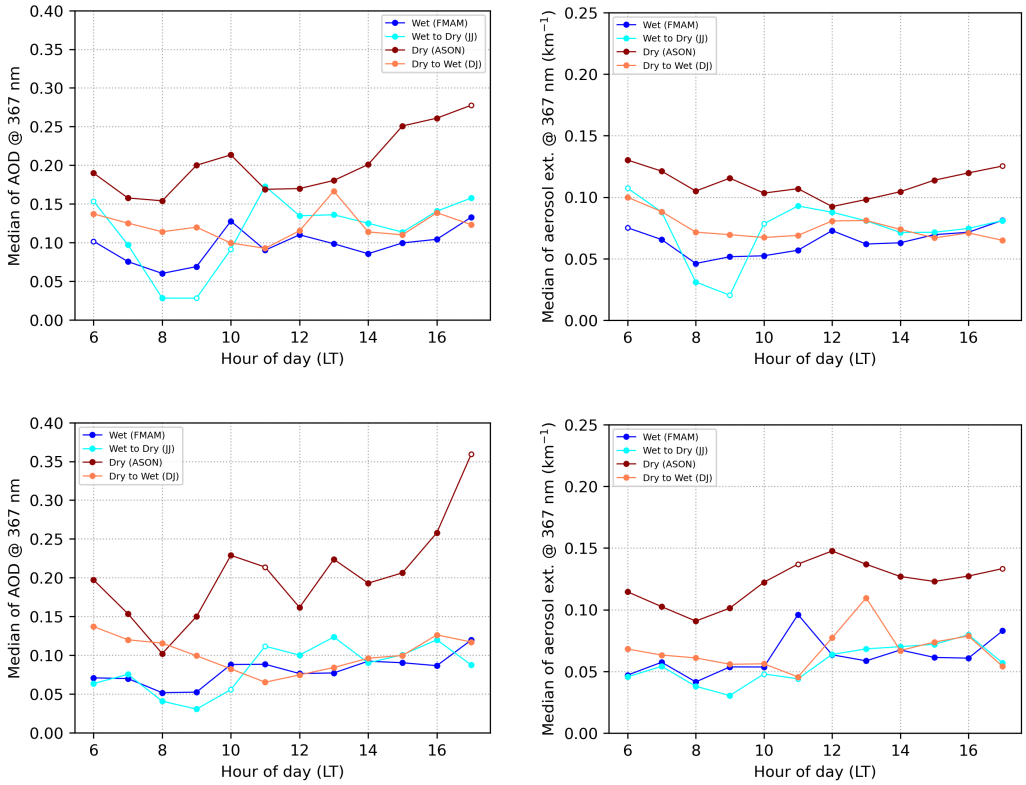


Figure 13.12: Diurnal cycles of AOD (left) and extinction at instrument altitude (right) at 367 nm for the ATTO 80 m (upper part) and 298 m (lower part) instruments for different seasons indicated by the different colours. Open circles label hours with less than 5 % valid MAPA retrievals in the respective season.

which additionally shows also a slight increase of the AODs during the course of the day in the wet season. However, since only one and a half years are covered by the measurements at 298 m, this has to be considered with care. Also for the aerosol extinctions at instrument altitude, no clear diurnal cycles are found in most of the seasons. Again only in the dry season systematic variations are obtained for both instruments. Nevertheless, these are not very strong compared to the diurnal variations of the AODs in the dry season and inverse behaviours are found for the two instruments. While a weak "U-shape" is observed for the 80 m instrument, the extinctions show an inverse "U" at 298 m. The shape for the 80 m instrument might be explained by a rising boundary layer in the course of the day with rather constant aerosol content until noon (indicated by the AODs) which leads to a slight decrease of the aerosol content in the lowest atmospheric layers. During the afternoon when the total aerosol abundances increase, also the extinctions at instrument altitude are increasing again. No clear explanation can be provided for the shape of the diurnal variations of the 298 m instrument.

A further observation is that especially in the transition period between wet and dry season (cyan lines), but also in the wet season itself, pronounced dips of both aerosol quantities around 8 to 9 a.m. occur which are stronger pronounced for the 80 m instrument. Also for this feature so far no conclusive explanation can be

given, but it is probably caused by systematic effects affecting the measurements in these 1-hour bins in the respective seasons or by measurement conditions that are not well represented by the LUTs used for the MAPA inversion (e.g. the aerosol scattering phase-function for small scattering angles). This is also supported by the fact that the 1-hour bins at 9 and 10 a.m. are flagged for having less than 5 % valid retrievals. For the scenario using a fixed O_4 SF, even less than 10 individual valid results are obtained in some of these 1-hour bins and thus they are not displayed at all for the 80 m instrument (compare upper part in Fig. 13.13).

Similar to Mohali, significantly different diurnal variations of the aerosol results are obtained using the fixed O_4 SF. Figure 13.13 presents the corresponding diurnal cycles of AOD (left) and extinction at instrument altitude (right).

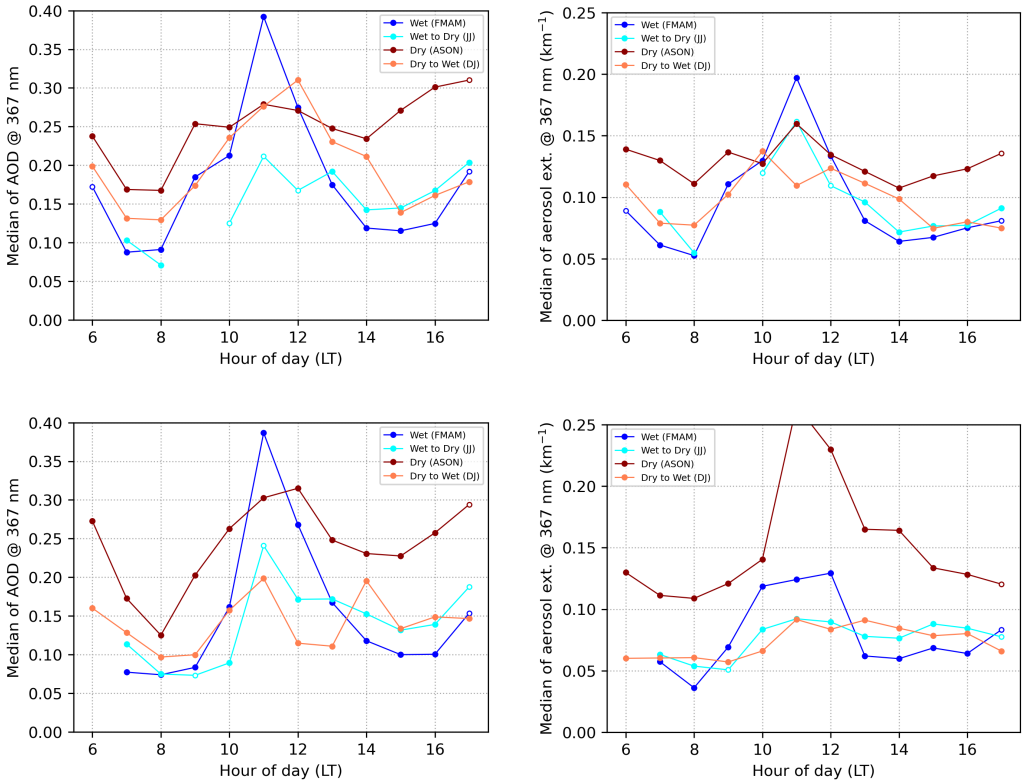


Figure 13.13: Same as Fig. 13.12 but for a fixed O_4 SF of 0.8 applied for the aerosol retrieval.

For both aerosol quantities and both instruments, unexpected and very strongly pronounced enhancements are found in the 1-hour bins centred around 11 a.m. which are most prominent in the wet season (blue lines). In addition, between 7 and 8 a.m. dips in the aerosol abundances can be observed for most cases. While in the transition period from wet to dry season the affected time bins have mostly below-average amounts of valid retrievals (open circles for the cyan curves or not displayed at all), this is not the case in the remaining seasons indicating suspicious valid retrievals. Also for HCHO an artefact was found around 11 a.m. which was only visible during the transition from wet to dry season (compare Fig. 13.3). As

already explained in detail for Mohali (compare Sect. 12.2), these unexpected patterns are attributed to effects that depend on SZA and RAA, probably caused by viewing geometries that are not well captured by the LUTs used for the MAPA retrievals. In particular, these problematic geometries occur at small RAAs and SZAs where the instruments point close to the sun and the forward peak of aerosol scattering becomes important. For details, refer to the Mohali section (Sect. 12.2). It should be mentioned that the unexpected diurnal patterns of the aerosol results for a fixed scaling factor were less pronounced and smoother at Mohali. Moreover, also the annual cycle was affected by the choice of the O_4 SF at Mohali which is not the case at ATTO. Assuming that these artefacts are indeed related to problematic viewing geometries, these differences can be explained by the fact that Mohali is located at higher latitudes than ATTO. Hence, solar positions change less rapidly during the course of the day at Mohali leading to a smoother diurnal pattern. In addition, the effects are generally less pronounced since the geometries are less extreme at Mohali, e.g. the minimum SZA is larger at Mohali. Lastly, weaker annual variations of the geometries occur at ATTO since it is located close to the equator which leaves the annual cycle mostly unchanged by the choice of the SF. As stated earlier, the variable SF probably compensates for such effects. Since such a compensation is potentially problematic, these findings require future investigations to fully understand these artefacts and the behaviour of the variable O_4 scaling factor.

Besides the species discussed in detail in this section, also HONO and water vapour retrievals were performed. Like for Bayfordbury, the HONO abundances at ATTO are below the detection limit of the used instruments and thus no enhanced HONO was measured by either instrument. Therefore, HONO is not considered in the further investigations of the ATTO results. For water vapour, surprising findings are obtained for the annual cycles which are discussed in Sect. 13.7, while practically no diurnal cycles are obtained for the H_2O abundances.

13.3 INFLUENCE OF METEOROLOGICAL CONDITIONS

Next, the dependencies of trace gas and aerosol abundances measured at ATTO on air temperature, wind speed and wind direction are investigated. For that, the data were selected as described in Sect. 9.3 and visualised as outlined in Sect. 10.3.

The uppermost panels of Fig. 13.14 and 13.15 show hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) as function of air temperature for the 80 m and 298 m instrument, respectively.

Reflecting the annual variations of the atmospheric HCHO content at ATTO (e.g. Fig. 13.1) with enhanced HCHO in the warmer dry season, both quantities show generally increasing values with increasing temperature. While this pattern is rather clear for the concentrations at instrument altitude, more variability is found for the VCDs probably indicating that the VCDs are, in addition to the temperature dependence, also influenced by transport processes, e.g. from biomass burning regions (compare below). Besides this general tendency, absolutely highest values of both quantities are observed at temperatures between

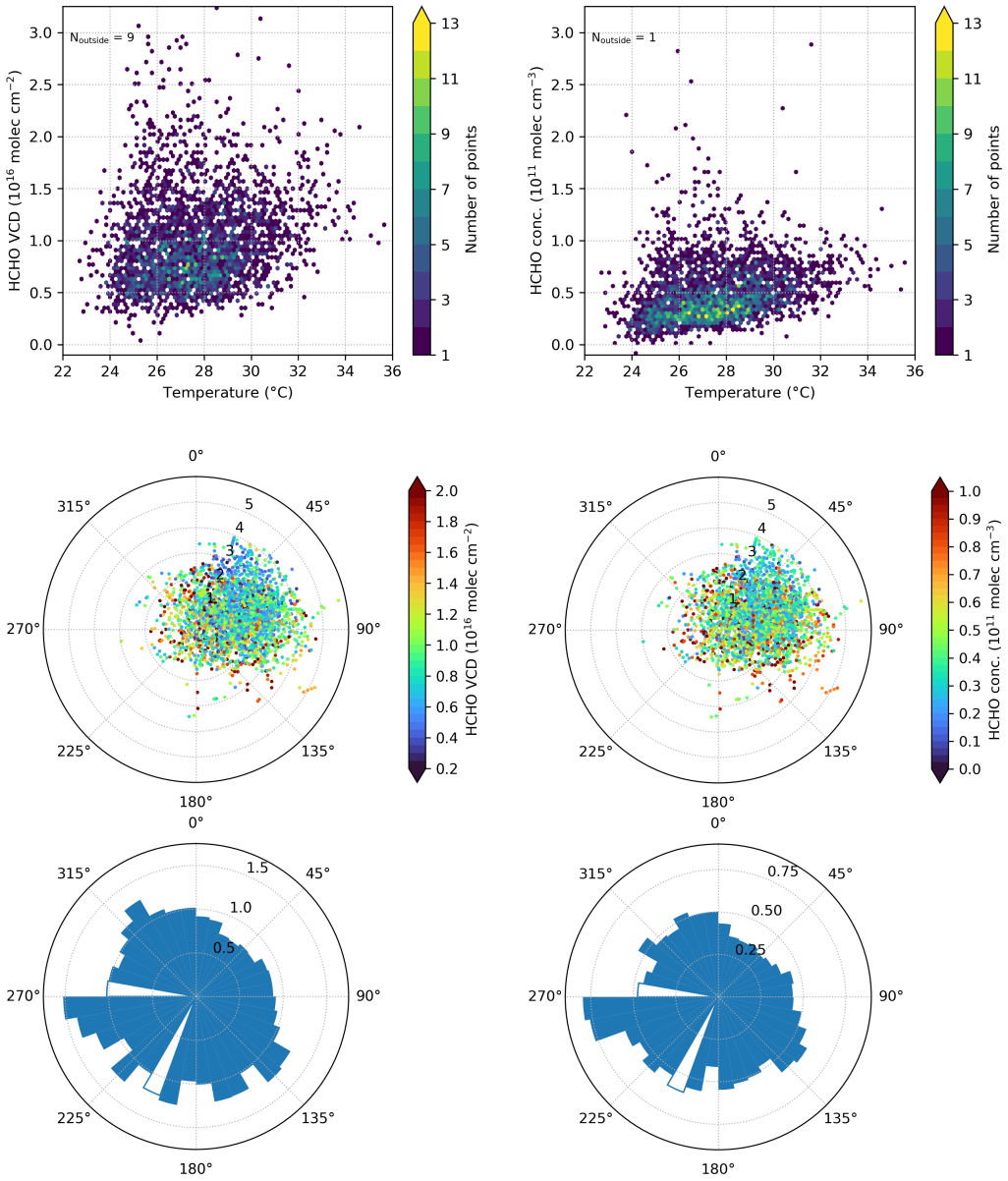


Figure 13.14: The two uppermost panels show hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (80 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HCHO VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

ca. 24 and 28 °C. This finding can be explained by the diurnal cycles of HCHO which show highest abundances in the morning and evening hours, whereas minima are observed around noon. As a consequence, absolutely highest HCHO abundances are observed in the mornings and evenings in the dry season at

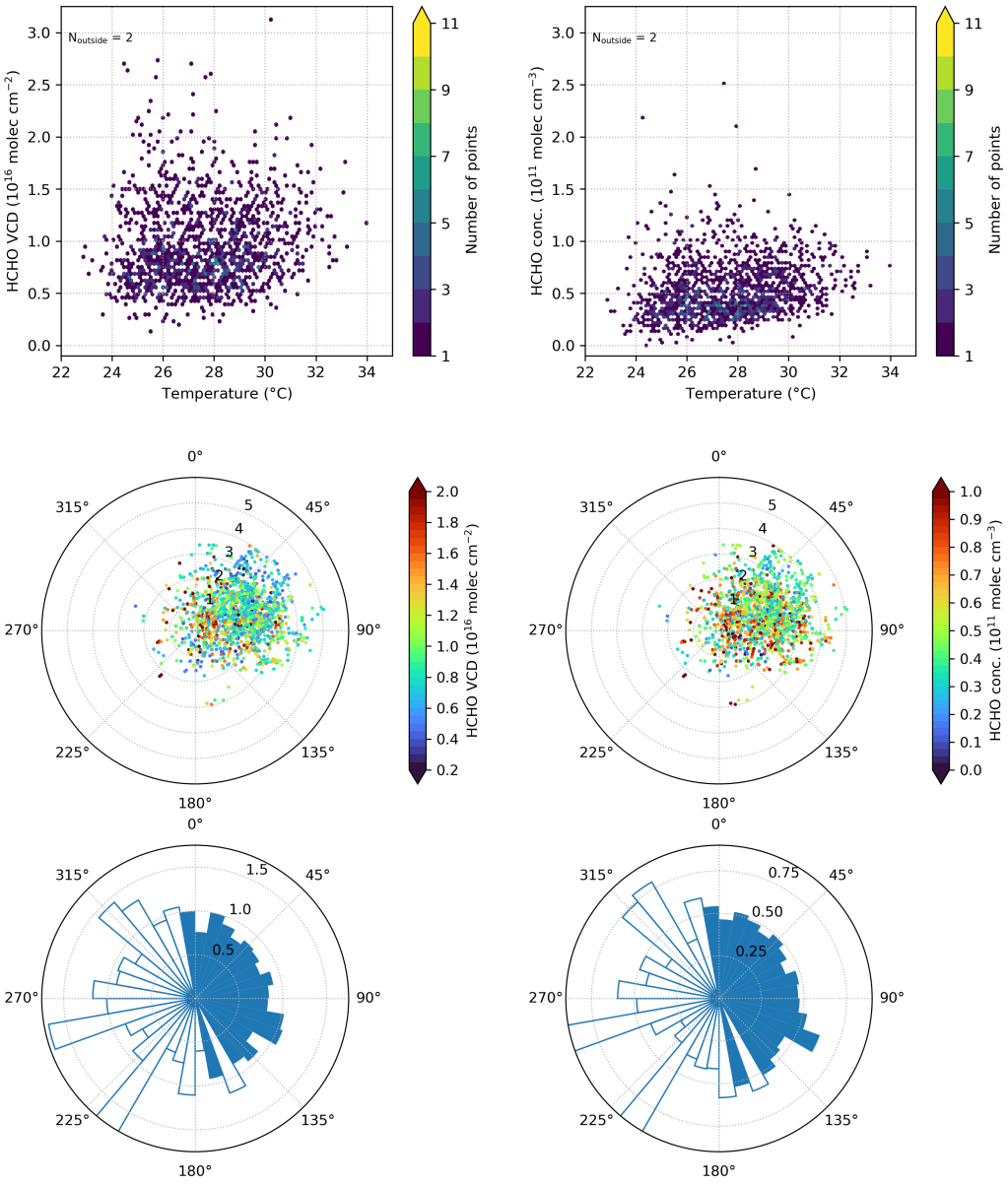


Figure 13.15: The two uppermost panels show hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (298 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median HCHO VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

relatively moderate temperatures. For both instruments, similar temperature dependencies are found, whereby more variability is observed for the 298 m instrument caused by the lower amount of available data. Unlike the other stations, the general temperature dependence at ATTO is almost linear which might be

explained by the different temperature ranges covered at the different stations. At ATTO, the temperature range covered by measurements is rather small ranging from approximately 22 to 36 °C. Also at Bayfordbury and Mainz, rather linear increases of the HCHO abundances are observed at temperatures above ca. 20 °C consistent to the temperature dependence of the emissions of biogenic precursor VOCs, such as for example isoprene or monoterpenes, in this temperature range (compare Fig. 10.9 and 11.9). However, while the highest of those temperatures occur only rarely at these two stations, temperatures below 20 °C are not covered by measurements at ATTO at all. At Mohali, the temperature dependence shows different characteristics indicating a strong effect of anthropogenic emissions there.

The middle panels of Fig. 13.14 and 13.15 depict hourly mean HCHO VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed and direction, while the lower panels summarise the general angular distributions of these hourly means in terms of median values in bins of 10°. Since at ATTO some of these bins contain only little (valid) data, bins with less than 10 data points are displayed as open patches. In contrast to the other stations, for ATTO the wind vectors from the model layer corresponding to an altitude of about 80 m above ground are used, while at all other stations the vectors from a height of ca. 10 m are used. This change was made since the canopy height at ATTO is roughly 40 m and the instrument having the longest data record is located at 80 m above ground. For consistency reasons, the same wind data are used for the 298 m instrument. However, for the temperature correlations still the 10 m temperatures are considered since the temperatures below the canopy probably determine biogenic activity.

For the 80 m instrument, both quantities show very similar angular patterns. The overall highest HCHO abundances occur under south-westerly to westerly winds. However, such wind directions prevail only rarely which is also stressed by the finding that the neighbouring 10°-bins are marked for containing less than 10 data points. In addition, wind speeds are rather low under such wind directions (compare middle panels of Fig. 13.14). As a consequence, both wind directions and medians should be considered with care as the medians in these bins might be biased high since low wind speeds are often associated with elevated trace gas abundances and only little data are available. Similar considerations probably hold for the pronounced maximum in the VCDs under winds from ca. 310 to 330°. Also minor impacts of local effects under such calm wind conditions cannot be ruled out completely. This, however, remains very speculative using the simple approach here. Besides these strong enhancements rather smooth angular patterns are obtained for both quantities indicating that HCHO at ATTO mainly originates from biogenic sources emitted by the forest itself and not from very distinct individual sources. This is also supported by the finding that both VCDs and concentrations at instrument altitude show very similar angular patterns. If only the two major wind directions at ATTO (north-easterly during the wet season and more easterly during the dry season) are considered, the annual variations of the dominating wind direction and the corresponding HCHO abundances in the different seasons are clearly reflected by the angular patterns.

In particular, the median VCDs and concentrations at instrument altitude are only ca. 0.75×10^{16} molec cm^{-2} and 0.35×10^{11} molec cm^{-3}, respectively, under north-easterly winds (wet season), while values of up to about 1.0×10^{16} molec cm^{-2} and 0.60×10^{11} molec cm^{-3} are obtained under easterly to south-easterly winds (dry season). Despite the generally smooth angular patterns under winds from the two major wind directions, local maxima are found for both quantities under winds from ca. 120 to 140° which potentially indicate effects of biomass burning activities although biogenic sources determine the overall variability of HCHO. The effects of biomass burning activities are discussed in more detail below in the paragraph about NO_2.

Also for the 298 m instrument, both quantities show very similar angular patterns. However, given the short measurement period covering only one and a half years and thus only one full annual cycle, no clear angular patterns are found although in principle rather clear annual variations of both quantities can be observed (compare Fig. 13.2). Consistent to the observations for the 80 m instrument, only little data are obtained under winds from westerly wind directions, where all bins are flagged for containing less than 10 data points. Lastly, no local maxima are observed under winds from ca. 120 to 140°.

Figures 13.16 and 13.17 show the dependence of the measured glyoxal VCDs (left) and concentrations at instrument altitude (right) on temperature (uppermost panels) and wind characteristics (middle and lowermost panels) for both ATTO instruments.

The temperature dependencies of the glyoxal abundances are in general similar to the ones of HCHO but show larger variability and differences in the details are found. Nevertheless, also for glyoxal the temperature dependence is clearer for the concentrations at instrument altitude which show an almost linear increase with increasing temperatures. For the VCDs, only a weak tendency of increasing values with increasing temperatures can be observed. Both findings are consistent with the annual cycle of the respective quantity which showed clearly elevated concentrations at instrument altitude in the early dry season (generally warmer), while a less pronounced annual cycle was found for the VCDs due to their large scatter. Further, the decline of the glyoxal abundances starts already in September where temperatures are on average highest (compare Fig. 4.16). As a consequence, the temperature dependencies are slightly less distinct for glyoxal compared to HCHO. Similar to HCHO, the absolutely highest values of both quantities occur at temperatures between about 24 and 28 °C which can again be explained by the diurnal cycles of glyoxal VCDs and concentrations at instrument altitude. Since only data until 21 April 2020 were considered for the 298 m instrument, even more scatter is found for the temperature dependencies. Nevertheless, for the concentrations at instrument altitude the same findings as for the 80 m instrument hold, while only a very weak pattern is found for the VCDs. Also for glyoxal the observed temperature dependencies are different to Mainz and Mohali. Since this is also similar to HCHO, refer to the paragraph above for a detailed explanation.

Under wind directions that prevail only rarely at ATTO, namely winds from the sector spanning (clockwise) from south-south-east to north-north-east, no consistent angular patterns are found for both quantities and both instruments

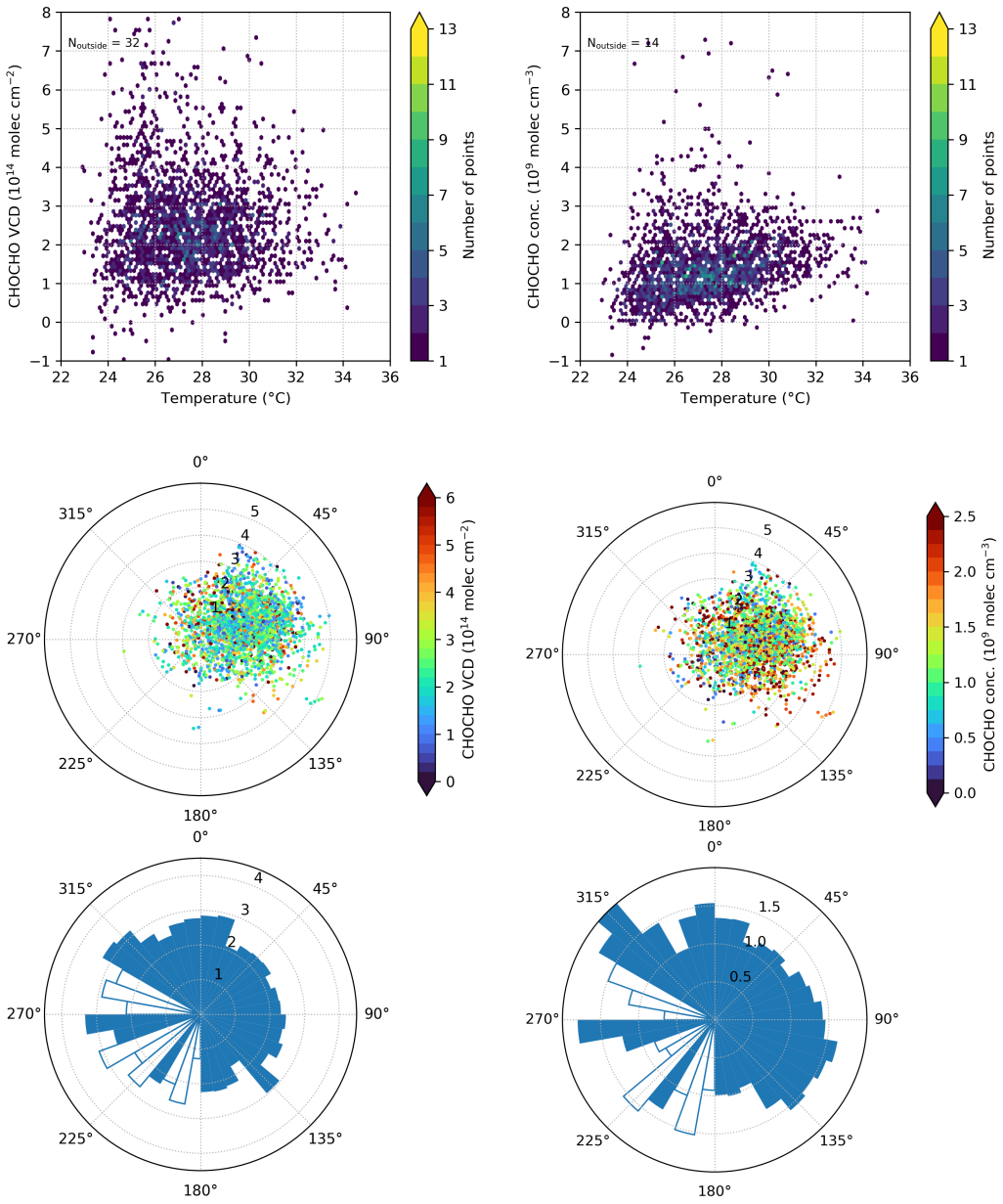


Figure 13.16: The two uppermost panels show hourly mean glyoxal VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (80 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean glyoxal VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median glyoxal VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

(compare lowermost panels in Fig. 13.16 and 13.17). Mainly strong fluctuations of the medians of the different 10°-bins are obtained in this sector and many of these bins are flagged for containing less than 10 data points. For the 298 m

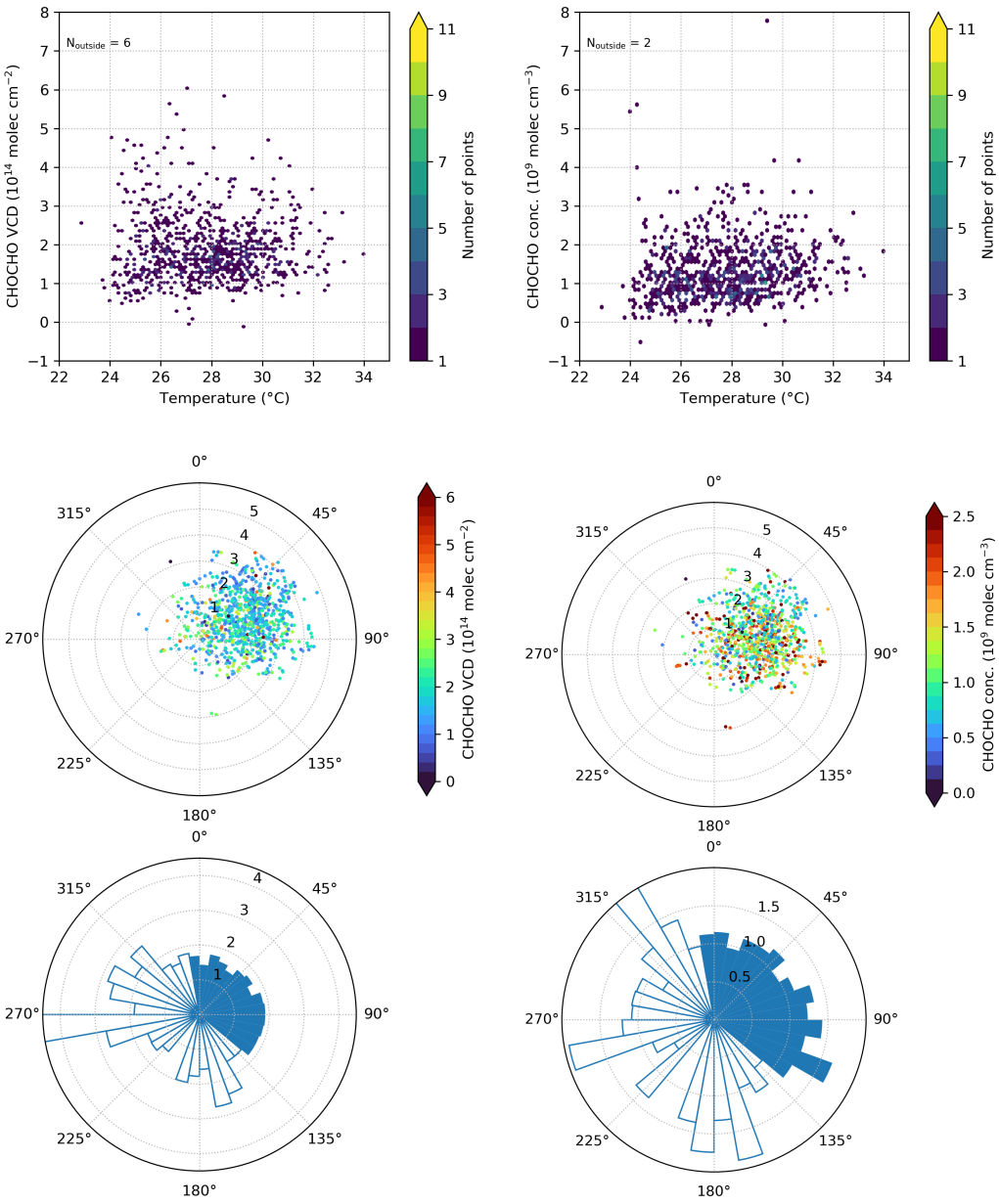


Figure 13.17: The two uppermost panels show hourly mean glyoxal VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (298 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean glyoxal VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median glyoxal VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points. Only data until 21 April 2020 are considered for all panels.

instrument (Fig. 13.17), all bins in this sector are flagged and two bins even contain not a single data point (bins centred around 205° and 215°). Also if only

the two major wind directions are considered, no clear angular patterns can be identified for the 298 m instrument. Only the concentrations at instrument altitude show some enhancements (with a lot of variability) under more easterly and south-easterly wind directions which are the dominant wind directions during the dry season. Similar findings hold for the 80 m instrument (Fig. 13.16). Here, elevated glyoxal concentrations at instrument altitude, slightly exceeding 1.5×10^9 molec cm^{-3}, occur under easterly to south-easterly winds in the dry season, while significantly smaller values are observed under north-easterly winds which dominate in the wet season. No such distinct patterns in the angular distributions of the glyoxal VCDs are found under the two major wind directions. Overall, these findings reflect the annual variations of both quantities which show clear enhancements of the concentrations at instrument altitude in the dry season and only weak annual variations for the VCDs (compare Fig. 13.4). Such differences in the angular distributions of both quantities were not observed for HCHO, where both quantities show very similar angular patterns for the 80 m instrument. As outlined above and assuming that no distinct individual sources of glyoxal exist in the Amazonian rainforest, this suggests that both species have similar precursors since the angular patterns of the glyoxal concentrations at 80 m are comparable to the ones of the HCHO concentrations. However, apparently both species are processed differently in the atmosphere as the glyoxal VCDs of the 80 m instrument and both quantities of the 298 m instrument show no distinct annual patterns in the two major wind directions which suggests that the glyoxal abundances decrease relatively fast with increasing height above ground. Details of the profiles shapes of the different species are discussed in Sect. 13.5.

Next, hourly mean NO_2 VCDs (left) and concentrations at instrument altitude (right) as function of air temperature are presented in the two uppermost panels of Fig. 13.18 and 13.19 for the 80 m and 298 m instrument, respectively.

For both instruments, no distinct temperature dependencies are found for NO_2 VCDs and concentrations at instrument altitude. Nevertheless, the majority of the highest NO_2 abundances for the 80 m instrument apparently occur at temperatures between 24 and 28 °C which is probably again related to the diurnal cycles presented in Fig. 13.9. However, given the large scatter of the NO_2 results this is less clear here. The finding that the NO_2 abundances show almost no systematic temperature dependencies is to some extent surprising since both quantities exhibit systematic annual variations for both instruments (Fig. 13.7 and 13.8) which more or less coincide with the small annual variations of the temperatures at ATTO (compare Fig. 4.16). However, although parts of the negligible correlation can probably be related to the large scatter of the NO_2 results, the low correlation probably also indicates that there is no direct causal relation between temperature and the NO_2 abundance. This interpretation is also supported by the pronounced diurnal cylces which show lower NO_2 abundances around noon and partly in the afternoon when temperatures are highest.

The middle and lower panels of Fig. 13.18 and 13.19 summarise the dependence of the NO_2 VCDs (left) and concentrations at instrument altitude (right) measured by both ATTO instruments on wind speed and direction.

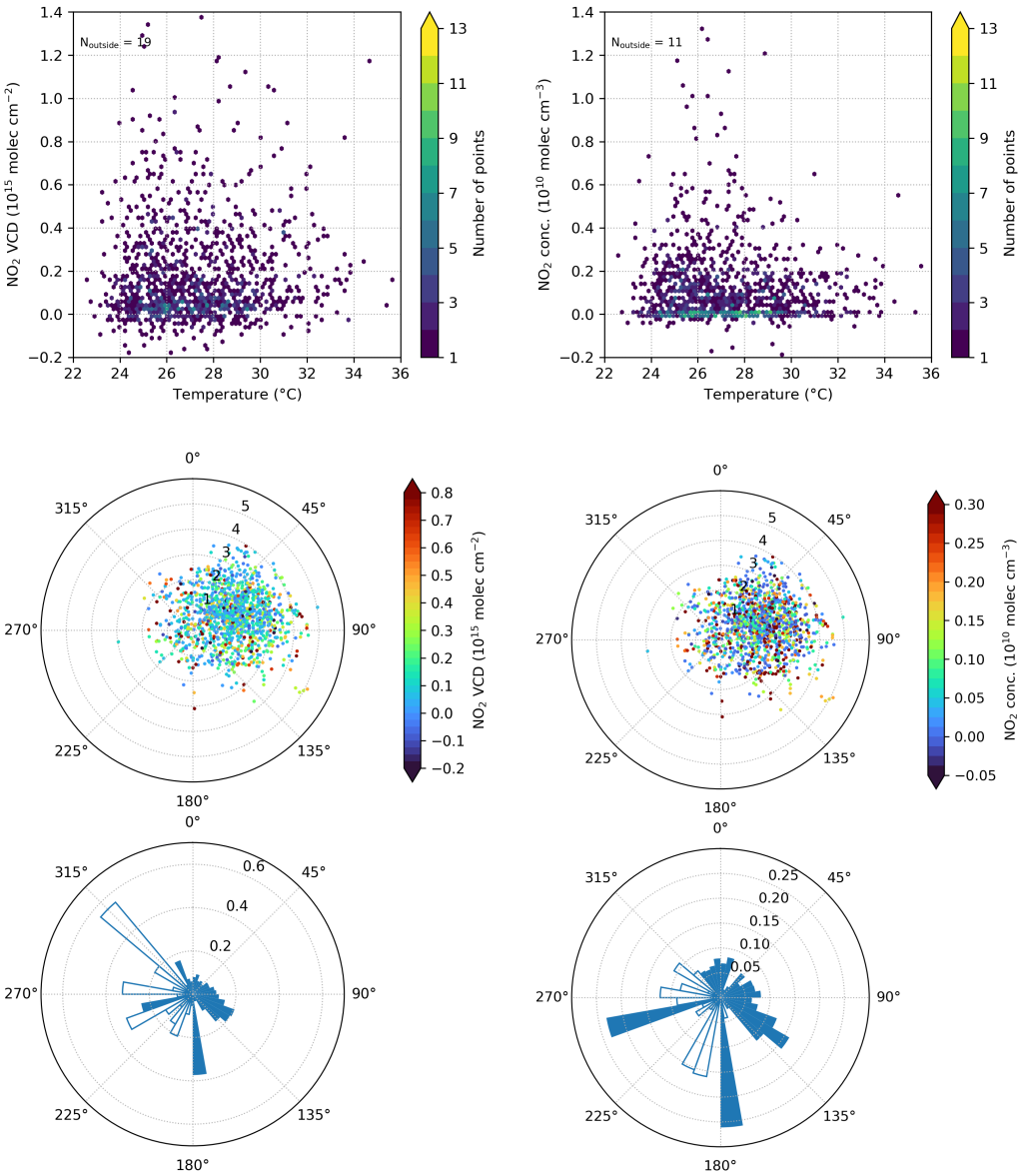


Figure 13.18: The two uppermost panels show hourly mean NO_2 VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (80 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean NO_2 VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median NO_2 VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

Given the low amount of valid NO_2 retrievals (for details refer to Sect. 13.2), rather large variability of the medians in the different 10°-bins is found for both ATTO instruments. This variability is particularly large under winds from directions that occur only rarely (clockwise from south-south-east to north-north-east)

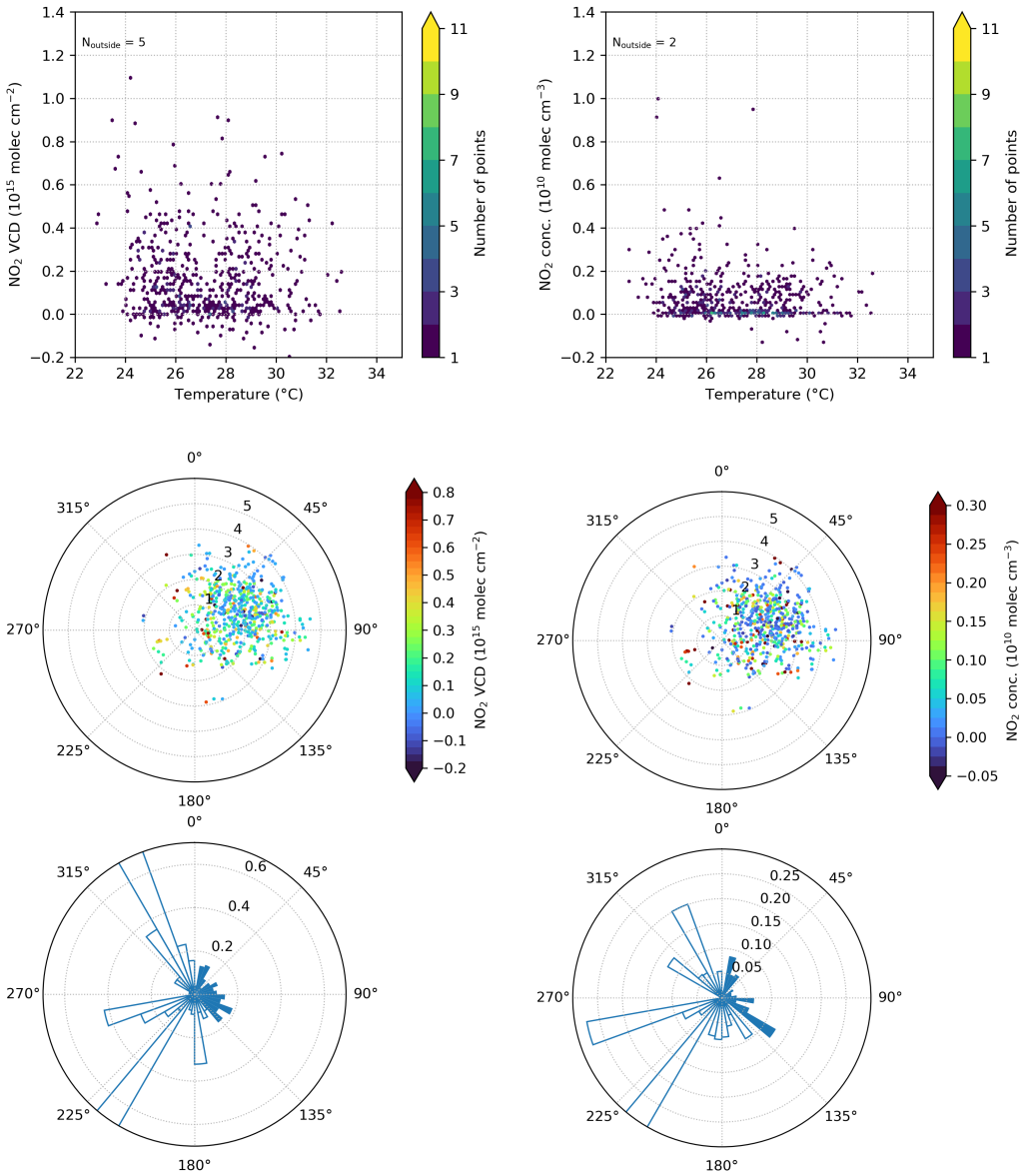


Figure 13.19: The two uppermost panels show hourly mean NO_2 VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (298 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean NO_2 VCDs (left) and concentrations at instrument altitude (right) as functions of wind speed (radius) and direction are displayed. In the bottom panels, median NO_2 VCDs (left) and concentrations at instrument altitude (right) obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

and even larger as compared to the fluctuations observed for glyoxal. Again many of these bins are flagged for their low data amount. For both instruments, rather high median NO_2 VCDs and concentrations at instrument altitude are obtained in this sector for some of the 10°-bins. Since these wind directions are associated

with comparably low wind speeds below $3\,\mathrm{m\,s^{-1}}$ (compare middle panels), small effects of local emissions cannot be excluded despite the remote location of the ATTO site. For instance, the site is equipped with Diesel generators for power supply which are located a few kilometres south-west of the tall ATTO tower. Further, several small communities are situated along the Uatumã river which is ca. 12 km south of the ATTO site and whose stem is from north-west to south-east. Also the closest town, Balbina, is at roughly 60 km distance towards north-west. An effect of emissions from Manaus is very unlikely given the distance of about 150 km towards south-west and the short lifetime of NO_2 in combination with low wind speeds under such wind directions. However, it should be mentioned that the influence of the outlined sources remains speculative since a very simple approach is used here and statistics are bad. This is also stressed by the finding that the angular patterns of both quantities and both instruments in the barely occurring wind sector differ substantially. If only the two major wind directions are considered, the angular distributions of both quantities are overall similar and the annual variations of the NO_2 abundances at ATTO are nicely represented. Higher NO_2 is measured under easterly to south-easterly winds dominating in the dry season, where median VCDs and concentrations at instrument altitude for the 80 m instrument can reach approximately $0.2 \times 10^{15}\,\mathrm{molec\,cm^{-2}}$ and $0.15 \times 10^{10}\,\mathrm{molec\,cm^{-3}}$, respectively, while notably lower values are observed under north-easterly winds which prevail during the wet season. Similar observations are made for the 298 m instrument, however, larger variability and generally lower NO_2 abundances in these wind sectors are found. Compared to HCHO and glyoxal, the differences between the abundances measured under the two main wind directions are (relatively seen) larger and the sector in which the enhancements occur seems to be more confined towards south-east, where the closest major deforestation fires occur. However, also HCHO shows a local maximum under winds from ca. 120 to 140° for the 80 m instrument. All these findings for the angular distributions of NO_2 measured at ATTO stress how pristine the ATTO site is in terms of anthropogenic pollution and that major emission sources are far away from the site as outlined in detail in Sect. 4.4. Nevertheless, also the influence of biomass burning activities on the trace gas abundances becomes evident.

Finally, the influence of meteorological conditions on AODs and aerosol extinctions at instrument altitude at 367 nm measured at ATTO is investigated. Hourly means of both quantities are plotted as function of temperature in the two uppermost panels of Fig. 13.20 and 13.21.

Despite the observed variability, both aerosol quantities show a tendency of generally increasing values with increasing temperatures which is similar to HCHO and basically reflects the annual variations of AODs and extinctions at instrument altitude with enhanced aerosol abundances during the warmer dry season and comparably low aerosol content in the wet season. The few very large extinctions at instrument altitude at temperatures around 25 °C which are observed for the 80 m instrument represent outliers occurring during the wet season (probably related to changing cloud cover) and are also visible in the monthly-box-whisker plots by long whiskers, while lower scatter is found during

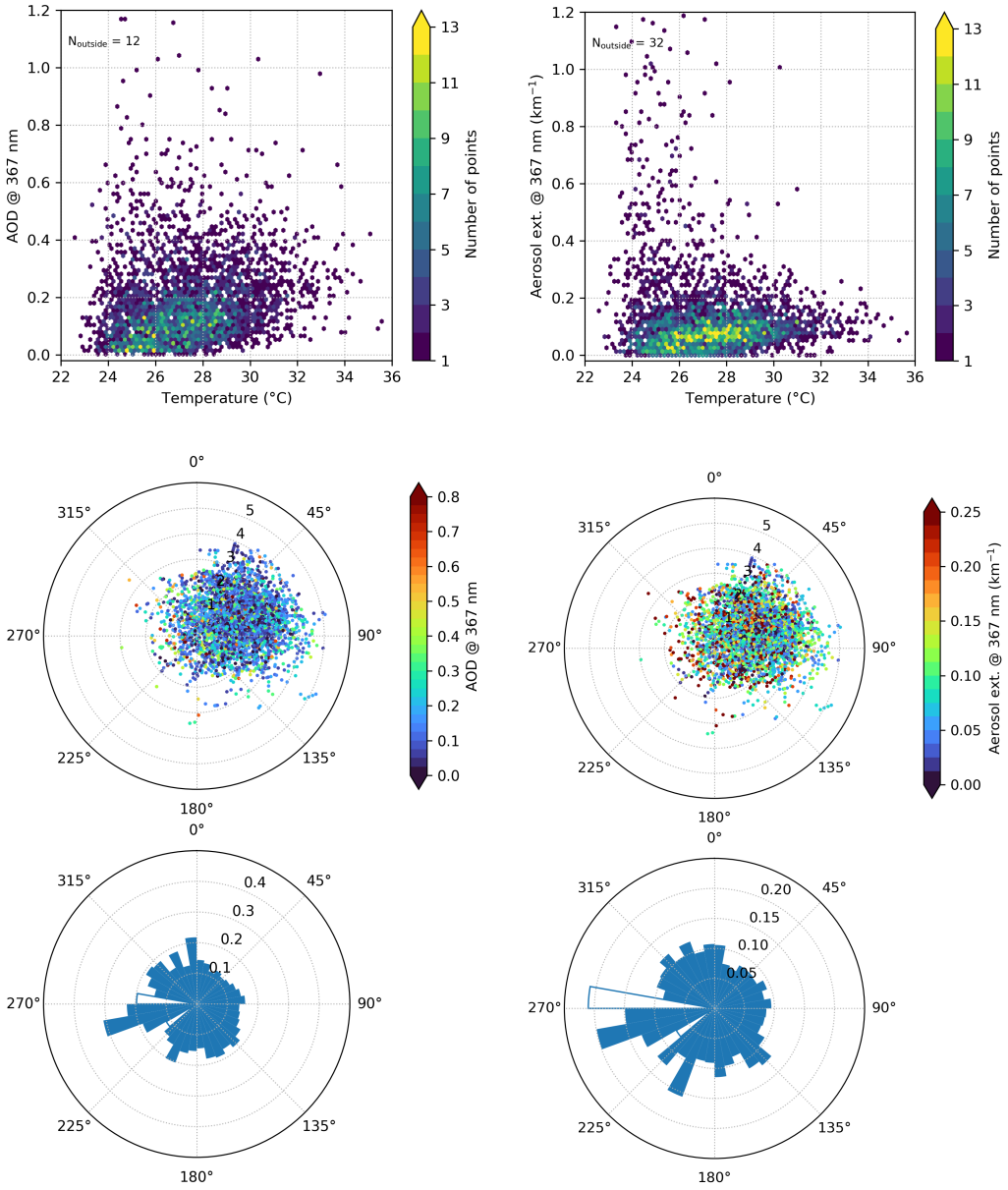


Figure 13.20: The two uppermost panels show hourly mean AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm measured at ATTO (80 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm as functions of wind speed (radius) and direction are displayed. In the bottom panels, median AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

the dry season (compare lower right panel in Fig. 13.10). All these observations hold for both instruments, however, the patterns are less clear for the 298 m instrument due to only little available data.

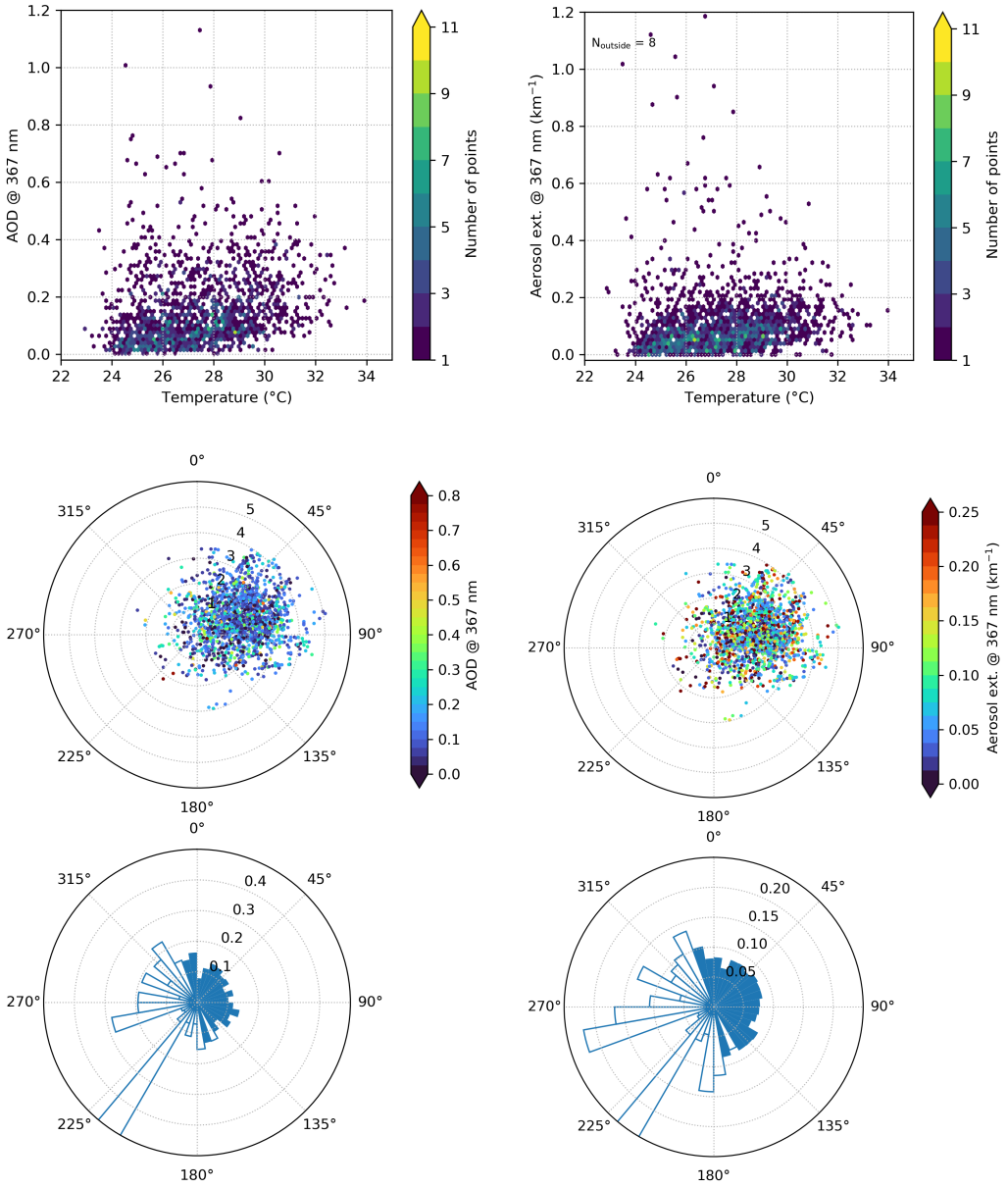


Figure 13.21: The two uppermost panels show hourly mean AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm measured at ATTO (298 m) as function of temperature. Point densities are colour-coded. In the middle panels, colour-coded hourly mean AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm as functions of wind speed (radius) and direction are displayed. In the bottom panels, median AODs (left) and aerosol extinctions at instrument altitude (right) at 367 nm obtained under different wind directions in bins of 10° are indicated by the radius. Open patches indicate bins with less than 10 data points.

The dependencies of AODs (left) and extinctions at instrument altitude (right) at 367 nm measured by the two ATTO instruments on wind characteristics are summarised in the middle and lowermost panels of Fig. 13.20 and 13.21.

Similar to HCHO, the absolutely highest values of both aerosol quantities are observed under westerly winds originating from a sector spanning from ca. 240 to 280°. As already outlined for HCHO, this observation might have two main reasons. First, only little data are available in these 10°-bins, where some are even flagged, and wind speeds are comparably low under such wind directions which are often associated with elevated abundances. In sum, this leads to large uncertainties in the medians of these bins which might be biased high. Second, effects of minor local pollution could contribute to these maxima as outlined in detail for NO_2 which is, however, a speculative hypothesis given the bad statistics in the affected 10°-bins. Similar effects might also cause the other minor local maxima in the angular distributions of both quantities which occur under winds from southerly to north-north-westerly directions. Similar findings are true for the 298 m instrument, however, all bins in this sector are flagged for containing less than 10 data points. In the two major wind directions, rather smooth angular distributions are found. For the AODs, the annual variations are reasonably captured and elevated values of up to approximately 0.18 are observed under easterly to south-easterly winds (dry season), while the median AODs stay below about 0.13 under north-easterly winds (wet season). Significant enhancements of the extinctions at instrument altitude occur only in a rather confined sector spanning from ca. 120 to 150°. Nevertheless, also the AODs and partly the other species (compare respective descriptions), especially for the 80 m instrument, show (local) maxima in these bins, whereby the strengths and the exact angular positions vary amongst the different species. This supports the hypothesis that under south-easterly winds (ca. 120 to 150°) the ATTO site can be affected by emissions originating from deforestation fires and by in general more polluted air masses originating from cities in the north-east of Brazil. Lastly, it should be mentioned again that many measurements with the highest aerosol loads are flagged by MAPA (compare Sect. 13.2) which might lead to a low bias of the reported AODs in this work, especially in the dry season.

For the 298 m instrument, both quantities show similar angular patterns. However, probably caused by the short measurement period and thus insufficient statistics, no clear angular patterns are found in the sectors of the two major wind directions although in principle both quantities show rather distinct annual variations (compare Fig. 13.11).

The dependencies of the water vapour abundances on meteorological parameters are discussed together with their annual cycles in Sect. 13.7.

13.4 TRACE GAS CORRELATIONS

In this section, VCDs/AODs and the corresponding quantities at instrument altitude are correlated amongst each other to further study the relations between the different species. Again, hourly means were used and obvious outliers were excluded from the correlations to reduce the scatter of the data. Further details of how the data were visualised for the correlation studies are outlined in Sect. 10.4.

The correlations of glyoxal and HCHO VCDs (upper parts) and concentrations at instrument altitude (lower parts) are presented in Fig. 13.22 and 13.23 for the

80 m and 298 m instrument, respectively. In both figures, the results of simple linear regressions which serve as orientation are shown in red.

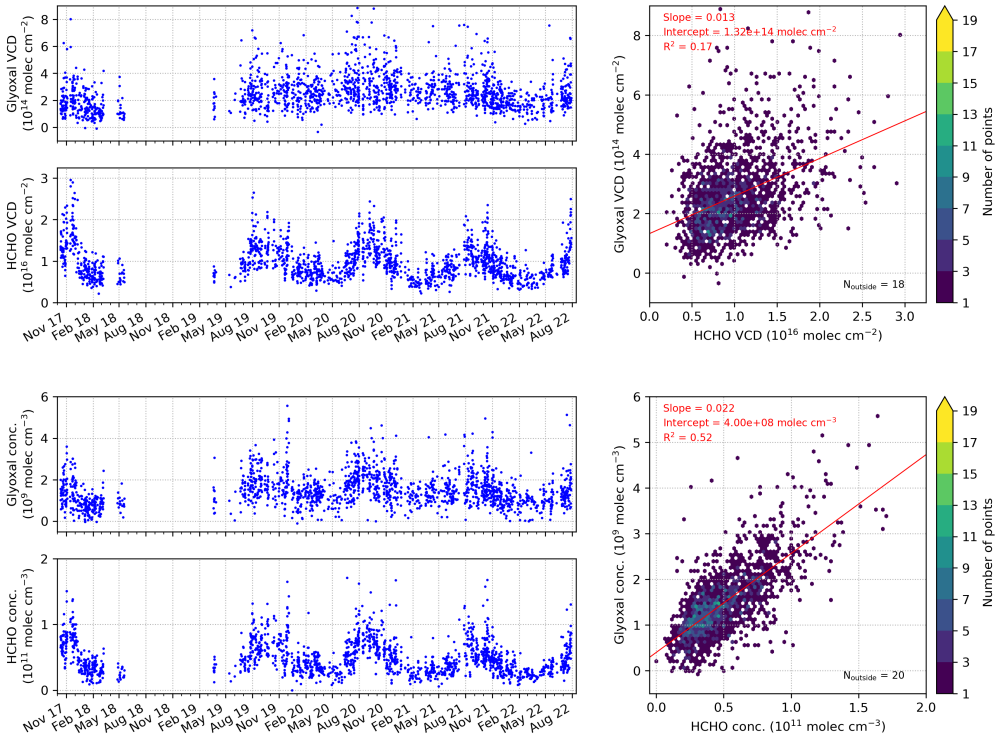


Figure 13.22: Comparison of glyoxal and formaldehyde VCDs (upper part) and concentrations at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The results of a simple linear regression are shown in red, while the number of excluded data points is indicated in the lower right corners.

For both instruments, rather good correlations between the glyoxal and HCHO concentrations at instrument altitude with R^2-values of 0.52 and 0.49 for the 80 m and 298 m instrument, respectively, are found. Contrary to that, the correlations are notably worse for the VCDs for which R^2-values of only 0.17 and 0.26 are obtained. The slopes of the linear regressions are clearly larger for the concentrations at instrument altitude, namely 2.2 % and 1.9 %, compared to the ones of the correlations of the VCDs (1.3 % and 1.0 %). These findings are consistent with the previous observations in this chapter which showed generally a close relation between the concentrations at instrument altitude of both species but larger differences in the behaviours of the corresponding VCDs (e.g. annual cycles). Moreover, this suggests again similar or common sources of glyoxal and HCHO but different processing in the atmosphere leading to different vertical distributions with glyoxal being concentrated closer to the ground, while the HCHO profiles apparently reach higher. This is also supported by the finding that the correlation of the concentrations is slightly better at 80 m. Details of the vertical profiles of the different species are discussed in the next section, whereas the differences between the results obtained with both instruments are examined

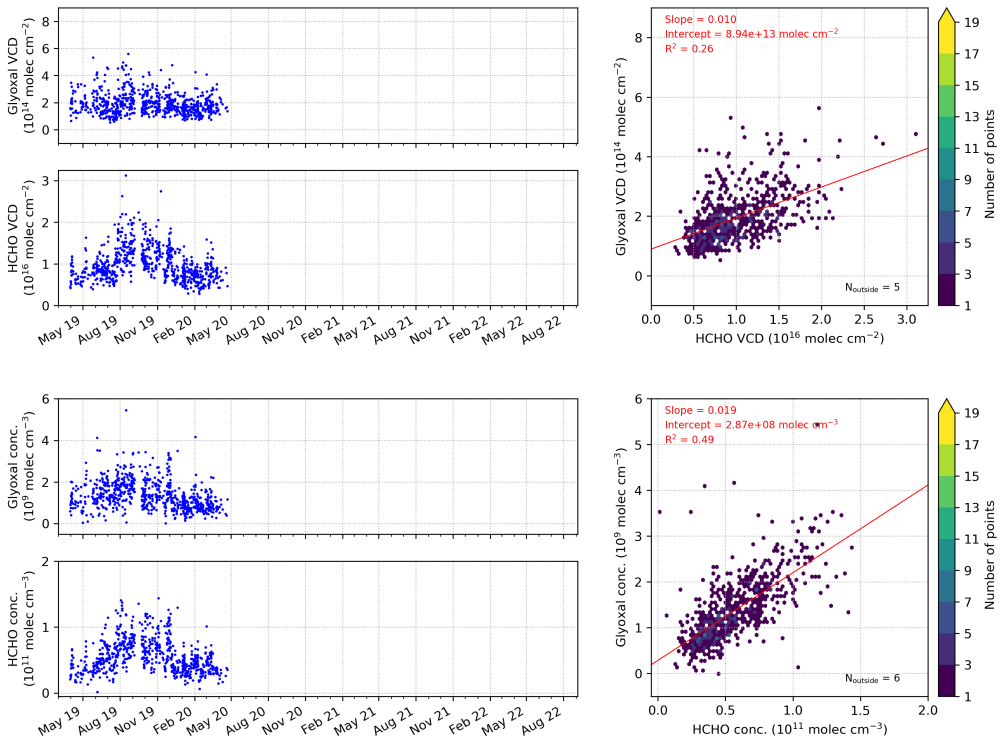


Figure 13.23: Comparison of glyoxal and formaldehyde VCDs (upper part) and concentrations at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The results of a simple linear regression are shown in red, while the number of excluded data points is indicated in the lower right corners. Only data until 21 April 2020 are considered.

in Sect. 13.8. Overall, similar patterns in the correlation plots are obtained for the dry and wet season[3], where mainly the seasonal variations of the abundances of both species are reflected. For instance, glyoxal is (relatively seen) slightly less reduced during the wet season leading to worse correlations in this season. This possibly indicates different chemical regimes in both seasons which is elaborated in more detail at the end of this section when the ratios of glyoxal and HCHO abundances are investigated.

Figures 13.24 and 13.25 summarise the correlations of HCHO and NO_2 VCDs and concentrations at instrument altitude for both ATTO instruments.

Reasonable correlation between both species, with comparable R^2-values ranging from 0.28 to 0.39 , for both quantities and both instruments is found. This finding is consistent to the annual cycles of both species which show generally similar seasonal variations. Nevertheless, highest point densities are found for NO_2 abundances close to 0, again stressing the overall low pollution levels at the remote ATTO site. The majority of these values close to 0 occur at also comparably low HCHO abundances with VCDs and concentrations at instrument altitude

3 No seasonal correlation plots are shown in the main part. For the resulting R^2-values of the correlations on seasonal basis, refer to Table A.2 in the appendix.

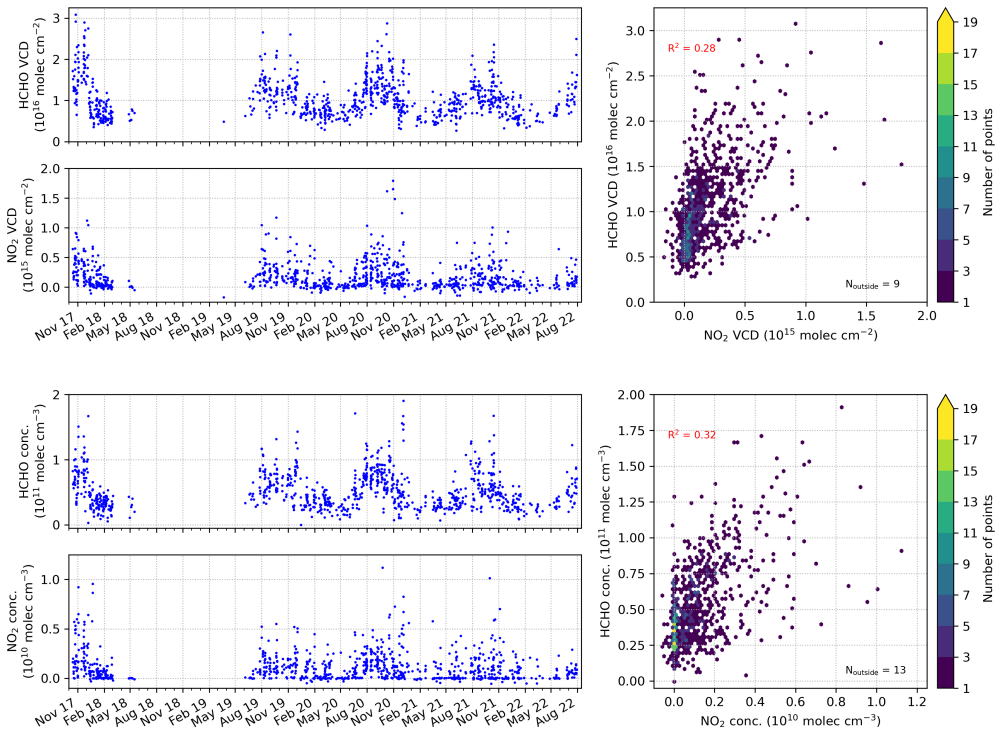


Figure 13.24: Comparison of formaldehyde and NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

below ca. 1×10^{16} molec cm^{-2} and 0.5×10^{11} molec cm^{-3}, respectively, mainly representing measurements during the wet season where the abundances of both species are clearly reduced (compare time series in both figures). Significantly enhanced NO_2 abundances occur only at relatively high HCHO abundances corresponding to dry season measurements in which the site might occasionally be affected by emissions from deforestation fires and in general more polluted air masses originating from cities in the north-east of Brazil as previously described. However, it should be mentioned that especially the concentrations measured at 80 m show small enhancements up to approximately 0.2×10^{10} molec cm^{-3} also during the wet season rather independently of the HCHO concentrations in this season probably indicating small local pollution events which do not largely affect the VCDs nor the measurements of the 298 m instrument. Another interesting aspect is that enhanced HCHO abundances are also observed without strong enhancements in NO_2 stressing again that overall biogenic sources dominate the HCHO levels at ATTO.

The correlations of glyoxal and NO_2 VCDs and concentrations at instrument altitude are shown in Fig. 13.26 and 13.27 for the 80 m and 298 m instrument, respectively.

For the 80 m instrument, the correlation patterns between glyoxal and NO_2 are similar to the ones obtained for the correlations between HCHO and NO_2 for

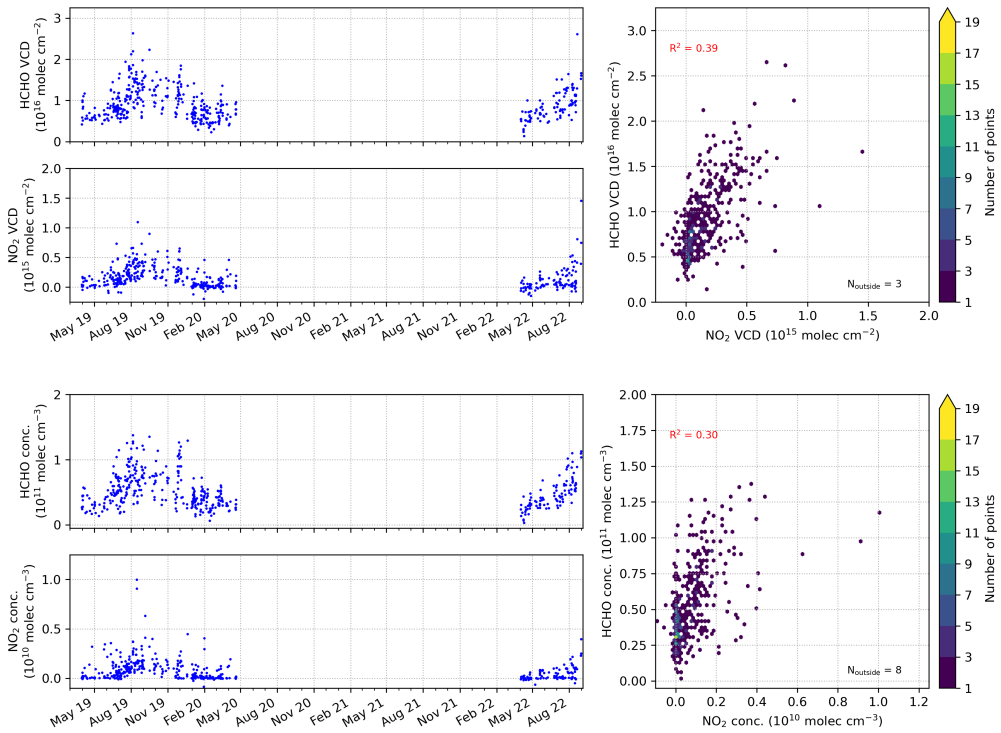


Figure 13.25: Comparison of formaldehyde and NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

both quantities. In particular, high point densities are found for NO_2 abundances close to 0 (wet season measurements), while high NO_2 abundances are mainly associated with rather high glyoxal abundances (dry season measurements). This basically reflects the seasonal variations of both species which are all in all comparable to the ones of HCHO except for details (compare Sect. 13.2) and stresses that elevated NO_2 abundances at ATTO occur mainly during the dry season. Again, these simultaneous enhancements of NO_2 and glyoxal (and HCHO) indicate the influence of biomass burning events on the pollution levels at ATTO. Further, apparently larger variability is observed for the correlations of glyoxal and NO_2 which is also consistent to the observations made above. For the concentrations at instrument altitude obtained with the 298 m instrument, also patterns similar to the ones of the correlation between HCHO and NO_2 concentrations at instrument altitude are found. However, since only one year of data is available for this instrument, the patterns are less distinct. Contrary to that, no consistent pattern emerges from the correlation of glyoxal and NO_2 VCDs measured with the 298 m instrument for which also the annual variations are rather noisy (compare time series). The differences in the correlation patterns of VCDs and concentrations at instrument altitude for the 298 m instrument can probably be explained by comparably shallow glyoxal profiles (compare

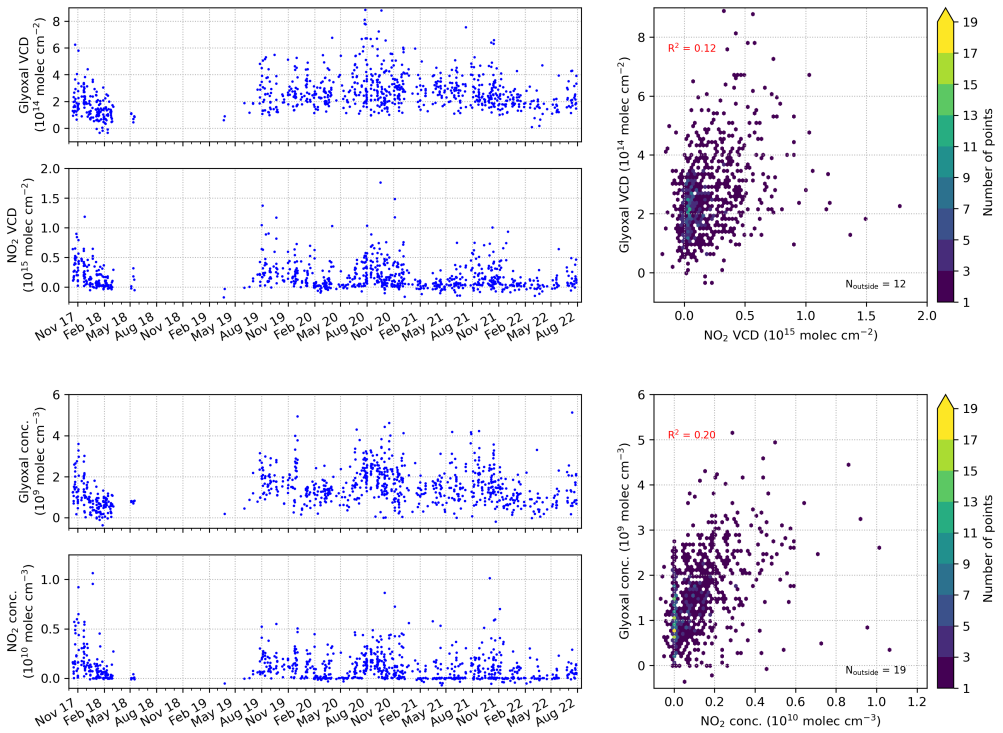


Figure 13.26: Comparison of glyoxal and NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

next section and Sect. 13.8) and the low amount of data available for the 298 m instrument.

Next, correlations between trace gas and aerosol results are examined. Figures 13.28 and 13.29 summarise the correlations of the HCHO and aerosol column quantities (upper parts) as well as the corresponding quantities at instrument altitude (lower parts) for the 80 m and 298 m instrument, respectively.

For both instruments and both quantities, enhanced aerosol loads are generally associated with elevated HCHO abundances which basically reflects the similar annual variations of both species as described in Sect. 13.2. The correlations are slightly better for the column quantities for which R^2-values of 0.23 and 0.29 are obtained for the 80 m and 298 m instrument, respectively. A separation of the data into dry and wet season reflects the seasonal variations of both species and shows that in the dry season similar patterns of the correlation plots are found, while the correlations between both species are significantly worse in the wet season. The latter can be explained by the increased variability of the aerosol results during the wet season (e.g. long whiskers in the lower panels of Fig. 13.10) caused by difficult measurement conditions as outlined previously, e.g. the frequent occurrence of clouds and precipitation. This effect is particularly strong for the aerosol extinctions at instrument altitude. Although some correlation between the HCHO and aerosol abundances is found, no clear conclusion can be made about

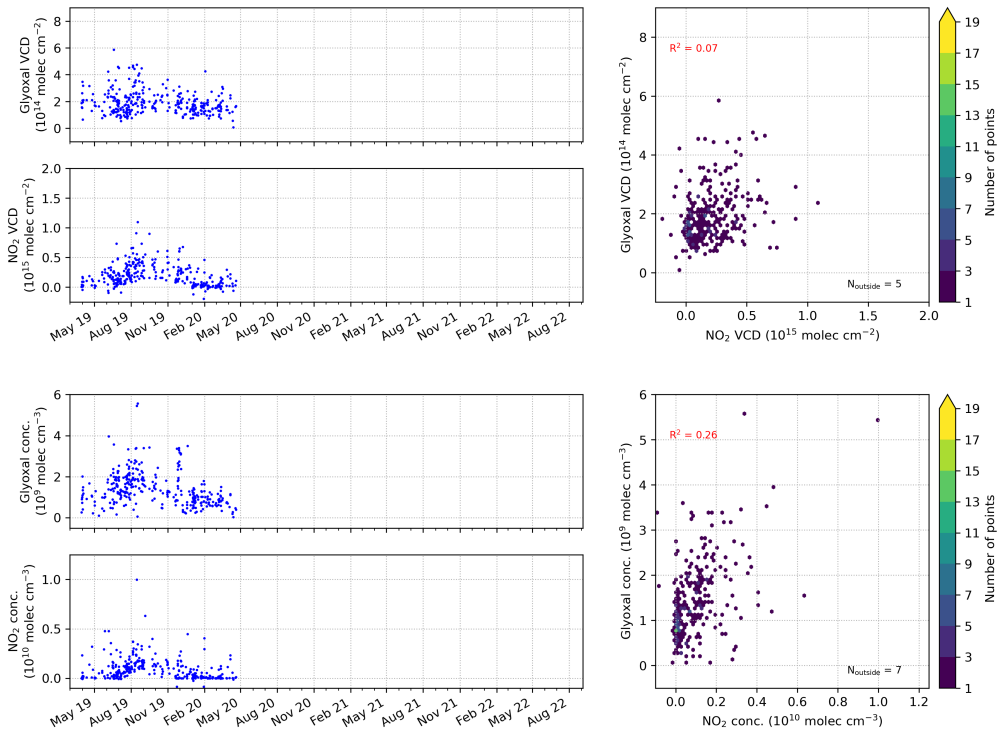


Figure 13.27: Comparison of glyoxal and NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners. Only data until 21 April 2020 are considered.

causal relations. For instance, the formation of SOA or the potential influence of biomass burning events on the HCHO abundances remain uncertain given the simple correlation study used here. Details of the latter were already discussed in Sect. 13.3.

The correlations between glyoxal and aerosol results for the two ATTO instruments are shown in Fig. 13.30 and 13.31.

For both instruments, the quantities at instrument altitude of glyoxal and aerosol correlate reasonably well showing that enhanced aerosol extinctions are typically associated with elevated glyoxal concentrations. However, given the relatively large variability of the retrieval results of both species, the R^2-values are considerably lower compared to the correlations between the HCHO concentrations and aerosol extinctions at instrument altitude. Unlike the observations for the quantities at instrument altitude, no clear patterns emerge from the correlations between glyoxal VCDs and AODs for both instruments. Here, enhanced abundances of either species can occur rather independently of the amount of the other species. Like for the correlations between the HCHO and aerosol results, no changes in the general patterns are found if only the dry season data are considered, while the largely enhanced scatter of the aerosol results leads to no distinct correlation patterns in the wet season also for the quantities at instrument altitude. The differences between the patterns of the correlations for the

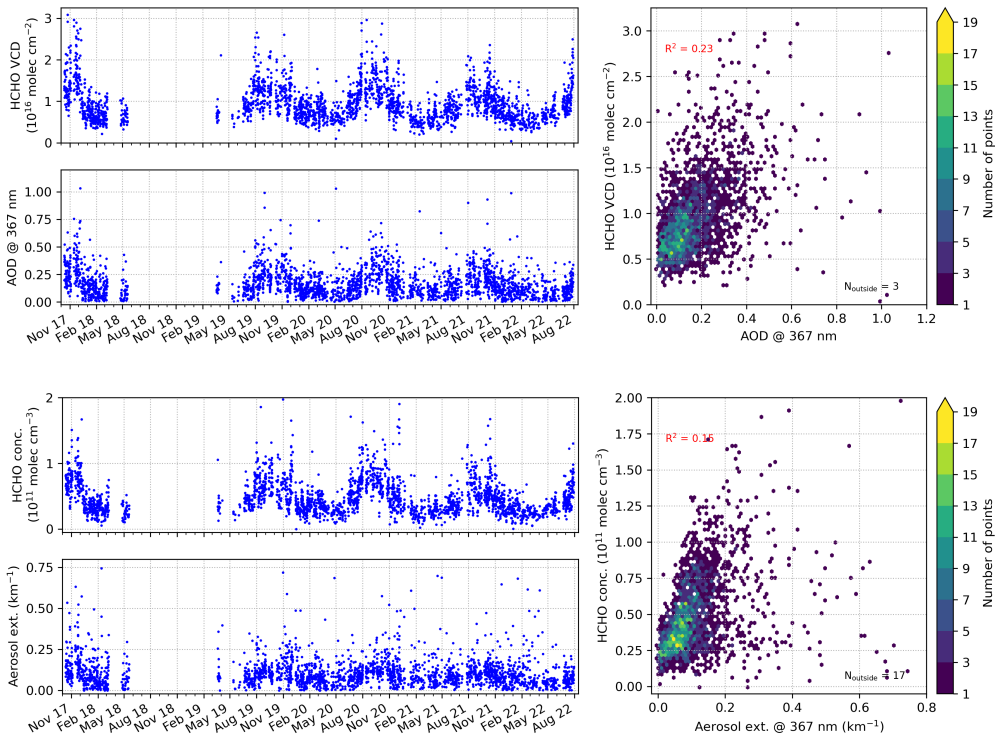


Figure 13.28: Comparison of formaldehyde and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

column quantities and quantities at instrument altitude of glyoxal and aerosol can be explained by the fact that the annual variations of both species are rather similar for the respective quantity at instrument altitude, while the seasonal patterns of glyoxal VCDs (which show large variability) and AODs differ more. This again indicates that glyoxal profiles reach less high compared to the other species (compare Sect. 13.5). Interestingly, for the correlations between HCHO and aerosol results, better correlation was observed for the column quantities. On the one hand, these differences are caused by the profile shapes of glyoxal and HCHO. On the other hand, in addition the HCHO concentrations probably increase within the lowest atmospheric layers. The latter aspect is discussed in more detail in Sect. 13.8.

Figures 13.32 and 13.33 depict the correlations of the NO_2 and aerosol results for the two ATTO instruments.

Unlike the correlations of the HCHO and glyoxal results with the aerosol data, no clear patterns are found for the correlations between the NO_2 and aerosol results for both column quantities and quantities at instrument altitude despite overall similar annual cycles of both species. Similar to the correlations of glyoxal and HCHO with NO_2, high point densities are found for NO_2 VCDs and concentrations at instrument altitude close to 0 which is especially pronounced

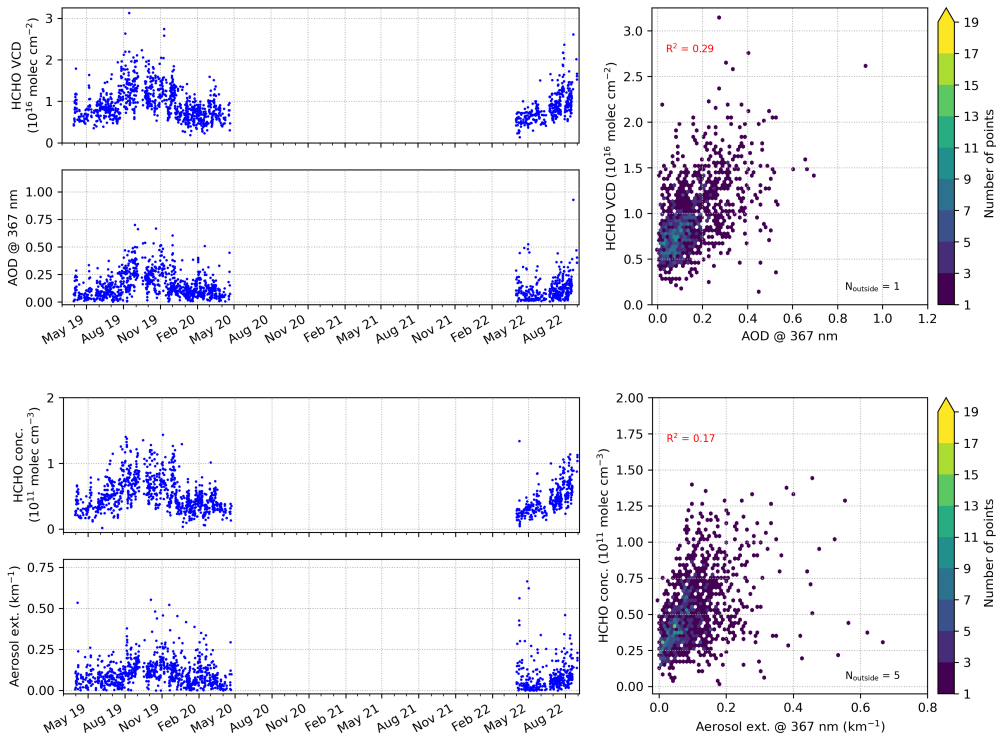


Figure 13.29: Comparison of formaldehyde and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

for the latter. When the whole data are taken into account and ignoring NO_2 abundances close to 0, only a weak tendency that elevated NO_2 abundances mainly occur under enhanced aerosol loads can be identified. This tendency is slightly more distinct if only the dry season data are considered and probably indicates episodes when the ATTO site is affected by emissions from deforestation fires. Contrary to that, no significant correlations between both species exist in the wet season, where the NO_2 abundances are practically 0. Further, the seasonal variations of the abundances of both species are reflected by this separation of the data set. These findings hold for both column quantities and quantities at instrument altitude and for both instruments. The observation that no clear correlations between NO_2 and aerosol are found can be explained by several overlapping effects. For instance, the annual cycles of both species do not agree perfectly and slight differences in the seasonal variations are found (compare for example Fig. 13.7 and 13.10). Also the wind rose plots of the NO_2 and aerosol abundances exhibit differences in the angular distributions of both species. Both findings suggest that NO_2 and aerosol do not necessarily have the same sources all of the times. In particular, NO_2 probably mainly originates from anthropogenic activities, such as deforestation fires, and more densely populated areas in the north-east of Brazil, while aerosol can originate from deforestation fires, secondary

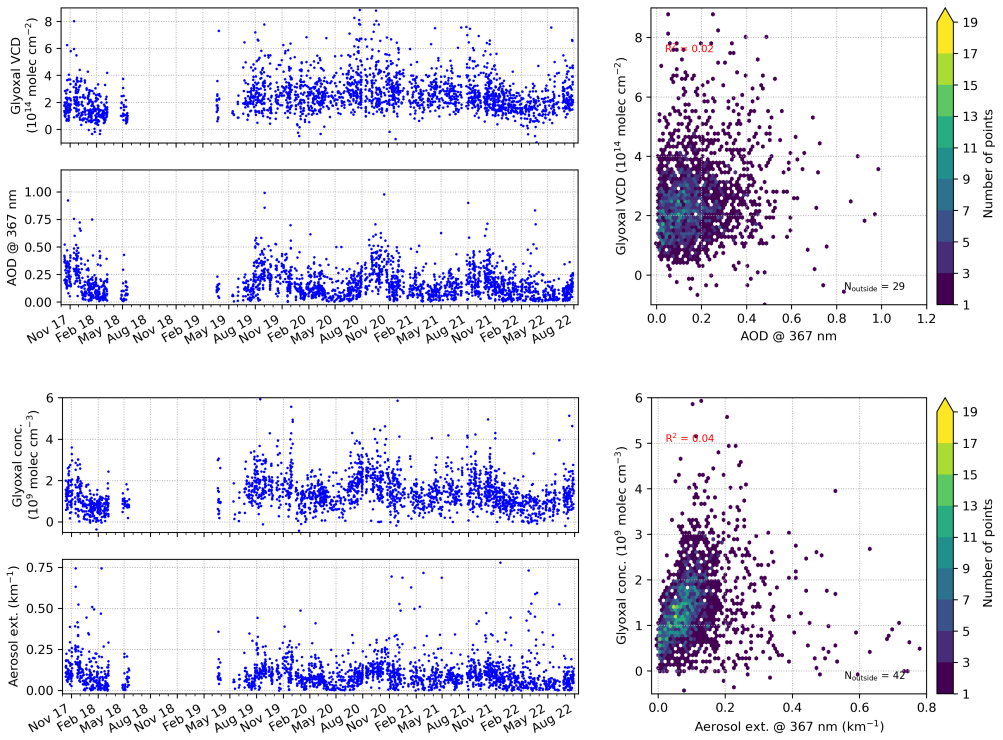


Figure 13.30: Comparison of glyoxal and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

(organic) aerosol formation or even the forest itself (e.g. pollen or fungal spores). While the anthropogenic effects are mainly important during the dry season, the biogenic sources of aerosol are also active in the wet season leading to non-zero aerosol abundances also in this season. Further reasons for the weak correlation between both species might be the variability of the retrieval results of both species which are comparably large as well as NO_2 abundances close to 0 in the wet season and close to the detection limit of the used MAX-DOAS instruments even in the dry season.

Lastly, annual cycles of the ratio between glyoxal and formaldehyde (R_{GF}) are examined to further study typical annual patterns of the relation between the abundances of both species and to potentially infer information on different chemical regimes present at the ATTO site. Details of how the ratios were calculated and visualised are summarised in Sect. 10.4. Like for Bayfordbury and Mohali, also at ATTO only a small fraction, namely roughly 6 % (for both instruments), of the whole data set yields overlapping valid HCHO and glyoxal results.

Average annual cycles of the R_{GF} for the 80 m (upper panels) and 298 m (lower panels) instruments are summarised as monthly box-whisker-plots in Fig. 13.34. Note that for the 298 m instrument again only glyoxal data until 21 April 2020 are considered. Given the particularly low fraction of valid retrievals for both

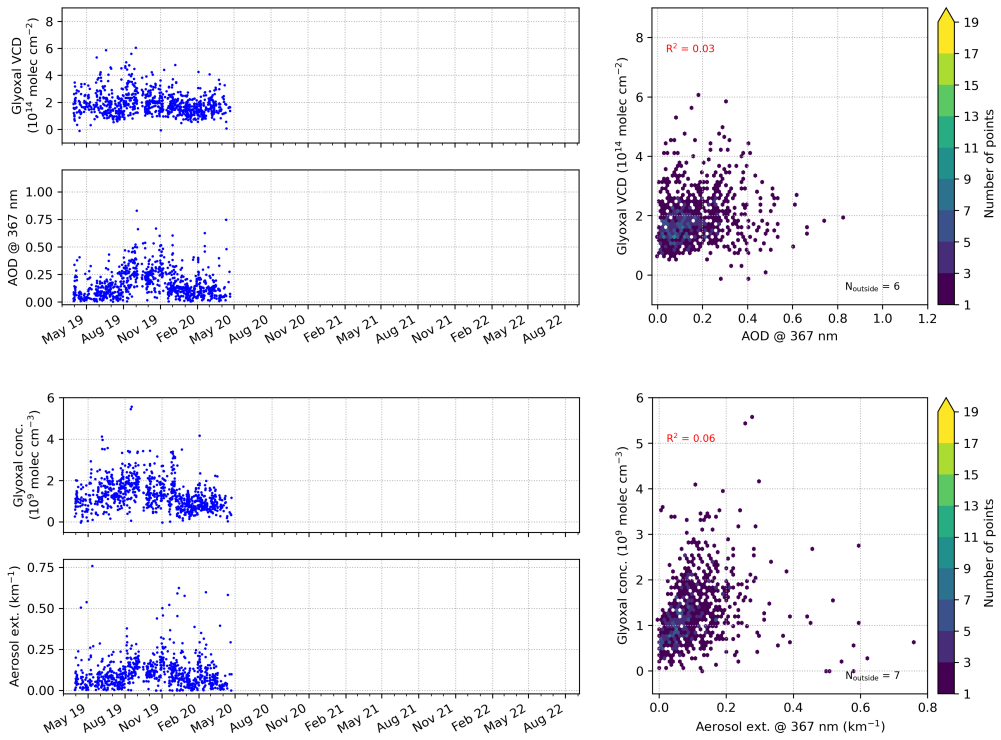


Figure 13.31: Comparison of glyoxal and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners. Only data until 21 April 2020 are considered.

species in the wet season caused by the frequent occurrence of unfavourable measurement conditions, several months are flagged for having less than 5 % available (valid) data and relatively large scatter indicated by long whiskers is found for both instruments. However, also in the dry season, where the overall variability of the results is smaller, the calculation of means is affected by outliers and the annual variations of the monthly means (green dots) are rather noisy. As a consequence, monthly medians (red lines) are considered in the following when absolute values are mentioned.

Systematic annual variations of the R_{GF} calculated from both quantities are found for both ATTO instruments, whereby the relative patterns of these variations are similar for both quantities and both instruments. Highest ratios are found in the wet season while significantly lower R_{GF}-values are found in the dry season months. This reflects the previous finding that the annual variations of the HCHO abundances are (relatively seen) larger than the annual changes of the glyoxal abundances with a stronger relative increase of the HCHO abundances in the dry season. Further, this supports the hypothesis that, despite similar sources, both species are processed differently in the atmosphere and that different chemical regimes prevail in the two main seasons as mentioned above

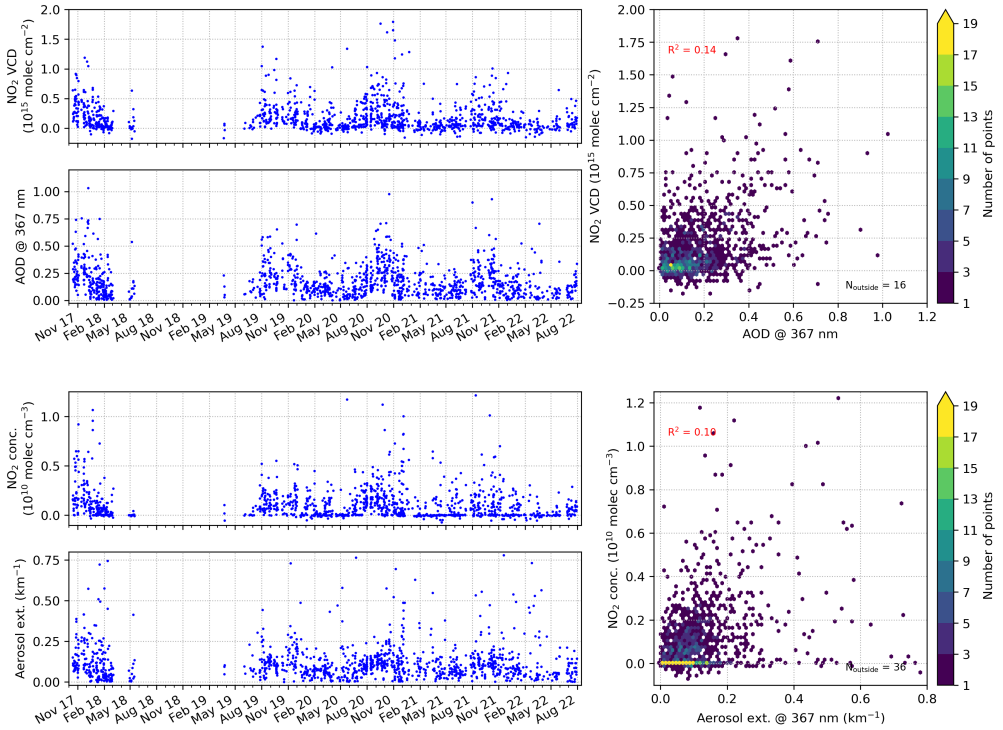


Figure 13.32: Comparison of NO_2 and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 80 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

during the trace gas correlation study. For instance, in the dryer part of the year the plants probably emit higher amounts of precursor VOCs that yield relatively more HCHO than glyoxal compared to the precursors dominating in the wet season. In particular, measurements by Gomes Alves et al. (2023) at the ATTO site showed that isoprene mixing ratios are higher during the dry season (and the transition from dry to wet) than in the wet season. Further, Kaiser et al. (2015) showed that high R_{GF}-values are typically related to a stronger relative contribution of monoterpene emissions, while lower R_{GF}-values are obtained when isoprene dominates the precursors of both species. Thus, the observed seasonal variations of the R_{GF} at ATTO are well in line with the current understanding of tropospheric chemistry.

Despite the generally similar annual variations, differences in the absolute values of the R_{GF} and details of the annual variations can be identified between the R_{GF} calculated from VCDs and concentrations at instrument altitude and both instruments which provide more detailed insights. For instance, especially in the dry season the ratios are higher for the concentrations at instrument altitude compared to the ones obtained from the VCDs for both instruments. Together with the finding that the R_{GF} is also generally slightly larger for the 80 m instrument (in all seasons and for both quantities), this again indicates that the bulk of glyoxal is typically concentrated closer to the ground than HCHO

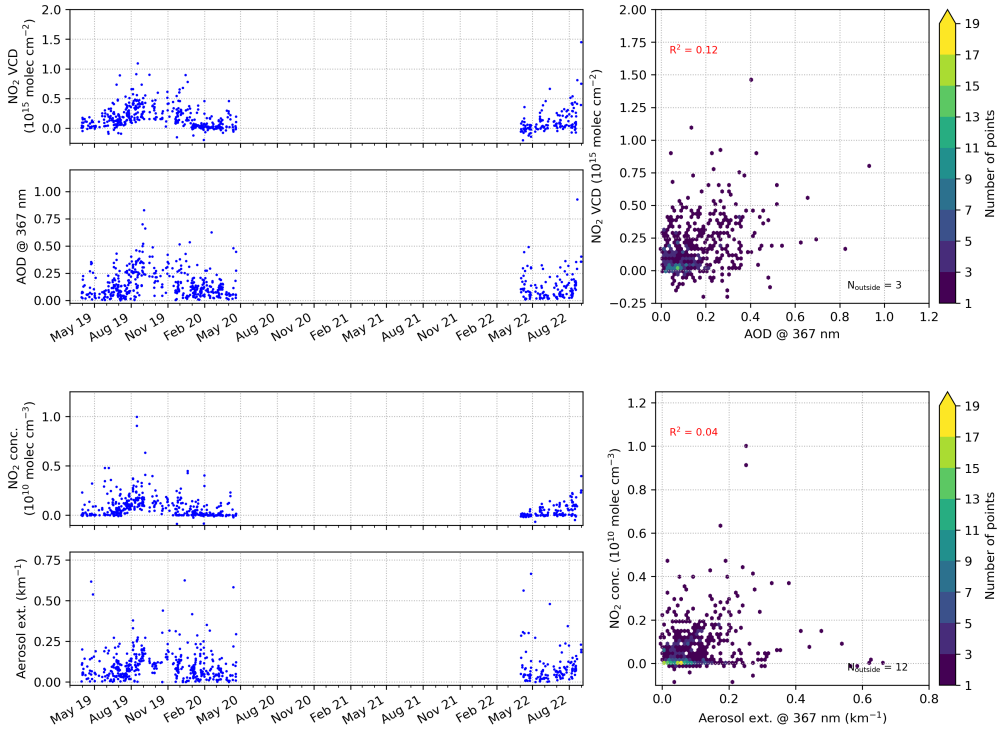


Figure 13.33: Comparison of NO_2 and aerosol results for VCDs and AODs (upper part) as well as concentrations and extinctions at instrument altitude (lower part) for the 298 m instrument. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. Point densities are colour-coded. The number of excluded data points is indicated in the lower right corners.

(compare next section). An overview of representative median R_{GF}-values in the dry and wet season for both instruments is provided in Table 13.2. Kaiser et al. (2015) suggest that under biogenic influence R_{GF}-values above 3 % indicate the dominance (or at least a strong contribution) of monoterpenes over isoprene and vice-versa[4]. Consistent to the considerations above, R_{GF}-values of ca. 3.3 % (dominance of monoterpenes) and 2.7 % (dominance of isoprene) are found for the ratios calculated from the concentrations of glyoxal and HCHO at 80 m in the wet and dry season, respectively. However, more detailed studies also including measurements of other VOCs are needed in the future to further characterise the VOC chemistry at ATTO.

Another aspect is that, especially for the ratios calculated from the concentrations at instrument altitude, the values apparently decrease slightly within the dry season itself (strongest from August to September) potentially indicating small changes of the chemical composition of the precursor substances even during the dry season itself yielding different shares of glyoxal and HCHO. This was already suggested by the individual annual cycles of the concentrations at instrument altitude of both species (compare Sect. 13.2). In particular, the glyoxal concentrations at instrument altitude start to decrease directly after reaching their

4 Actually, the threshold for the dominance of isoprene is stated as 2.5 %.

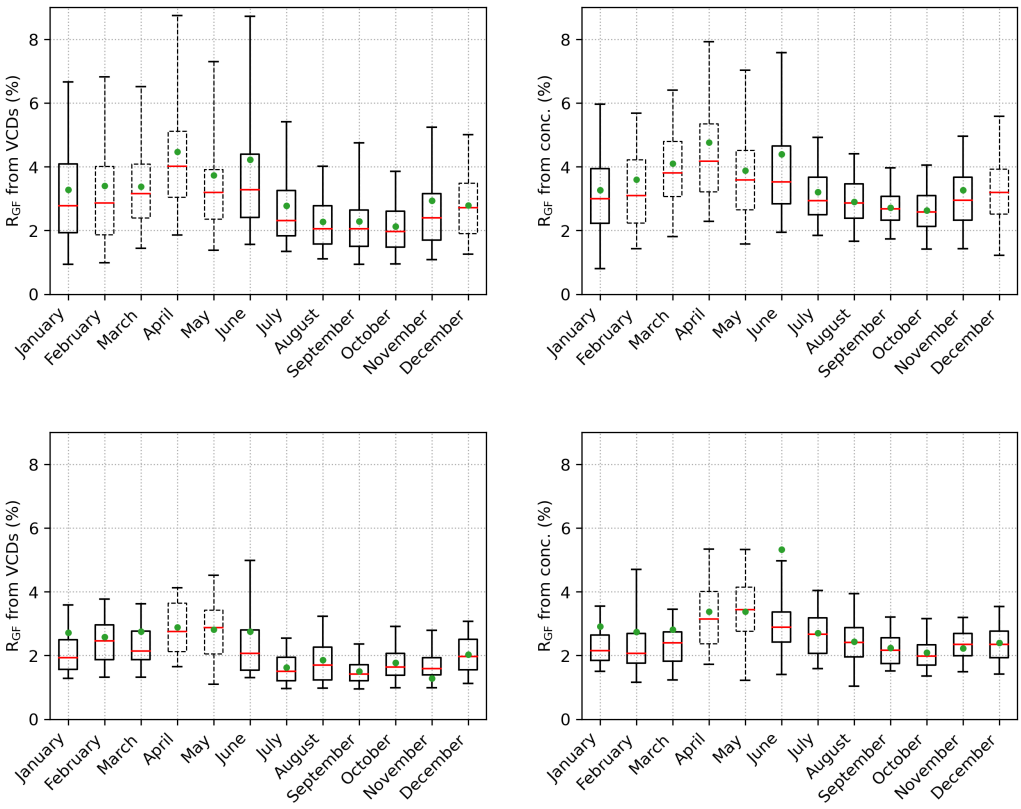


Figure 13.34: Monthly box-whisker-plots of the annual cycles of the ratio R_{GF} obtained from VCDs (left) and concentrations at instrument altitude (right) for the 80 m (upper panels) and 298 m (lower panels) instruments. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Dashed boxes indicate months with less than 5 % remaining data. Only data until 21 April 2020 are considered for the 298 m instrument.

Table 13.2: Overview of representative monthly median R_{GF}-values calculated from VCDs and concentrations at instrument altitude (SFC) in the two main seasons.

INSTRUMENT	WET SEASON		DRY SEASON	
	VCD	SFC	VCD	SFC
80 m	3.1 %	3.3 %	2.1 %	2.7 %
298 m	2.5 %	2.8 %	1.6 %	2.2 %

maximum in August, while the HCHO concentrations at instrument altitude stay at similar levels until October. Only in November, when also the HCHO concentrations at instrument altitude decrease significantly, the R_{GF}-values increase again. A possible explanation for this behaviour might be the additional influence of biomass burning events as suggested by the annual variations of NO_2 and aerosol. It should be mentioned that also in the wet season variations

of the R_{GF} suggest changes in the chemical properties, however, given the large uncertainties there, this should be considered with care and probably needs more measurements to improve the statistics.

Similar to Bayfordbury and Mohali, absolute median R_{GF}-values are determined for the 80 m instrument[5] and values of 2.5 % and 3.0 % are obtained for VCDs and concentrations at instrument altitude, respectively. Compared to the overall median R_{GF} at the other stations, at ATTO the highest ratios between the glyoxal and HCHO abundances are found stressing how different the chemical regime at ATTO is compared to the other two strongly anthropogenically influenced stations.

13.5 VERTICAL PROFILES

To study the vertical distributions of the different species at ATTO, median vertical profiles of HCHO, glyoxal and NO_2 concentrations as well as aerosol extinctions at 367 nm were calculated for each of the four seasons. For the glyoxal profiles of the 298 m instrument, again only data until 21 April 2020 were considered. The resulting profiles for the 80 m and 298 m instruments are presented in Fig. 13.35 and 13.36, respectively. It should be stressed again that MAPA yields vertical profiles relative to the altitude of the instrument. Since at the other stations considered in this work the instruments are placed at maximum 25 m above the ground (at Mainz) which is in the order of the vertical resolution of the default MAPA output grid in the lowest layers (20 m; compare Sect. 3.3), this aspect is not relevant for those. At ATTO, however, the instruments are located at altitudes of 80 m and 298 m above ground and the corresponding height differences become important. Therefore, in this chapter the profiles are displayed in terms of geometric height above ground and also the height information provided below is given relative to the ground. As orientation, the altitudes of the instruments are indicated by dashed black lines in the figures below.

Compared to the other three stations, the formaldehyde profiles reach clearly higher at ATTO. For both instruments, the majority of HCHO is distributed up to about 2 km, whereby the maximum concentrations are found up to a height of approximately 1 km. Also the layer in which the bulk of HCHO is concentrated is significantly thicker at ATTO with a vertical extension of around 700 m for the 80 m instrument. Reflecting the height difference between both instruments, the layer in which the bulk of HCHO is concentrated is vertically less extended (ca. 500 m) for the 298 m instrument as it is "blind" to trace gases located below the instrument altitude. Hence, both instruments yield consistent profile results which reflect the height difference between them reasonably well. Another interesting aspect is that, despite the artefact in the lowest layers, especially in the dry season the concentrations at the instrument altitudes and in the lowest layers above are slightly higher for the 298 m instrument suggesting HCHO formation in the altitude range between both instruments. Further details of the differences between the trace gas results of both instruments are discussed in Sect. 13.8. No

5 The results of the 298 m instrument are likely not representative for the chemical regime close to the ground.

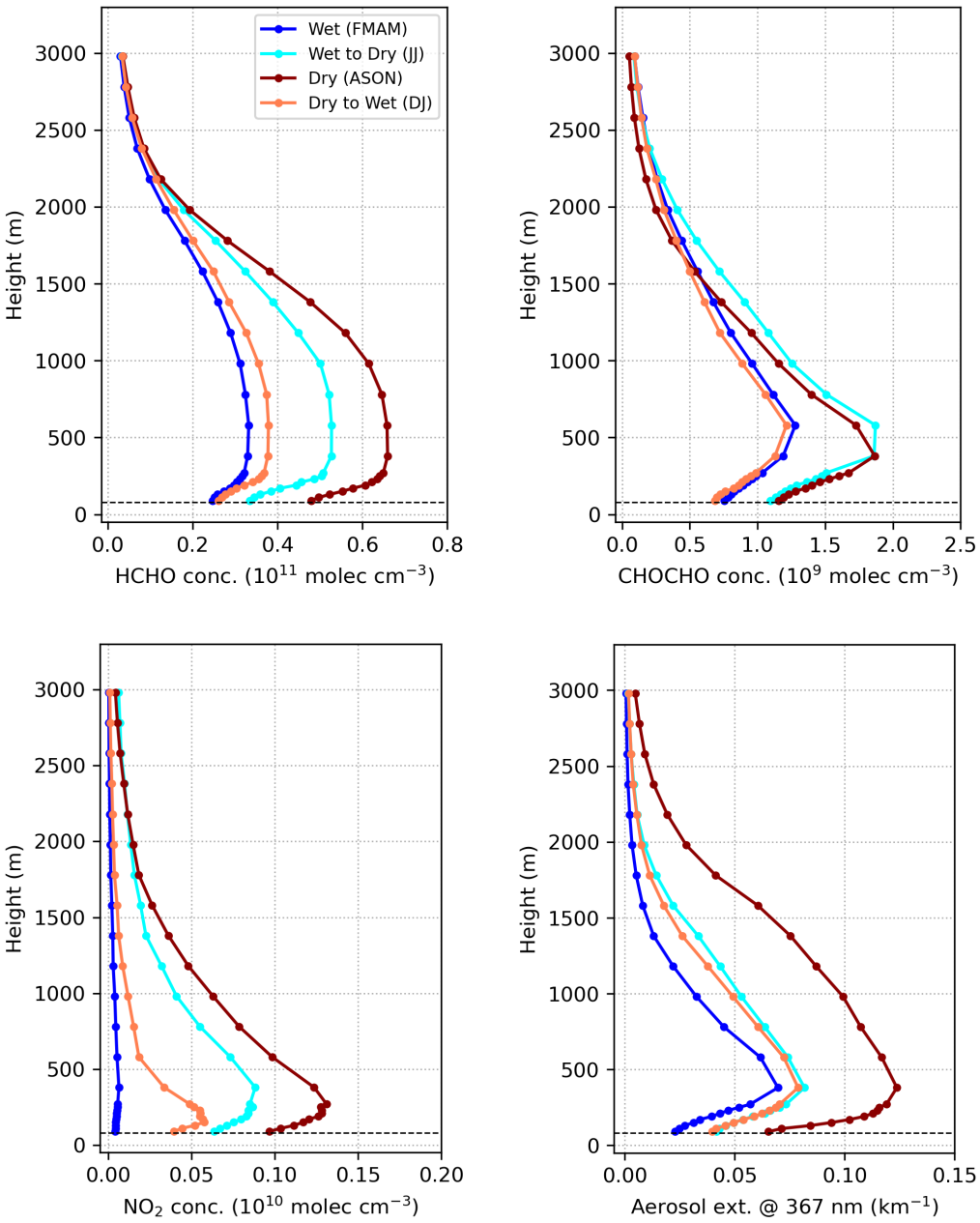


Figure 13.35: Seasonal median profiles of formaldehyde (upper left), glyoxal (upper right) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the ATTO measurements (80 m). Profiles for the different seasons are indicated by different colours. The instrument altitude is marked by dashed black lines.

significant variations of the relative profile shapes between the four seasons can be observed and mainly the seasonal variations of the HCHO abundances are reflected. Nevertheless, for both instruments the decreases of the concentrations after reaching their maxima are apparently less steep in the wet season and in the transition period between dry and wet season potentially indicating longer HCHO

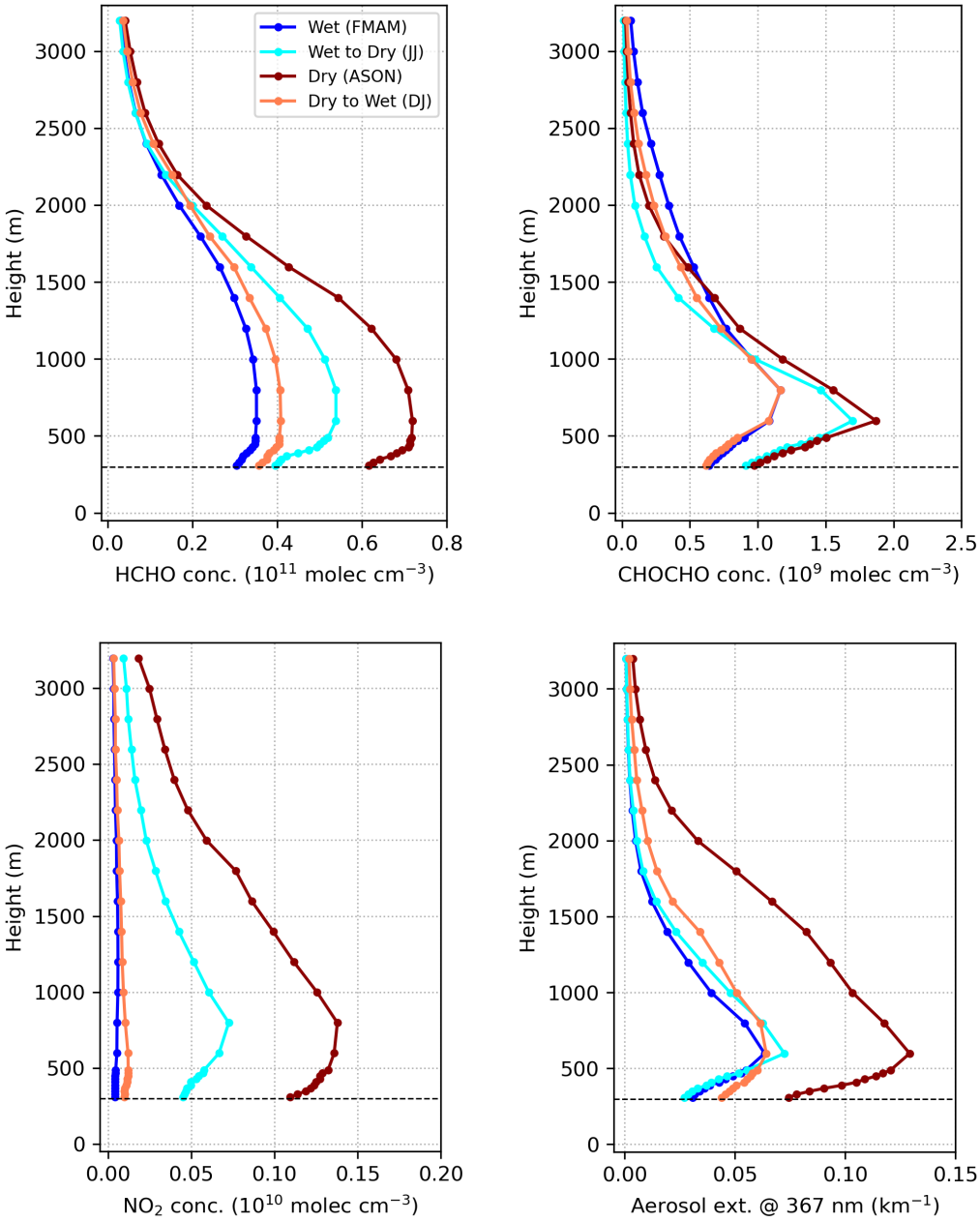


Figure 13.36: Seasonal median profiles of formaldehyde (upper left), glyoxal (upper right) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the ATTO measurements (298 m). Profiles for the different seasons are indicated by different colours. The instrument altitude is marked by dashed black lines. Only data until 21 April 2020 are considered for glyoxal.

lifetimes in these two seasons. The higher reaching profiles at ATTO compared to the other stations can be explained by continuous emissions of HCHO precursors at the instrument location itself and a comparably high reaching boundary layer in the tropics. A last finding is that the concentrations at the altitudes of the

ATTO instruments are comparable or even slightly below the ones observed at Mainz. An effect probably contributing to this is that the ATTO instruments do not yield information on the trace gas abundances below their respective altitudes. However, given the high reaching profiles, the VCDs are overall slightly higher at ATTO.

Also the vertical profiles of glyoxal reach higher at ATTO compared to Bayfordbury and Mohali. However, compared to HCHO the profiles are shallower with the maximum concentrations occurring below about 600 m and the majority of the total glyoxal column being distributed below roughly 1.5 to 1.75 km. Furthermore, the layer containing the bulk of glyoxal is more confined than for HCHO. Similar to HCHO, the profiles obtained for both instruments are overall consistent. However, given the larger variability of the glyoxal results and the rather shallow profiles in combination with the artefact in the lowest layers above the instruments, some discrepancies are found. For instance, the altitude below which the maximum concentrations occur is around 100 to 200 m higher for the 298 m instrument. Nevertheless, consistent to the findings of the previous sections, the layer in which the bulk of glyoxal is concentrated is vertically faintly less extended for the 298 m instrument and the decline of the concentrations after reaching their maxima is steeper, especially in the transition towards the dry season and the dry season itself. The latter two findings probably suggest that the glyoxal concentrations decrease between the heights of the two instruments. For details, refer to Sect. 13.8. No strong variations of the relative profile shapes between the different seasons are found. Nonetheless, in the transition towards the wet season and in the wet season itself the maximum glyoxal concentrations occur at slightly higher altitudes and, similar to HCHO, the decreases of the concentrations after reaching their maxima are apparently less steep indicating longer glyoxal lifetimes in these two seasons. As already mentioned, the finding that the HCHO profiles extend over a rather large vertical range while glyoxal is concentrated closer to the ground is consistent to earlier findings in this chapter and suggests different chemical processing of both species despite similar sources/precursors. For instance, larger R_{GF}-values and better correlations between HCHO and glyoxal were found for the concentrations at instrument altitude. Also at Bayfordbury and Mohali, the glyoxal profiles are apparently shallower than the HCHO profiles which seems to be a general finding.

Like for the previous three stations, the NO_2 profiles measured with the 80 m instrument are the shallowest of all species. The maximum concentrations are found below approximately 400 m in the transition towards the dry season and in the dry season itself, while the maximum concentrations reach only up to ca. 250 m in the transition between dry and wet season. Similar to glyoxal, the majority of NO_2 is distributed up to about 1.5 to 1.75 km (1 km in dry to wet). No enhanced NO_2 abundances are measured in the wet season at all (compare Fig. 13.7), while similar relative profile shapes are found in the dry season and in the transition towards the dry season. In addition to shallower profiles, in the transition towards the wet season also the layer containing the bulk of NO_2 is vertically less extended. These slightly different profile shapes are caused by at least two effects. First, temperatures are typically higher in the months from June

to November (compare Fig. 4.16), corresponding to the transition period towards the dry season and the dry season itself, than in the rest of the year. Hence, the boundary layer reaches higher and vertical mixing is enhanced in these seasons. Interestingly, for HCHO and glyoxal the effect of longer lifetimes is apparently stronger than the less effective mixing in the wet season and transition towards the wet season. Second, in the seasons in which higher reaching profiles are obtained, the measured pollution originates mostly from sources located a few hundred of kilometres away from the ATTO site. Thus, mostly long-range-transported NO_2 is measured which is typically not concentrated close to the ground like for example observed at Mohali where mostly freshly emitted pollution is measured (compare Fig. 12.21). Despite the observation that NO_2 shows the shallowest profiles amongst all species also at ATTO, the profiles reach higher compared to the other three more polluted stations. Also this difference can be explained by the reasons outlined before. For the 298 m instrument, overall similar results are obtained, i.e. only very low NO_2 concentrations are found in the wet season and the transition season towards the wetter part of the year. The difference between the findings of both instruments in the transition period towards the wet season might be partly explained with the height difference between both instruments as the NO_2 profiles are rather shallow as described above. Contrary to that, the profiles reach apparently higher in the transition towards the dry season and in the dry season itself with the maximum concentrations occurring up to almost 1 km and the majority of the total NO_2 being distributed up to ca. 2 km. No clear explanation for this finding can be given so far. One aspect is probably that only one dry season is covered by the 298 m instrument and consequently statistics are rather bad. However, also characteristics specific for the dry season 2020 are imaginable which get lost in the long-term medians used for the 80 m instrument. Lastly, also different sensitivities for different layers of (transported) NO_2 for both instruments might contribute. Details of this have to be discussed when more measurements for the 298 m instrument are available.

For both ATTO instruments, comparable profiles of aerosol extinction at 367 nm are obtained and in all seasons the highest extinctions are found below a height of around 600 to 750 m. While the majority of the aerosol content is distributed up to a height of roughly 2 km in the dry season, the profiles reach less high during the rest of the year, where the majority of aerosol is located below around 1.5 to 1.75 km. Overall, consistent results are found for both instruments. However, similar to the findings for glyoxal the altitude below which the maximum extinctions occur is around 100 to 200 m higher for the 298 m instrument which is likely related to the artefact in the lowest layers above the instruments. Nevertheless, the annual variations of the aerosol load at ATTO are nicely reflected by the seasonal median profiles for both instruments, with highest aerosol abundances in the dry season and minimum aerosol abundances during the wet season. Values in between, but closer to the ones in the wet season, are found for the two transition periods. Also the layer in which the bulk of aerosol is concentrated is notably thicker in the dry season especially for the 80 m instrument (probably related to the height difference between both instruments). The vertically more extended aerosol extinction profiles in the dry season can be explained in a similar way

as outlined above for NO_2, namely a combination of sources located rather far away from the ATTO site and enhanced mixing in the warmer part of the year. However, unlike NO_2 the profiles for the transition period towards the dry season are not affected for the aerosol extinctions. This finding might be explained by the timing between the change of the dominant wind direction and the occurrence of deforestation fires as already outlined in Sect. 13.2. Similar to the previous species, also the profiles of aerosol extinction at 367 nm reach higher at ATTO compared to the other stations most likely caused by a higher boundary layer and enhanced mixing in the tropics. Further, also the formation of SOA might contribute.

As already observed at the previously discussed stations, also for the ATTO measurements consistently higher concentrations at instrument altitude are obtained for all trace gases using the fixed O_4 scaling factor of 0.8, while mainly enhanced scatter is observed for the VCDs. For the aerosol retrieval, both quantities show higher values for the fixed SF compared to the variable one. As explained in detail for Bayfordbury in Sect. 10.5, these findings can be explained by slight differences in the vertical distributions of the total trace gas column using either SF, while the entire profiles are shifted towards higher extinctions for aerosol. Despite these differences, also at ATTO the main results are not affected by the choice of the SF. Only the diurnal cycles of both aerosol quantities were significantly different for both SF scenarios as outlined in Sect. 13.2, whereas the annual cycle is almost unchanged.

The seasonal median profiles of water vapour concentrations are presented in Sect. 13.7.

13.6 COMPARISON TO TROPOMI

This section provides a brief comparison between formaldehyde VCDs measured by the two ATTO MAX-DOAS instruments and the TROPOMI satellite instrument. Since TROPOMI HCHO VCDs are only available from May 2018 onwards (compare Sect. 5.3) and a large data gap between mid of May 2018 and mid of March 2019 exists for the 80 m instrument (compare Fig. 13.1), overlapping data are available for around three and a half years for this instrument. For the 298 m instrument, the overlap is limited by the performance of the MAX-DOAS instrument and overlapping data exist for only one and a half years. Details of how the data were selected and handled for the comparison are described in Sect. 10.6. Comparisons on daily (upper parts) and monthly (lower parts) basis are shown in Fig. 13.37 and 13.38 for the 80 m and 298 m instrument, respectively.

Also for the two ATTO instruments reasonable correlations between both data sets are found. Although comparable R^2-values of 0.27 and 0.24 are obtained for the comparisons on a daily basis for the 80 m and 298 m instrument, respectively, the slopes of the linear regressions differ more between both instruments with values of 0.64 (80 m) and 0.49 (298 m). This finding might be explained by the comparably low data amount (only 65 data points) available for the 298 m instrument, together with a rather large variability of the data (compare upper correlation plot in Fig. 13.38). For both instruments, no distinct systematic over-

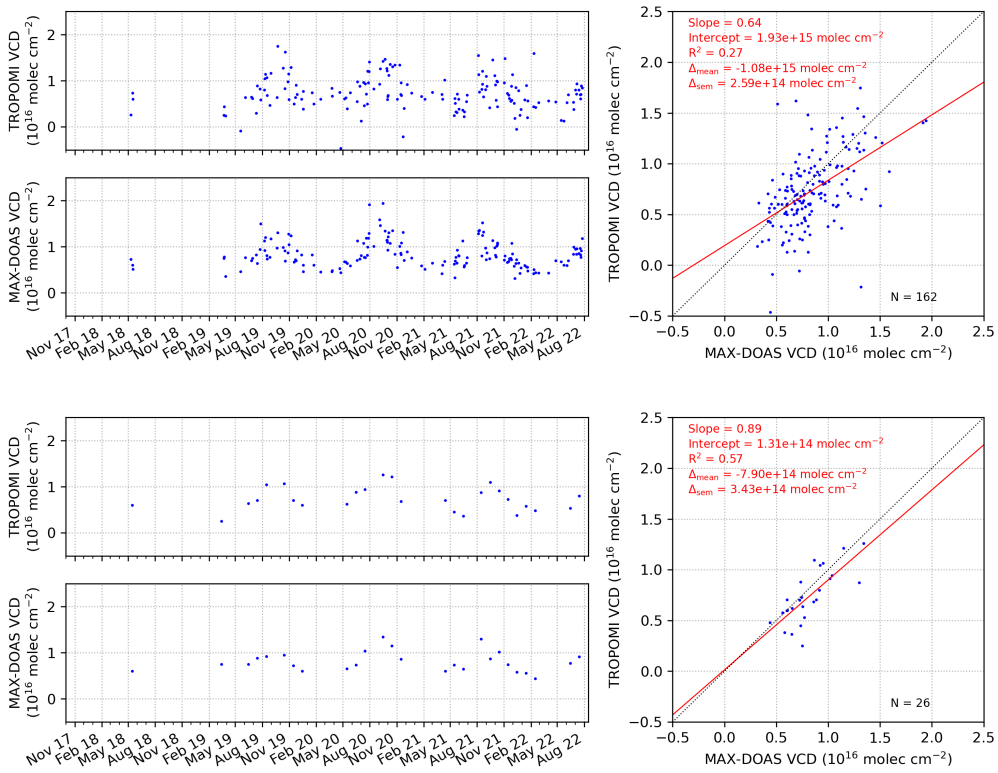


Figure 13.37: Comparison of MAX-DOAS and TROPOMI HCHO VCDs for the ATTO 80 m instrument on daily (upper part) and monthly basis (lower part). Each part consists of time series of TROPOMI (upper left panels) and MAX-DOAS (lower left panels) VCDs as well as the corresponding correlation plots (right panels). The results of a simple linear regression are shown in red and the number of data points used in the comparisons is indicated in the lower right corners.

or underestimation of particularly high or low VCDs by TROPOMI is found and the bulk of the data scatters around the 1-to-1 lines of the correlation plots. Nevertheless, slightly more outliers showing underestimations by TROPOMI exist especially for the 80 m. In sum, a general underestimation of the HCHO VCDs at ATTO by the satellite measurements is observed which is also reflected by the mean differences between the data sets of $(-1.08 \pm 0.26) \times 10^{15}$ molec cm^{-2} (80 m) and $(-8.59 \pm 4.57) \times 10^{14}$ molec cm^{-2} (298 m) corresponding to ca. 13 % (80 m) and 11 % (298 m) of the mean MAX-DOAS HCHO VCDs around the satellite overpass (80 m: 8.59×10^{15} molec cm^{-2}, 298 m: 8.16×10^{15} molec cm^{-2}). No strong gradients in the horizontal distribution of HCHO are expected at ATTO given its remote location and mostly homogeneously distributed biogenic sources. Therefore, mainly different sensitivities of both measurement types to absorbers located close to the ground together with imperfect a-priori assumptions in the TROPOMI HCHO retrieval and cloud effects shielding some of the HCHO in the lowermost parts of the tropospheric column might explain the observed underestimation by TROPOMI.

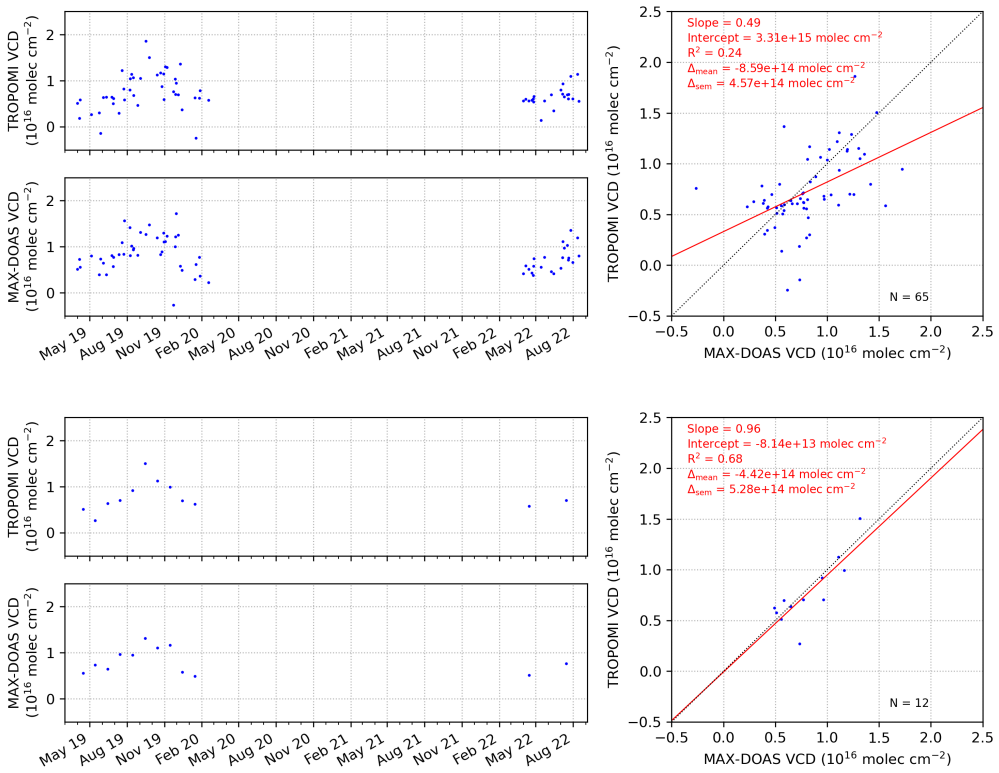


Figure 13.38: Comparison of MAX-DOAS and TROPOMI HCHO VCDs for the ATTO 298 m instrument on daily (upper part) and monthly basis (lower part). Each part consists of time series of TROPOMI (upper left panels) and MAX-DOAS (lower left panels) VCDs as well as the corresponding correlation plots (right panels). The results of a simple linear regression are shown in red and the number of data points used in the comparisons is indicated in the lower right corners.

The comparisons on monthly basis yield better agreements between MAX-DOAS and satellite VCDs for both instruments, whereby the R^2-values are increased to 0.57 (80 m) and 0.68 (298 m). Also the slopes of the linear regressions are improved (closer to 1) for both instruments, while the mean absolute differences decrease to $(-7.90 \pm 3.43) \times 10^{14}$ molec cm^{-2} and $(-4.42 \pm 5.28) \times 10^{14}$ molec cm^{-2} corresponding to approximately 9 % and 5 % of the mean MAX-DOAS VCDs for the 80 m and 298 m instrument, respectively. These improvements are mainly caused by the fact that single daily means showing large deviations between both data sets have less weight in the calculation of monthly medians. A further finding is that the annual variations of the HCHO abundances are nicely captured by TROPOMI. However, it should be noted that only sparse data are included in the comparisons from December to April, corresponding to the transition period towards the wet season and the main part of the wet season itself, caused by only little valid MAX-DOAS retrievals (because of clouds) in the affected periods.

While the daily comparisons perform slightly better for the 80 m instrument given the large scatter of the results for the 298 m instrument, the monthly com-

parisons are seemingly better for the 298 m instrument when only the quantities used for the evaluation of the comparisons are considered. However, the standard error of the mean difference is rather large for the latter caused by the low amount of available data (12 data points). The consistently smaller mean differences for the 298 m instrument are most likely related to the in general smaller VCDs for this instrument. Altogether, it can be concluded that similar results are obtained for both instruments.

If solely the mean differences between MAX-DOAS and satellite HCHO VCDs are considered, the comparisons between both data sets are generally comparable to the ones for Mohali[6] and Bayfordbury and consequently slightly better than for Mainz. This applies in principle to both daily and monthly comparisons, however, on a monthly basis in fact the agreement is best amongst all stations for ATTO. This can probably be explained by the homogenous horizontal HCHO distribution at ATTO. When the R^2 is considered, ATTO performs slightly worse than the other stations especially on a daily scale which might be related to the frequent occurrence of convective clouds at ATTO leading to larger scatter of the data whose effect is minimised by the usage of monthly medians where the effect of a smooth horizontal distribution becomes more important.

Like already done for the other stations, also comparisons using the retrieval results for a fixed O_4 SF of 0.8 or applying different selection criteria for the MAX-DOAS data were performed. The corresponding results of the daily comparisons for both ATTO instruments are summarised in Tables 13.3 and 13.4.

Table 13.3: Comparison results between daily TROPOMI and MAX-DOAS data at ATTO (80 m) for different selection criteria.

SCENARIO	QUANTITY				
	R^2	Slope	Intercept (molec cm^{-2})	Δ_{mean} (molec cm^{-2})	Δ_{sem} (molec cm^{-2})
Standard	0.27	0.64	1.93×10^{15}	-1.08×10^{15}	2.59×10^{14}
SF 0.8	0.29	0.57	2.62×10^{15}	-8.43×10^{14}	3.24×10^{14}
$\pm$60 min	0.28	0.64	1.98×10^{15}	-9.49×10^{14}	2.01×10^{14}
Warnings incl.	0.27	0.60	2.27×10^{15}	-1.13×10^{15}	2.01×10^{14}

For the 80 m instrument, none of the different scenarios significantly changes the comparison results within their uncertainty margins. The slightly different values for the individual comparison quantities are caused by faintly varying amounts of data included in the comparisons or the in-/exclusion of data which are most likely outliers. For the 298 m instrument, larger variations of the comparison quantities between the different scenarios are found. For instance, a notably smaller mean difference between MAX-DOAS and TROPOMI VCDs is found for the fixed SF which is caused by the exclusion of some obvious outliers from the comparison. Further, a largely increased R^2 and a slightly smaller mean

6 An overestimation by TROPOMI is observed at Mohali but the absolute values of the mean differences are comparable.

Table 13.4: Comparison results between daily TROPOMI and MAX-DOAS data at ATTO (298 m) for different selection criteria.

SCENARIO	QUANTITY				
	R^2	Slope	Intercept (molec cm^{-2})	Δ_{mean} (molec cm^{-2})	Δ_{sem} (molec cm^{-2})
Standard	0.24	0.49	3.31×10^{15}	-8.59×10^{14}	4.57×10^{14}
SF 0.8	0.30	0.62	2.74×10^{15}	-3.58×10^{14}	4.08×10^{14}
$\pm$60 min	0.25	0.52	3.11×10^{15}	-6.51×10^{14}	3.01×10^{14}
Warnings incl.	0.50	0.68	2.03×10^{15}	-6.99×10^{14}	3.54×10^{14}

difference are found for the case when also results flagged with a warning are included in the comparison. This can be explained by almost twice as many daily means included in the comparison in this case which improves the statistics. Nevertheless, overall only little data are available for the 298 m instrument and consequently the standard errors of the means are rather large. In conclusion, large overall variability is found for the 298 m instrument making a detailed conclusion difficult.

13.7 FURTHER RESULTS: WATER VAPOUR

For the three stations discussed so far, expected results were obtained for water vapour reflecting the respective large-scale meteorological patterns and the resulting variations of the absolute atmospheric water vapour content at the individual stations. At ATTO, however, interesting differences between the annual cycles of water vapour VCDs and the corresponding concentrations at instrument altitude are found. These differences are discussed in this section using only results for the 80 m instrument as similar results are obtained for both ATTO instruments.

First, the average annual cycles of H_2O VCDs (left) and concentrations at instrument altitude (right) are presented in Fig. 13.39.

The VCDs show a distinct annual cycle with maximum values in the late wet season where a monthly mean of 1.66×10^{23} molec cm^{-2} is obtained in May. During the transition period towards the dry season, the VCDs are decreasing and reach their minimum in the (early) dry season with a monthly mean of ca. 1.23×10^{23} molec cm^{-2} in September. Towards the end of the dry season, the values increase again and stay at comparably high values from December to May, whereby a slight increase from January onwards until May is visible. Contrary to these distinct annual variations, the annual cycle is interestingly clearly weaker for the concentrations at instrument altitude especially in view of the variability within the months themselves. Lowest concentrations at instrument altitude are observed in January where the monthly mean is 0.60×10^{18} molec cm^{-3}. Afterwards, the values are increasing reaching a maximum in the late wet season with a monthly mean of 0.67×10^{18} molec cm^{-3} in May. During the dry season,

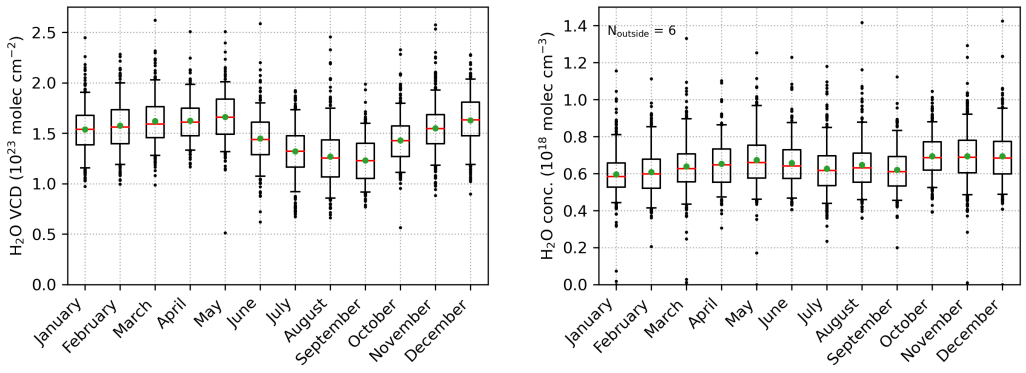


Figure 13.39: Box-whisker-plots of the annual cycles of H_2O VCDs (left) and concentrations at instrument altitude (right) for the 80 m instrument. Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

the concentrations at instrument altitude are only slightly reduced and monthly means are still around 0.64×10^{18} molec cm^{-3} which is still higher than in January. Towards the end of the dry season, the values increase rather abruptly and values similar to the ones in the wet season are found already in October. Actually, the values in the months from October to December are the absolutely highest. However, given their variability they are still comparable to the wet season concentrations at instrument altitude.

As already mentioned above, no diurnal cycle can be observed for neither quantity in none of the seasons.

The finding that a rather clear annual cycle is found for the VCDs, whereas only weak annual variations occur for the concentrations at instrument altitude is also expressed by the temperature dependencies of both quantities which are summarised in Fig. 13.40.

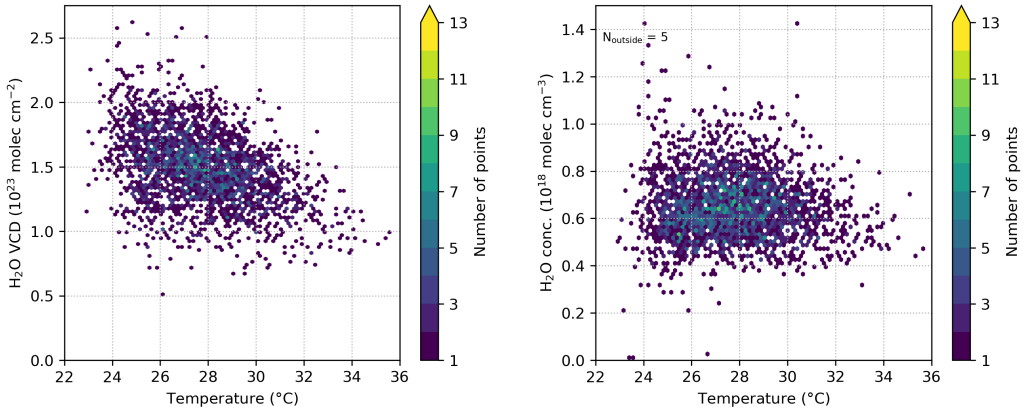


Figure 13.40: Hourly mean H_2O VCDs (left) and concentrations at instrument altitude (right) measured at ATTO (80 m) as function of temperature. Point densities are colour-coded.

A clear tendency showing generally decreasing values with increasing temperatures is found for the VCDs which basically reflects the annual cycle of the VCDs and indicates an overall anti-correlated relation between water vapour VCDs and air temperatures (compare Fig. 4.16 and 13.39). For the concentrations at instrument altitude, no clear temperature dependence is visible. At first glance, these observed temperature dependencies, especially the anti-correlation between water vapour VCDs and air temperature, are surprising since according to the Clausius–Clapeyron equation warmer air masses can contain more water vapour than colder ones. However, these findings suggest that at ATTO the change of the main fetch region in the course of the year is more important for the total atmospheric water vapour content than the temperature dependence (compare discussions below).

Interestingly, the observed annual cycle of the VCDs is not reflected by the wind rose plots and thus no clear dependence on wind speed and wind direction exists for neither quantity. Therefore, no wind rose plots for water vapour are presented here.

So far, the differences in the observed annual cycles of both quantities as well as the unexpected temperature dependencies are explained by a combination of the general meteorological conditions at ATTO located in an undisturbed rainforest with a tropical climate and the change in the origin of the probed air masses during the course of the year. In particular, the main fetch region is north-east during the wet season and air masses reaching the ATTO site passed over the Atlantic Ocean resulting in high atmospheric water vapour contents. Contrary to that, air masses probed during the dry season originate from more easterly to south-easterly directions and travelled longer over continental areas before reaching the ATTO site leading to an absolutely seen less humid character in this season. However, the tropical rainforest itself is rather moist throughout the year since convective precipitation occurs frequently in all seasons. Hence, strong evaporation that moistens the lowest atmospheric layers has to be considered throughout the year. Consequently, the absolute water vapour content in the lowermost layers of the troposphere is constantly high in all seasons leading to comparable concentrations at instrument altitude. The VCDs, however, are dominated by the general character of the air masses which significantly varies between the two main seasons.

This hypothesis is also supported by the median profiles for the different seasons which are shown in Fig. 13.41.

Similar to the other stations, also at ATTO water vapour shows the highest reaching profiles amongst all species, whereby the ATTO measurements yield the highest H_2O profiles even amongst all stations reaching above 2.5 km in all seasons. Only small variations of the H_2O concentrations within the lowest kilometre are visible between the different seasons with the highest concentrations occurring in the dry season. The latter can be explained by typically more sunshine at higher temperatures in the dry season probably leading to strongest evaporation in that part of the year. Above ca. 1 km, the concentrations decrease more rapidly in the seasons for which clearly reduced VCDs are found, namely the transition between wet and dry season and the dry season itself, compared to

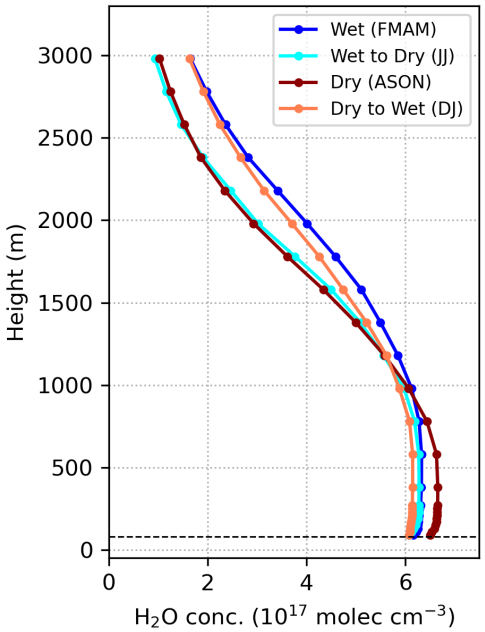


Figure 13.41: Seasonal median profiles of H_2O concentrations obtained for the ATTO measurements (80 m). Profiles for the different seasons are indicated by different colours. The instrument altitude is marked by a dashed black line.

the rest of the year in which higher VCDs are found. Together, this leads to only slightly varying concentrations at instrument altitude while the VCDs are clearly reduced in the drier part of the year as outlined above.

Finally, it can be stated that the discussions in this section also hold for the results obtained with a fixed O_4 SF of 0.8 as well as the results of the 298 m instrument.

Despite overall conclusive explanations for the presented difference between the annual cylces of both quantities, independent measurements, such as radiosonde data, would help to deepen the understanding which could be an aspect for further studies.

13.8 DIRECT COMPARISON OF THE TWO INSTRUMENTS

The basic idea behind operating two MAX-DOAS instruments at two different heights on the tall ATTO tower is to be able to compare their results directly with each other which allows to identify small-scale differences in the vertical distributions of the different species that are hardly resolved by the full vertical profiles. Already some of the previously presented results of both ATTO instruments showed small but systematic differences and suggested slightly different atmospheric processing of some of the investigated species in the rather small altitude range (ca. 200 m) between both instruments. In this section, the trace gas VCDs and AODs as well as the corresponding quantities at instrument altitude obtained with both ATTO instruments are directly compared to each other. For that, hourly mean VCDs/AODs and concentrations/extinctions at instrument

altitude were calculated for the results of both instruments after filtering the data according to the criteria defined in Sect. 9.3. Then the data were visualised as time series and correlation plots for which also simple linear regressions were performed as orientation. Further, the mean differences between the column quantities and quantities at instrument altitude obtained with the 298 m and 80 m instruments (Δ_{mean}) and the standard errors of these means (Δ_{sem}) are provided in the correlation plots. Note that overlapping data for both instruments exist just for roughly one and a half years (only one year for glyoxal) which covers only one full annual cycle.

The comparison results for HCHO VCDs (upper part) and concentrations at instrument altitude (lower part) are presented in Fig. 13.42.

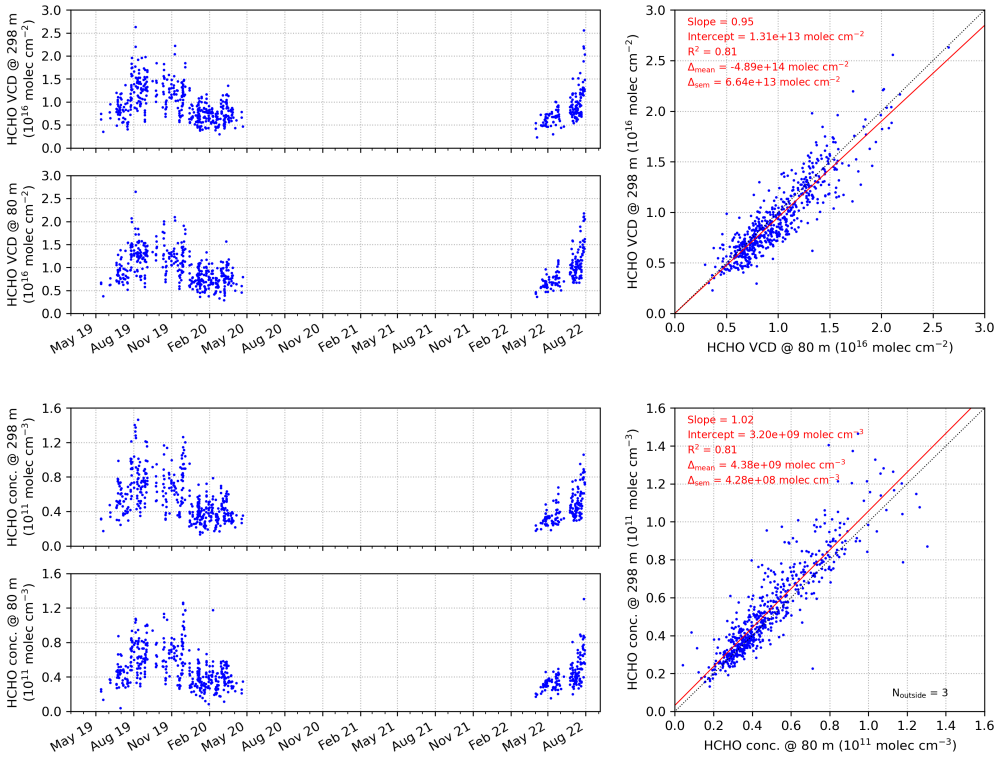


Figure 13.42: Comparison of the formaldehyde VCDs (upper part) and concentrations at instrument altitude (lower part) obtained with both ATTO instruments. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. The results of simple linear regressions are shown in red, while the number of excluded data points is indicated in the lower right corners.

Consistent with the results presented above, the formaldehyde results of both ATTO instruments agree well with slopes close to 1 and an R^2 of 0.81 for both quantities. However, the VCDs are systematically slightly higher for the 80 m instrument and a mean difference of $(-4.89 \pm 0.66) \times 10^{14}$ molec cm^{-2}, corresponding to 5 % of the mean HCHO VCD of the 80 m instrument in the considered measurement period (9.83×10^{15} molec cm^{-2}), is obtained. Interestingly, the opposite is found for the concentrations at instrument altitude which show slightly higher values at 298 m yielding a mean difference of $(4.38 \pm 0.43) \times 10^{9}$ molec cm^{-3}

which is 9 % of the mean concentration at 80 m (4.86×10^{10} molec cm^{-3}). These two findings are consistent with the observations in the previous sections of this chapter and strongly suggest that the HCHO concentrations increase within the lowest 200 m above the canopy because of secondary formation of HCHO from biogenic precursor VOCs leading eventually to high reaching HCHO profiles. However, a non-negligible part of the total column is located below the 298 m instrument and thus not detected by this instrument leading to lower VCDs for the 298 m instrument despite higher concentrations close to the altitude of the instrument (compare profile shapes in Sect. 13.5). Lastly, it can be stated that these findings hold also for the comparison of daily means and the results obtained with a fixed O_4 SF.

Figure 13.43 summarises the comparison of the glyoxal results obtained for both instruments.

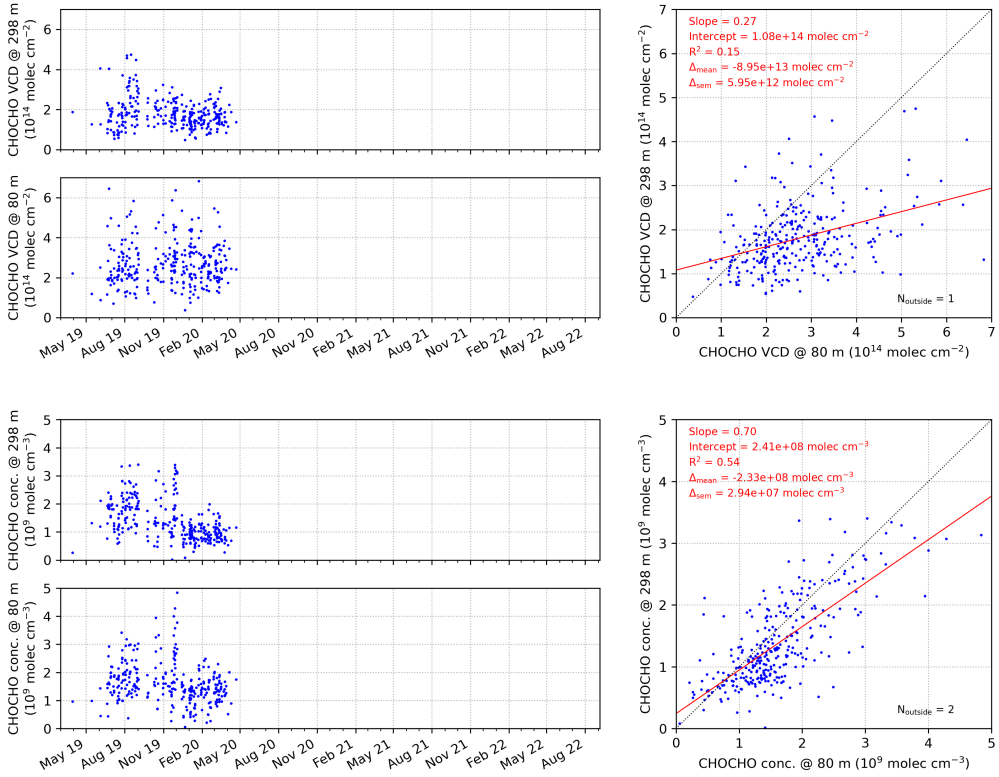


Figure 13.43: Comparison of the glyoxal VCDs (upper part) and concentrations at instrument altitude (lower part) obtained with both ATTO instruments. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. The results of simple linear regressions are shown in red, while the number of excluded data points is indicated in the lower right corners. Note the limited time period for the glyoxal results.

Compared to formaldehyde, clearly worse correlation between the VCDs obtained with both instruments with an R^2 of only 0.15 is found for glyoxal. The mean difference is $(-8.95 \pm 0.60) \times 10^{13}$ molec cm^{-2} which is 33 % of the mean glyoxal VCD of the 80 m instrument (2.69×10^{14} molec cm^{-2}) and indicates notably lower VCDs for the 298 m instrument. Better agreement ($R^2 = 0.54$) is

observed between the concentrations at the altitudes of both instruments. However, again markedly smaller values are found for the 298 m instrument which is reflected by a mean difference of $(-2.33 \pm 0.29) \times 10^8$ molec cm^{-2} corresponding to ca. 15 % of the mean concentration at 80 m (1.60×10^9 molec cm^{-3}). Together, these findings indicate that a considerable fraction of the total lower tropospheric glyoxal column is located below the 298 m instrument and that the glyoxal concentrations decrease significantly already in the lowest 200 m above the canopy. Further, this leads to rather shallow glyoxal profiles and confirms the previous hypothesis that, despite most likely common sources, HCHO and glyoxal are processed (formed and degraded) differently in the atmosphere. Another aspect might be a larger fraction of directly emitted glyoxal compared to HCHO or probably a more rapid glyoxal formation from common precursors. Again, similar results are obtained for the comparison on a daily basis and the results obtained with a fixed O_4 SF.

The comparisons of NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) obtained with both ATTO instruments are shown in Fig. 13.44.

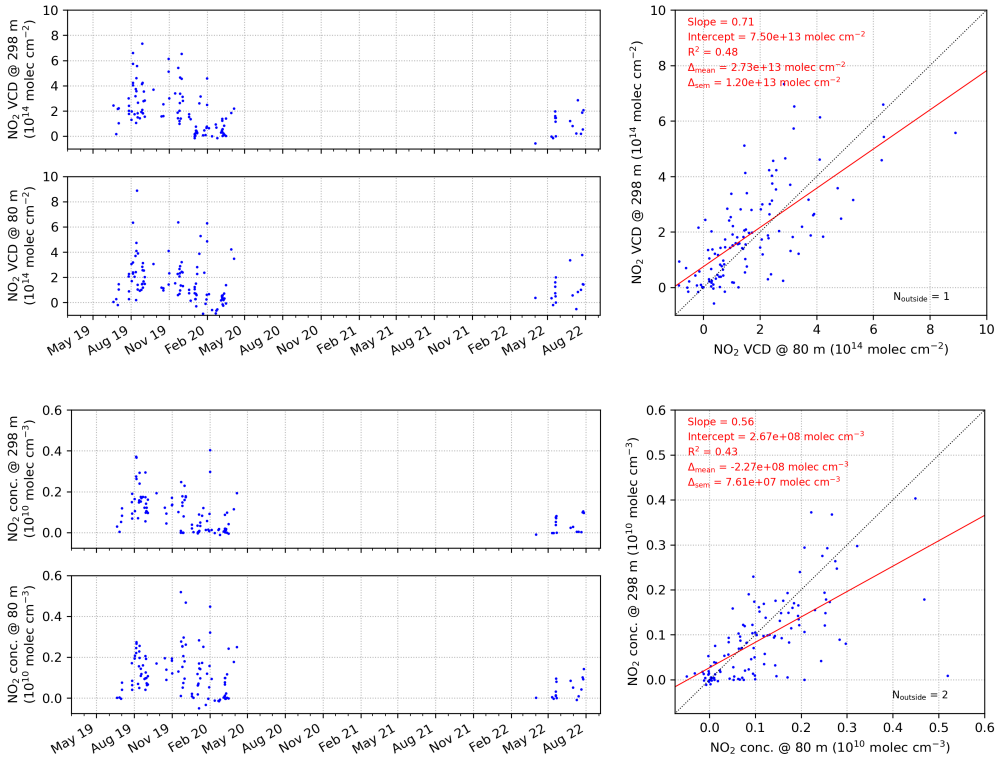


Figure 13.44: Comparison of the NO_2 VCDs (upper part) and concentrations at instrument altitude (lower part) obtained with both ATTO instruments. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. The results of simple linear regressions are shown in red, while the number of excluded data points is indicated in the lower right corners.

The NO_2 results obtained with both ATTO instruments agree overall reasonably well, whereby R^2-values of 0.48 and 0.43 are obtained for VCDs and

concentrations at instrument altitude, respectively. On average, slightly higher VCDs are obtained for the 298 m instrument with a mean difference of $(2.73 \pm 1.20) \times 10^{13}$ molec cm^{-2} corresponding to approximately 17 % of the mean NO_2 VCD of the 80 m instrument (1.62×10^{14} molec cm^{-2}). For the concentrations at instrument altitude, however, higher values are measured at 80 m and a mean difference of $(-2.27 \pm 0.76) \times 10^8$ molec cm^{-3} is found (20 % of corresponding mean; 80 m: 1.14×10^9 molec cm^{-3}). These observations are somehow contradicting and no clear explanation can be given. However, the differences are (absolutely seen) very small and many measurements with small NO_2 abundances were filtered out by MAPA (compare above). Especially the latter can lead to different selections for both instruments and probably explain the contradicting results. Nevertheless, they are similar to the findings obtained for the vertical profiles of NO_2 which showed higher reaching profiles for the 298 m instrument in the transition towards the dry season and in the dry season itself, while significant NO_2 in the transition period towards the wet season was only found for the 80 m instrument (compare Sect. 13.5). Here, it should be stressed that only one dry season was covered by the 298 m instrument so far. Thus, these unclear results can probably be improved in future when more measurements are available. This is also supported by the finding that different results showing larger VCDs for the 80 m instrument are obtained for the fixed scaling factor. Thus, future comparisons including longer time series are needed for NO_2.

Finally, Fig. 13.45 summarises the comparison of the aerosol results obtained with both ATTO instruments.

Also for the aerosol results, rather good agreement between the results of both instruments is found and R^2-values of 0.59 and 0.45 are obtained for AODs and extinctions at instrument altitude, respectively. Consistently slightly larger values are measured with the 80 m instrument for both quantities with mean differences of -8.46×10^{-3} and -5.74×10^{-3} km^{-1} corresponding to about 5 % of the respective means of the results for the 80 m instrument in the considered measurement period (0.15 and 0.11 km^{-1}). Especially, the AODs show rather large variability for the highest aerosol loads probably due to (valid) measurements which are affected by clouds despite the rather strict filter criteria. The observation that slightly larger AODs and aerosol extinctions at instrument altitude are found for the 80 m instrument is consistent with the previous findings in this chapter, e.g. the results for the profiles obtained for both instruments (compare Sect. 13.5), and indicates again that a certain part of the total lower tropospheric aerosol column is located below the 298 m instrument and thus not measurable for this instrument (similar to the observations made for HCHO). The same findings hold for the daily comparison and the results obtained with the fixed O_4 SF of 0.8, whereby the mean difference of the AODs is more than doubled for the latter case.

Already the simple comparisons between the results obtained with both MAX-DOAS instruments installed at ATTO for this rather short time period of only one and a half years presented in this section provide interesting insights in small-scale variations of the trace gases within the lowest 200 m above the canopy. Potential further studies should include longer time series and investigations on a

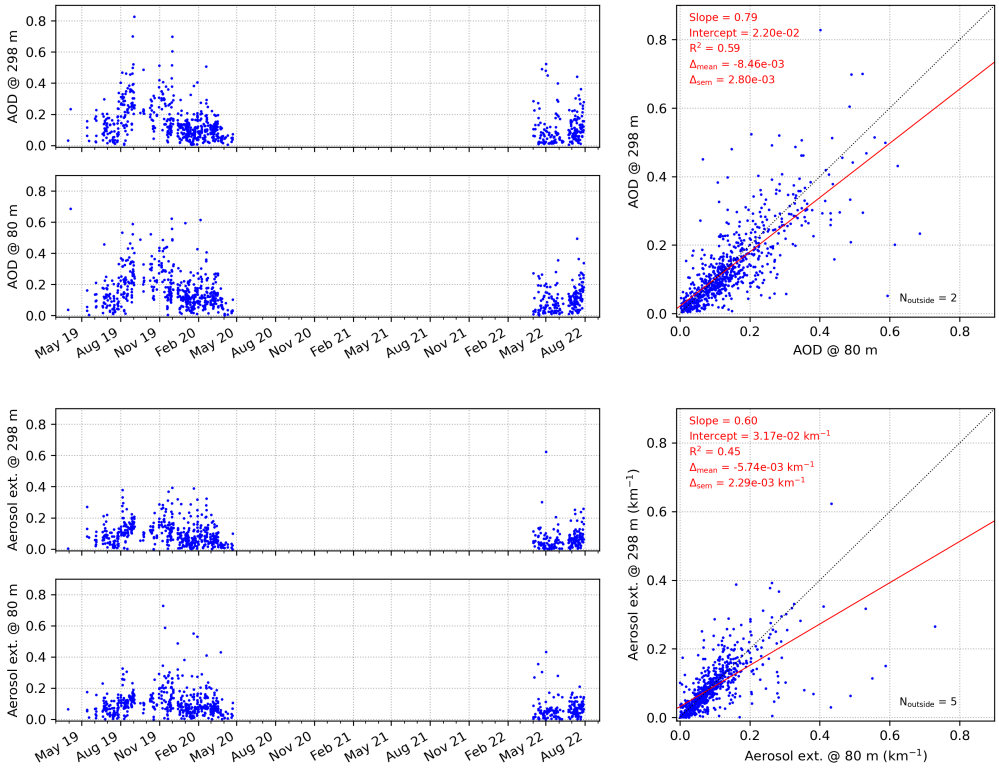


Figure 13.45: Comparison of AODs (upper part) and aerosol extinctions at instrument altitude (lower part) at 367 nm obtained with both ATTO instruments. Left panels show time series of hourly mean values, while the right panels show the corresponding correlation plots. The results of simple linear regressions are shown in red, while the number of excluded data points is indicated in the lower right corners.

seasonal basis. Moreover, the investigation of diurnal variations of the differences between the results of both instruments as well as the comparison to the vertical distribution of potential precursor substances of glyoxal and HCHO such as isoprene and monoterpenes might yield interesting insights.

Part V

CONCLUSIONS AND OUTLOOK

14

CONCLUSIONS

The overall aim of this work is to contribute to a better understanding of the sources and the atmospheric processing of VOCs. For this, ground-based MAX-DOAS measurements at four different measurement sites were used to create a consistent long-term data set of HCHO concentration profiles and vertical column densities (VCDs) in different parts of the world covering different climatic zones and environmental conditions. The HCHO data set was amended by glyoxal, NO_2 and aerosol results.

Although satellite measurements of HCHO (and the other species) exist providing a global view on its distribution, they lack many details such as vertical distributions (profiles) and diurnal cycles (for most satellites). Further, satellite retrievals, especially of HCHO, are associated with large uncertainties and a decreasing sensitivity towards the ground.

These limitations can be addressed by MAX-DOAS measurements which are usually more than ten times more sensitive to trace gases located close to the ground compared to satellites. However, in the past MAX-DOAS measurements have been analysed with a great variety of analysis settings leading to difficulties when different data sets from different stations and/or data providers are compared.

Therefore, a major focus of this work was to develop consistent analysis settings. Hence, a large part was dedicated to a sensitivity study in which the DOAS analysis settings of HCHO were systematically assessed and optimised with a great emphasis on consistency for many atmospheric conditions using both simulated and measured spectra. Finally, recommendations for uniform DOAS analysis settings for HCHO were derived. Here, a major improvement was the consistent use of Puķīte terms for NO_2 and O_3. The usage of Puķīte terms for NO_2 was shown to be essential under high NO_2 conditions while they have almost no adverse effects on the fit results under low NO_2 conditions. Moreover, systematic uncertainties of the HCHO analysis were investigated, minimised and the remaining uncertainties estimated. Here, limitations of the DOAS method itself were identified to be the main reason for the remaining uncertainties. Improvements were also made for the analysis settings of the other species and the DOAS retrieval in general. One important result in this regard was the inclusion of a water vapour cross-section in the O_4 fit to improve the fit results significantly especially under very humid conditions (like in the Amazon region). In addition, the glyoxal (and NO_2) fit was improved by minimising strong residual structures caused by strong water vapour absorptions using a pseudo absorption cross-section and by the optimisation of the fitting interval as well as the inclusion of Puķīte terms for NO_2 which was again shown to be essential under high NO_2 conditions. The DOAS retrieval can be generally improved by using a Super Gaussian instrumental slit function instead of a simple Gaussian

one. Further, it was found that also in this thesis the usage of an O_4 scaling factor (SF) is needed and that small improvements of the MAPA profile inversion can be achieved by using the LUTs at so-called effective wavelengths.

These optimised retrieval settings were then used to analyse the MAX-DOAS measurements conducted at the four different sites resulting in a consistent data set of HCHO concentration profiles and VCDs. The HCHO results were also compared with the results of the other analysed species, namely glyoxal, NO_2 and aerosol, whereby especially the comparisons with glyoxal and aerosol showed many similarities. One speciality for the measurements at ATTO is that high resolution (small-scale) concentration gradients could be determined by the direct comparison of the results of both instruments installed at altitudes of 80 m and 298 m. A last aspect addressed in this thesis was the comparison of the HCHO VCDs to satellite (TROPOMI) measurements.

Based on the obtained consistent data set, several scientific questions were addressed and implications for atmospheric chemistry were derived.

In general, HCHO and glyoxal have a large variety of sources which can be categorised into three main categories, namely anthropogenic and biogenic sources as well as biomass burning which can be of both anthropogenic (e.g. crop residue or deforestation fires) and natural (forest/bush fires caused by lightning in tropical regions) origin. Hence, a major aspect of this work was the identification and to some extent also the differentiation of the different sources of HCHO and glyoxal (and partly aerosol) under different environmental conditions. This was attempted by considering several facets, such as annual and diurnal cycles of the trace gas (and aerosol) abundances, their dependencies on air temperature and wind direction as well as vertical concentration profiles, and combining the results of the different species mentioned above.

Anthropogenic sources of HCHO and glyoxal as well as the influence of biomass burning can be identified by the simultaneous occurrence of enhanced NO_2 abundances in both temporal (e.g. annual cycles) and spatial (e.g. wind dependencies) dimension. For example, even at ATTO where overall biogenic precursor VOCs are the dominant HCHO (and glyoxal) source, the similar annual cycles of the NO_2 and HCHO abundances, both showing higher values in the dry season, suggested a substantial effect of biomass burning events on the HCHO abundances. This was further supported by rather confined patterns in the wind dependencies of both species under south-easterly winds, when the air masses probed at ATTO originate from deforestation fire regions and the more densely populated north-eastern part of Brazil (compare below).

Additional ways to discriminate between the different source categories are the investigation of temperature dependencies or trace gas correlations. For example, a strong temperature dependence indicates a strong contribution of biogenic precursors because the release of precursor species from vegetation strongly depends on temperature. For the data set analysed here, a strong (non-linear) temperature dependence of the HCHO (and glyoxal) abundances indicated the importance of biogenic precursors at Bayfordbury and Mainz. At the same time, reasonable correlations between HCHO and NO_2 during winter suggested that in this season anthropogenic emissions play an important role for the measured

HCHO abundances. Based on similar investigations for all measurement sites, it was concluded that both biogenic and anthropogenic sources can play a role at all stations, whereby biogenic sources are often dominant particularly in the warmer parts of the year. However, as outlined in detail in the main part of this thesis, the relative contributions of biogenic and anthropogenic sources vary significantly between the different sites and seasons.

All three source categories can be associated with comparably small amounts of directly emitted HCHO (even smaller for glyoxal) and, to a larger extent, emissions of precursors forming HCHO and glyoxal during subsequent chemical processing. Consistent to existing literature, the results of this work indicate that overall secondary formation from precursors dominates, regardless of whether the precursors are of anthropogenic or biogenic origin. An important finding leading to this conclusion were the pronounced annual variations of the HCHO and glyoxal abundances that showed highest values during parts of the year in which solar radiation and temperatures are high leading to enhanced sunlight-driven photochemistry. In general, this is valid for both biogenically and anthropogenically dominated measurement sites. However, there were several indications that especially during summer biogenic precursors dominate (in particular at Bayfordbury and Mainz) or have at least a large contribution (Mohali). Given its location in a pristine rainforest region, at ATTO biogenic emissions dominate throughout the year. For HCHO, the dominance of secondary formation was also deduced from the finding that often the different dependencies are rather similar for near-surface concentrations and VCDs. Another important finding was that HCHO typically exhibits the highest reaching profiles amongst all species (with aerosol often reaching to similar altitudes) at all four measurement sites indicating the large effect of secondary formation of HCHO from precursors. Also at Mohali, HCHO showed high reaching profiles showing that also here, despite strong anthropogenic influence, typically secondary formation dominates the HCHO abundances. These observations were further supported by the direct comparison of the HCHO results of both ATTO instruments (compare below). Although for glyoxal the profiles are notably shallower, the close connection between HCHO and glyoxal strongly suggested that also for glyoxal overall secondary formation dominates which is in particular supported by similar temperature dependencies of both species.

Despite the conclusion that secondary formation generally dominates the sources, there were indications that also direct (anthropogenic) emissions can sometimes significantly contribute to the prevailing HCHO abundances especially in the absence of strong biogenic activity during winter at Bayfordbury and particularly at Mainz and Mohali where several anthropogenic sources lie in the vicinity of the instruments. As a consequence of the latter, especially at these two sites the effect of long-range transport is smaller. Further, at these two sites the profiles are markedly shallower and the layer containing the bulk of HCHO is more confined in winter compared to the other seasons which is not the case at Bayfordbury where major anthropogenic emission sources are located at larger distances from the site. Although the HCHO profiles are still the highest amongst all species and likely also less efficient vertical mixing as well as

a shallower boundary layer contribute to the shallower profiles, it was concluded that especially at Mainz and Mohali direct emissions of HCHO from mostly traffic and fires for domestic heating (at Mohali) can play an important role during winter. This was additionally supported by the close relation of the near-surface concentrations of HCHO and NO_2 during winter. While the latter is also the case at Bayfordbury, here the profiles did not show clearly different shapes in the different seasons. These findings might also apply to other (especially mid-latitude) cities which have a similar character like Mainz and the Rhine-Main region or Mohali, i.e. heavy traffic and dense population.

One of the most evident findings of the studies regarding the sources of HCHO and glyoxal was that both species have similar or even common sources which is reflected in a number of different aspects. For instance, both species showed comparable annual/seasonal variations at all measurement sites and their respective 1-hour means were often highly correlated. Furthermore, the dependencies of the abundances of both species on air temperature and the wind direction (and partly speed) showed often very similar patterns. These observations were particularly pronounced for the near-surface concentrations which are supposed to be stronger affected by the sources. To a smaller extent, similar findings also hold for aerosol.

Although HCHO and glyoxal apparently have mostly similar or even common sources, the results of this work also hinted that both species are photochemically processed differently in the atmosphere likely affecting both the formation from precursors and the degradation. Several findings of this work suggested that usually the secondary formation of HCHO occurs over a longer time period while glyoxal is formed for a shorter time and/or faster. On the other hand, also the degradation of glyoxal apparently occurs faster. In particular, at all measurement sites (with only few exceptions, e.g. at Mohali in winter) the vertical profiles of the glyoxal concentrations are shallower and the layers containing the bulk of glyoxal are vertically less extended compared to the high-reaching HCHO profiles indicating that the majority of glyoxal is typically located closer to the ground. Consistent to that, higher correlations between the abundances of both species were generally found for the near-surface concentrations which showed overall more similar patterns in various other aspects (e.g. annual variations) compared to the VCDs. In addition, the ratio of glyoxal and HCHO (R_{GF}) calculated from the near-surface concentrations was typically higher compared to the corresponding R_{GF} calculated from the VCDs. At ATTO, also the direct comparison of the HCHO and glyoxal results of both instruments suggested different chemical processing of both species since the HCHO concentrations increase (still ongoing formation of HCHO) in the altitude range between both instruments (ca. 200 m) while the glyoxal concentrations already decline indicating a net production of HCHO and a net degradation of glyoxal in this altitude range.

The investigation of the seasonal variation of the R_{GF} provides insights into the composition of the precursor species of glyoxal and HCHO. This is particularly the case for chemical regimes that are dominated by VOCs of biogenic origin since the precursor composition can become very complex under anthropogenic influence and more detailed characterisations are needed in such cases. For

instance, from the annual variations of the R_{GF} at ATTO it was concluded that, consistent to previous studies, different compositions of precursor VOCs prevail in the two main seasons with a larger relative contribution of monoterpenes in the wet season compared to the dry season in which isoprene alone dominates the overall regime.

Another interesting aspect is to which extent the measured HCHO abundances at the different sites are affected by transport processes from more distant sources. Given the rather short lifetime of HCHO (and glyoxal) of a few hours during day time, often local effects, such as biogenic activity (compare temperature dependence), availability of precursor VOCs or local pollution levels, dominate the abundances. Nevertheless, there were indications that in several cases long-range transport or transport from more nearby sources can become important although the exact contribution of transported emissions (from different distances) to the measured HCHO abundances is difficult to quantify. For instance, the dependence on the wind direction at Bayfordbury clearly showed the effect of transported anthropogenic emissions on the measured HCHO abundances and partly even allowed an attribution to distinct sources located within at maximum ca. 60 km to the measurement site. At the same time, there were also indications for emissions contributing to the local HCHO abundances originating from sources up to 350 km away. Even at ATTO where the rainforest itself is the main HCHO source, rather confined patterns in the wind rose plots suggested additional effects of transported emissions from deforestation fire regions and the more densely populated north-eastern part of Brazil (most likely HCHO is formed from precursors during the transport) during the dry season over distances of a few hundred kilometres. Besides these examples, indications of effects of transport processes on the measured HCHO abundances also existed for the other two sites where the effects are, however, apparently smaller given the vicinity of possible (anthropogenic) emission sources.

Finally, the MAX-DOAS HCHO data from the four sites were used to perform a consistent and comprehensive comparison to satellite HCHO data from the TROPOMI instrument. In general, reasonable to good agreement between both data sets was found for all stations independent of the specific environmental conditions. The agreement was further improved when monthly median VCDs were considered instead of daily values. For the two mid-latitude stations Bayfordbury and Mainz, best agreement was obtained for medium VCDs, whereas larger discrepancies between both data sets occurred for particularly high or low HCHO abundances. Here, especially the highest HCHO VCDs were systematically underestimated by TROPOMI. Effects contributing to these observations are different sensitivities of both measurement types to absorbers close to the ground and horizontal gradients (very important in the vicinity of cities) which are not captured by satellite measurements due to their limited spatial resolution. The fact that the satellite data were averaged 20 km around the sites as well as imperfect a-priori assumptions in the TROPOMI retrieval might additionally contribute to the differences between both data sets. Further, at these two stations the pronounced annual cycles observed for the MAX-DOAS measurements are not well captured by TROPOMI likely caused by the small amount of available

data in late autumn and during winter as well as underestimations of the highest VCDs in summer. While TROPOMI yields smaller VCDs than the MAX-DOAS measurements at Bayfordbury, Mainz and ATTO, the opposite was found at Mohali where surprisingly higher VCDs were retrieved from the satellite measurements. This finding is likely related to an inappropriate correction of cloud effects in the TROPOMI retrieval. In particular, high aerosol loads at Mohali are probably interpreted as clouds by the retrieval algorithm and thus a too small air mass factor (AMF) is applied while usually high aerosol amounts cause an increase of the AMF. At ATTO, where usually no large horizontal gradients in the HCHO distribution are expected, best agreement (in terms of mean differences) between both data sets was found although only little data were included in the comparison in the transition period towards the wet season and in the wet season itself due to the frequent occurrence of convective clouds. The comparison between TROPOMI and MAX-DOAS data also indicated a strong influence of the cloud screening of the satellite data. Other selection criteria (e.g. including MAX-DOAS results flagged with a warning) did not lead to notably different comparison results.

Altogether, the results of this work showed that biogenic emissions of precursors of HCHO are important under all environmental conditions. At the same time, there was evidence for the influence of anthropogenic emissions at all stations which shows how mankind is affecting air chemistry and quality all over the world. The overall dominance of secondary formation was also proven at all sites while at the more anthropogenically influenced sites Mainz and Mohali also direct emissions can play a role during winter when biogenic activity and photochemistry are low. Lastly, the data set was successfully used for satellite validation.

15 OUTLOOK

Although successful efforts to improve the quality and consistency of the retrieval results were made in this work, there are several aspects that should be addressed in the future to further improve MAX-DOAS results.

Regarding the spectral retrieval, the usage of accurate instrumental slit functions measured using spectral line lamps is recommended based on the results in Sect. 7.2 and 8.1 to minimise or even avoid residual structures caused for instance by strong water vapour absorptions. This should in particular be done for the spectral retrieval of glyoxal but could also improve the fit results of other species. Another aspect is to investigate the influence of the so-called tilt effect on the spectral retrieval. Especially, the remaining residuals in the analyses of synthetic spectra can likely be further reduced by the usage of "tilt-correction" spectra.

Apart from the spectral retrieval, the profile retrieval using MAPA should be further improved. In particular, modifications of the profile parametrisation are suggested to represent also purely exponential profiles which is of particular importance for species directly emitted at the ground like NO_2. Also the flagging scheme of the MAPA results should be updated. Firstly, more valid results for small VCDs (e.g. NO_2 at ATTO) are desirable. Such inversions should not be filtered out in future although the relative errors of such measurements are typically large (while the absolute errors are small). Secondly, the aerosol results at ATTO showed that many retrievals yielding high AODs are flagged (mostly warnings) by MAPA, even if the AOD is clearly below the threshold of the corresponding AOD flag. In that way, especially episodes of biomass burning events were affected. Hence, corresponding modifications of the MAPA flags are recommended.

There is also the need to better understand the O_4 scaling factor. Here, the main question is why sometimes the usage of an O_4 SF is necessary and sometimes not. These studies should also address the way the scaling factor apparently compensates for other effects, e.g. inaccurate elevation calibrations or problematic results for specific viewing geometries (e.g. artefacts in the diurnal cycles observed at Mohali and ATTO).

In addition to these technical aspects aiming to improve the DOAS retrieval, efforts should be made to extend and refine the interpretation of the results.

Foremost, the results obtained from the MAX-DOAS measurements should be compared to other (independent) measurements. Comparisons with measurements of the same species using other measurement techniques, e.g. in-situ measurements of NO_2 (or NO_x) or HCHO, could be used to verify the annual or diurnal patterns observed for the MAX-DOAS data. More importantly, comparisons with measurements of other species are needed. For instance, comparisons of the HCHO and glyoxal results at ATTO (but not exclusively) with in-situ measurements of precursor VOCs (e.g. isoprene) are important, especially to further

confirm changes in the chemical regimes (or VOC compositions) between the different seasons. Here, more detailed studies of the R_{GF} should be additionally performed. These studies of the R_{GF} might include investigations of diurnal cycles and vertical profiles as well as dependencies on wind patterns or temperature.

Another aspect for the measurements at ATTO is the comparison of the height profiles to measurements performed on an elevator providing further insights into small-scale gradients. In this regard, also more detailed analyses of the small-scale gradients obtained by the direct comparison of the results from both ATTO instruments are needed. For instance, these small-scale differences in the trace gas concentrations between both instruments could be investigated with respect to seasonal or diurnal variations. Further, dependencies on pollution levels or meteorological parameters should be assessed.

Besides that, the MAX-DOAS measurements should also be compared to model simulations. While this is generally of interest for all stations, in particular the unexpected diurnal cycles of the trace gas abundances at ATTO (especially the "U-shape" of the HCHO abundances) could be further studied by such comparisons to model results.

This work also showed that for further species important and unexpected results can be obtained. For instance, enhanced HONO abundances were successfully retrieved at Mohali and Mainz. Further studies should be conducted to explore in how far MAX-DOAS measurements can contribute to gain deeper insights into the production and degradation mechanisms of HONO which are closely linked to radical chemistry. Besides that, studies to confirm the explanation for the unexpected annual cycles of water vapour VCDs and concentrations at ATTO are suggested including the comparison to radiosonde data.

Part VI

APPENDIX

A

SUPPLEMENTARY TABLES

Here, two supplementary tables are provided to complement the information provided in the main part.

Table A.1 gives an overview of the instrument type (compare Sect. 4.2.1 in the main part) operated at each of the different stations. Further, the respective elevation sequences employed at the different stations are listed together with the corresponding total integration times used to acquire an accumulated spectrum. For instruments using different elevation sequences, the sequences employed in the respective periods are listed. Note that the instrument operated at Mainz has a completely different setup than the TubeMAX-DOAS instruments used at the other stations (compare Sect. 4.1.1). As a consequence, also the measurement routine is different in which the number of scans is prescribed rather than the total integration time. Hence, the latter varies accordingly. Further, it should be mentioned that for the ATTO instrument at 80 m various elevation sequences were tested in the second measurement period (03/2018 – 09/2019). The sequence listed in the table below represents the most frequent one in this time period. Similarly, the elevation sequence employed by the Mainz instrument varied sometimes due to technical and/or software issues. However, during nearly the entire measurement period the sequence listed below was used. For the profile retrievals using MAPA, only positive elevation angles are considered for all instruments. In addition, also (low) elevations which are likely affected by obstacles are excluded from the profile retrieval (compare Table 9.7 in Sect. 9.2.2 in the main text). Lastly, for the ATTO instruments no 90° spectra can be recorded due to limitations of the mounting on the tower. Hence, spectra at 88° and 87° are used as FRS for the 80 m and 298 m instrument, respectively.

Table A.2 provides an overview of the R^2-values obtained for the trace gas correlations on a seasonal basis at the different measurement sites for VCDs/AODs (VCD) and near-surface concentrations/aerosol extinctions (SFC). The corresponding descriptions can be found in the respective (trace gas correlation) sections in the main part of this work. Also explanations of the filter criteria that were applied to the data before the correlations were performed are given in the main text. Note that only values for the two main seasons occurring at the different stations are provided and that for the ATTO 298 m instrument only glyoxal data until 21 April 2020 were considered. Since the glyoxal data at Mainz are of insufficient quality, no correlations including glyoxal were performed at all.

Table A.1: Overview of the instrument types, elevation sequences and total integration times at the different stations. For instruments using different elevation sequences, the sequences of the respective periods are listed.

STATION		TYPE	ELEVATION SEQUENCE	INTEGRATION TIME
Bayfordbury	11/2018 – 07/2022	RaspberryPi-based	−1°, −0.5°, 1°, 2°, 3°, 4°, 5°, 6°, 7°, 8°, 9°, 10°, 15°, 30°, 90°	60 s
Mainz	06/2013 – 03/2022	4-azimuth	−1.9°, 0°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 90°	variable
Mohali	04/2019 – 07/2022	Laptop-based	0°, 1°, 2°, 3°, 5°, 8°, 10°, 15°, 30°, 90°	30 s
ATTO 80 m	10/2017 – 03/2018	RaspberryPi-based	−4°, −1°, −0.2°, 0°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 45°, 80°, 88°	60 s
	03/2018 – 09/2019		−15°, −8°, −4°, −2°, −1°, −0.5°, 0.5°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 45°, 80°, 88°	
	09/2019 – 07/2022		−10°, −8°, −4°, −2°, −1°, −0.5°, 0.5°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 45°, 80°, 88°	
ATTO 298 m	03/2019 – 09/2019	RaspberryPi-based	−15°, −8°, −4°, −2°, −1°, −0.5°, 0.5°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 45°, 80°, 87°	60 s
	09/2019 – 09/2022		−10°, −8°, −4°, −2°, −1°, −0.5°, 0.5°, 1°, 2°, 3°, 4°, 5°, 6°, 8°, 10°, 15°, 30°, 45°, 80°, 87°	

Table A.2: Overview of the R^2-values obtained for the trace gas correlations on a seasonal basis at the different measurement sites for VCDs/AODs (VCD) and near-surface concentrations/aerosol extinctions (SFC). The corresponding descriptions can be found in the respective (trace gas correlation) sections in the main part. Only values for the two main seasons occurring at the different stations are provided.

STATION	SEASON	GLYOXAL/HCHO		HCHO/NO_2		GLYOXAL/NO_2		HCHO/AEROSOL		GLYOXAL/AEROSOL		NO_2/AEROSOL	
		VCD	SFC	VCD	SFC	VCD	SFC	VCD	SFC	VCD	SFC	VCD	SFC
Bayfordbury	Summer	0.58	0.58	0.10	0.12	0.19	0.16	0.32	0.17	0.15	0.06	0.05	0.13
	Winter	0.19	0.11	0.24	0.34	0.14	0.12	0.05	0.10	0.05	0.09	0.01	0.15
Mainz	Summer	–	–	0.01	0.09	–	–	0.16	0.27	–	–	0.03	0.17
	Winter	–	–	0.11	0.33	–	–	0.17	0.33	–	–	0.10	0.26
Mohali	Pre-monsoon	0.54	0.64	0.20	0.43	0.19	0.33	0.37	0.61	0.20	0.42	0.08	0.37
	Winter	0.34	0.43	0.08	0.27	0.07	0.13	0.31	0.43	0.15	0.11	0.06	0.18
ATTO 80 m	Dry season	0.16	0.49	0.18	0.24	0.07	0.16	0.20	0.21	0.03	0.04	0.10	0.12
	Wet season	0.10	0.21	0.04	0.04	0.10	0.21	0.01	0.01	0.00	0.00	0.00	0.01
ATTO 298 m	Dry season	0.27	0.49	0.23	0.20	0.08	0.18	0.25	0.22	0.00	0.04	0.09	0.17
	Wet season	0.12	0.24	0.03	0.02	0.00	0.02	0.06	0.06	0.00	0.01	0.00	0.01

B

SUPPLEMENTARY FIGURES

Here, supplementary figures are provided to complement the information provided in the main part.

B.1 SENSITIVITY STUDY FOR THE DOAS FIT SETTINGS OF HCHO

Figure B.1 presents an example for the RMS (upper panel) and sums (lower panel) of the absolute deviations from the true reference dSCDs (calculated as described in the main part) for all 16 combinations of NO_2 cross-sections for an individual sequence in the case of high NO_2 and HCHO, an SZA of 45° and an RAA of 180° (scenario used in the main part). The most prominent feature is that the four combinations not including any Puḳīte term for NO_2 clearly yield the largest RMS_{abs} indicating large deviations from the true reference dSCDs. As shown in Fig. 6.5 in the main part (compare Sect. 6.2.2.1), large underestimations of the reference dSCDs are found for the lowest elevations which is reflected by negative sums of the absolute deviations for these four combinations of NO_2 cross-sections.

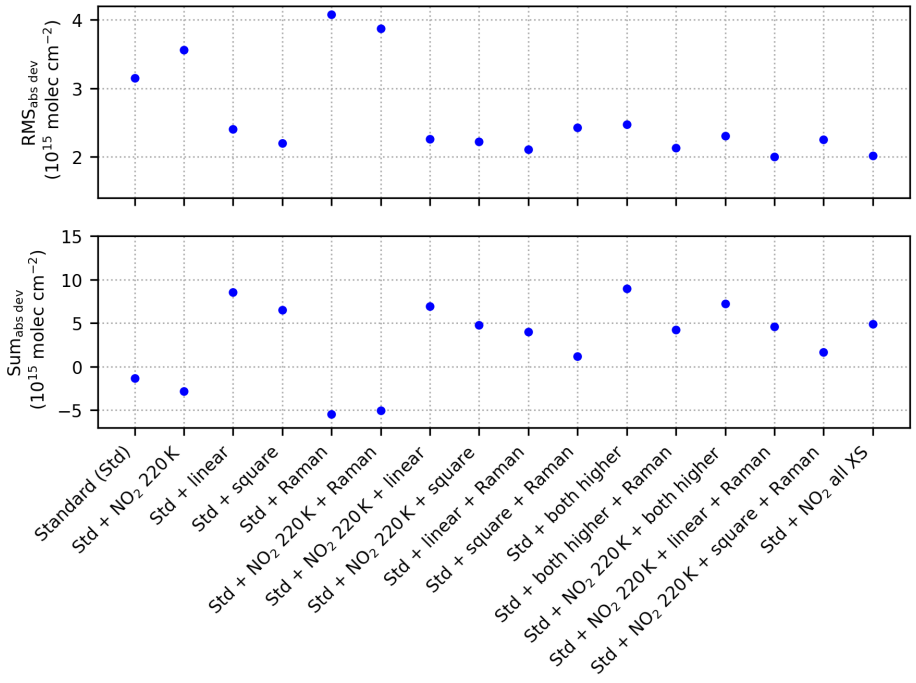


Figure B.1: RMS (upper panel) and sums (lower panel) of the absolute deviations from the true reference dSCDs for all 16 combinations of NO_2 cross-sections for the case of high NO_2 and HCHO, an SZA of 45° and an RAA of 180°.

Figure B.2 presents an example for the RMS (upper panel) and sums (lower panel) of the *CQ*-values (compare description in the main part) for all 16 combinations of NO_2 cross-sections for an individual sequence in the case of high HCHO, an SZA of 45° and an RAA of 180° (scenario used in the main part). Again, the four combinations not including any Puķīte term for NO_2 yield the largest RMS_{CQ} indicating overall large differences between the absolute deviations obtained for the high and low NO_2 scenarios for these combinations of NO_2 cross-sections. As shown in Fig. 6.8 in the main part (compare Sect. 6.2.2.1), for these combinations large underestimations of the reference dSCDs are found in the case of high NO_2 at the lowest elevations. Since this is not the case for the low NO_2 scenario, for these four combinations of NO_2 cross-sections *CQ* becomes largely negative at low elevations which is reflected by largely negative sums in the lower part of Fig. B.2. Note that many combinations yield slightly negative sums since slightly smaller deviations from the reference dSCDs are found for several combinations of NO_2 cross-sections (compare Fig. 6.8).

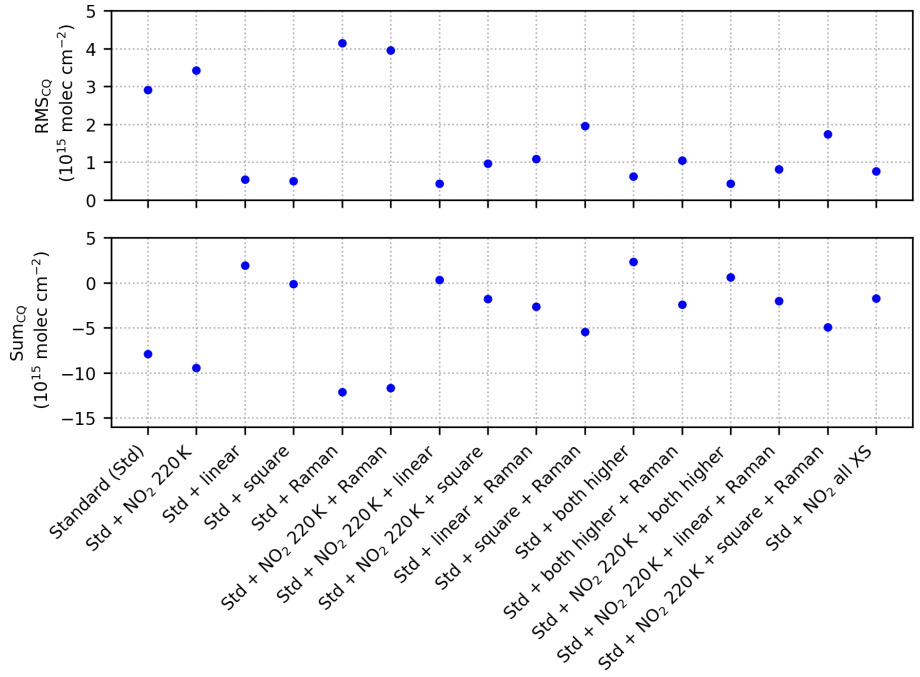


Figure B.2: RMS (upper panel) and sums (lower panel) of the *CQ*-values for all 16 combinations of NO_2 cross-sections for the case of high HCHO, an SZA of 45° and an RAA of 180°.

Figure B.3 depicts an example for the RMS (upper panel) and sums (lower panel) of the absolute deviations from the true reference dSCDs (calculated as described in the main part) for all 16 combinations of O_3 cross-sections for an individual sequence in the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180° (scenario used in the main part). As described in the main part, the need of two O_3 cross-sections at different temperatures becomes evident since the combination not including the cross-section at 243 K yields a larger RMS_{abs} compared to the standard settings. Further, several combinations of O_3

cross-sections yield slightly smaller RMS_{abs}-values than the standard settings indicating that the agreement between simulations and retrievals can be improved by including additional O_3 cross-sections.

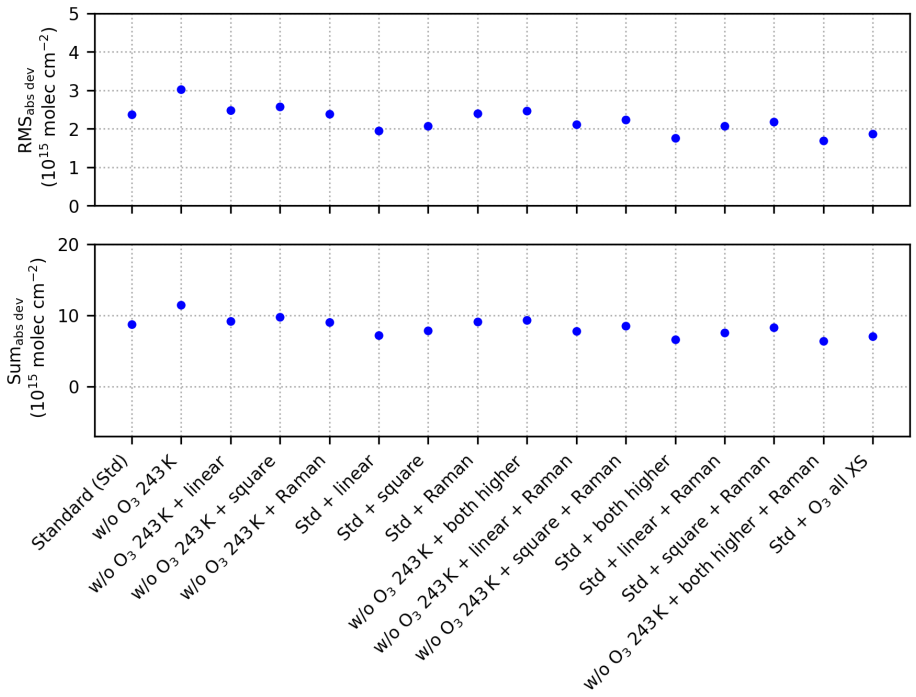


Figure B.3: RMS (upper panel) and sums (lower panel) of the absolute deviations from the true reference dSCDs for all 16 combinations of O_3 cross-sections for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°.

B.2 IMPROVEMENT OF THE DOAS FIT SETTINGS OF FURTHER SPECIES

Figure B.4 provides an example for the glyoxal (and NO_2) fit results before the modifications described in the main part are applied (compare Sect. 7.2.1). Strong residual structures caused by strong water vapour absorptions are clearly visible (compare panels marked in red).

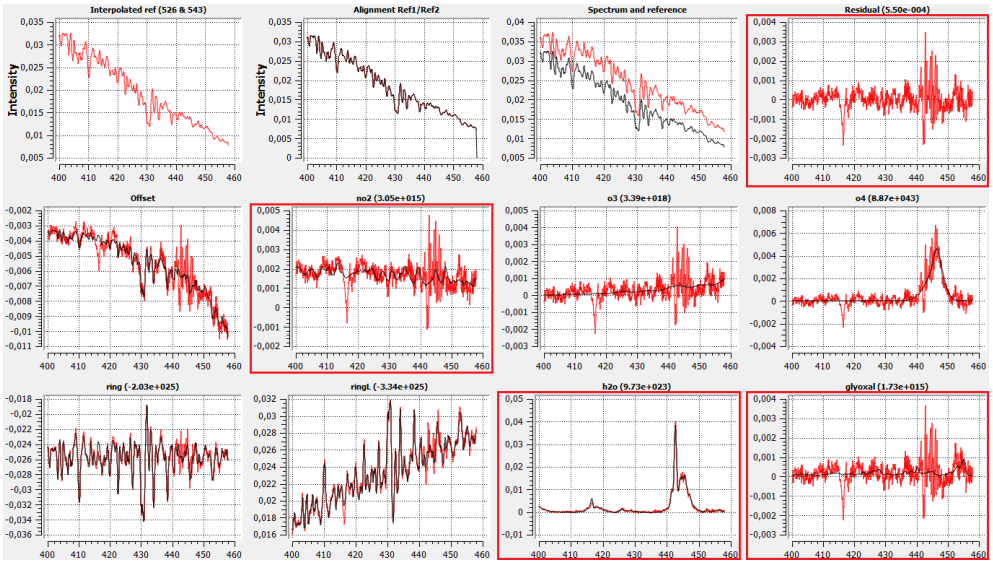


Figure B.4: Fit example for a glyoxal fit for a spectrum recorded by the ATTO 80 m instrument at 2° elevation on 22 December 2017 at 15:30 LT. The different panels show the fits of the trace gas absorption cross-sections included in the fit. Panels marked in red show the fit residual as well as the NO_2, water vapour and glyoxal fits. As described in Sect. 7.2.1 in the main part, strong residual structures caused by strong water vapour absorptions are clearly visible.

Figure B.5 shows time series of glyoxal dSCDs (a), differences (b) and ratios (c) of the retrieved glyoxal dSCDs and the fit RMS (d) for the different fit intervals described in the main part (compare Sect. 7.2.3) for the ATTO 298 m instrument at 30° elevation.

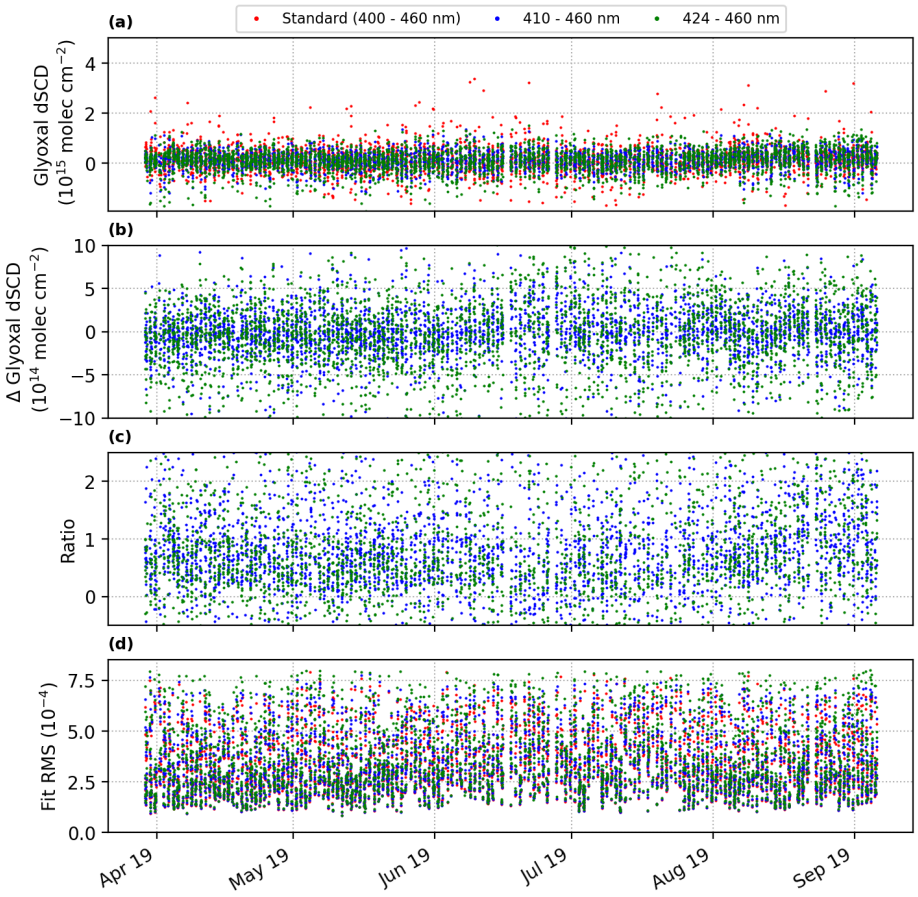


Figure B.5: Time series of glyoxal dSCDs (a), differences (b) and ratios (c) of retrieved glyoxal dSCDs and the fit RMS (d) for the different fit intervals described in the main part for the ATTO 298 m instrument at 30° elevation.

B.3 DATA ANALYSIS

Figure B.6 shows the time series of the (apparent) horizon elevations at Bayfordbury retrieved from the daily horizon scans. The horizon elevations retrieved from the intensities at different wavelengths are indicated by different colours. In addition, the expected horizon elevation (0.52°) and the horizon elevation corresponding to the elevation bias which was finally used for the MAPA retrieval for the Bayfordbury measurements (bias of 1°) are marked by a red and black dashed line, respectively. Further details regarding the analysis and the interpretation of horizon scans as tool to monitor the elevation calibration of a MAX-DOAS instrument are summarised in Donner et al. (2020).

As described in the main part (compare Sect. 9.2.2), the varying (apparent) horizon elevations showing both smooth and abrupt changes indicate too high horizon elevations with values reaching up to ca. 2.25° suggesting that a correction of the reported elevation angles by applying an elevation bias in MAPA is necessary.

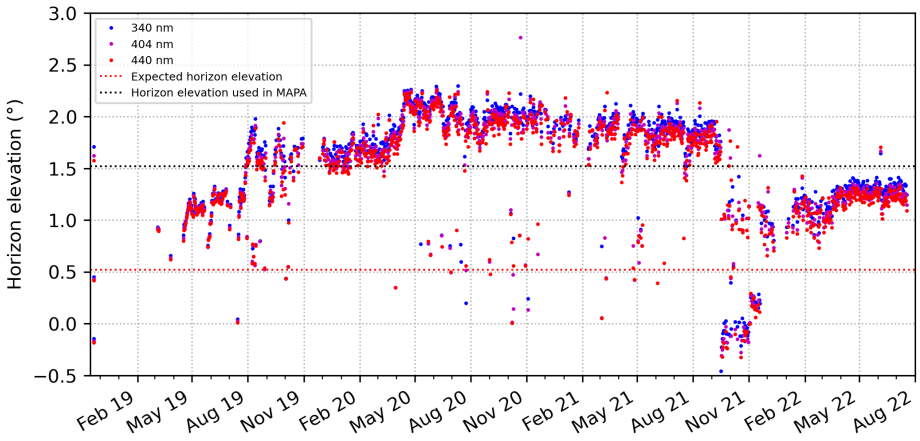


Figure B.6: Time series of the (apparent) horizon elevations at Bayfordbury retrieved from daily horizon scans. Horizon elevations retrieved from the intensities at different wavelengths are indicated by different colours. The expected horizon elevation (0.52°) and the horizon elevation corresponding to the elevation bias which was finally used for the MAPA retrieval for the Bayfordbury measurements (bias of 1°) are marked by a red and black dashed line, respectively.

Figures B.7 to B.9 depict the correlations of the near-surface concentrations/aerosol extinctions and the corresponding VCDs/AODs of glyoxal, NO_2 and aerosol to investigate the information content of the retrieved MAX-DOAS profiles (compare Sect. 9.2.3 in the main part). In each figure, the four panels show individual measurements as well as hourly, daily and monthly means. Note that only valid results for the retrievals using a variable O_4 SF are presented.

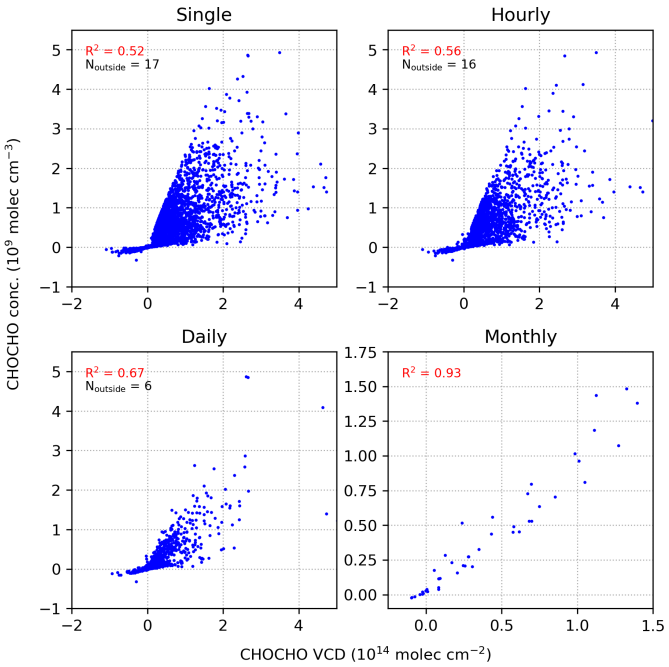


Figure B.7: Correlation of glyoxal near-surface concentrations and corresponding VCDs measured at Bayfordbury. Only valid results for a variable O_4 SF are shown. The four panels present individual measurements as well as hourly, daily and monthly means.

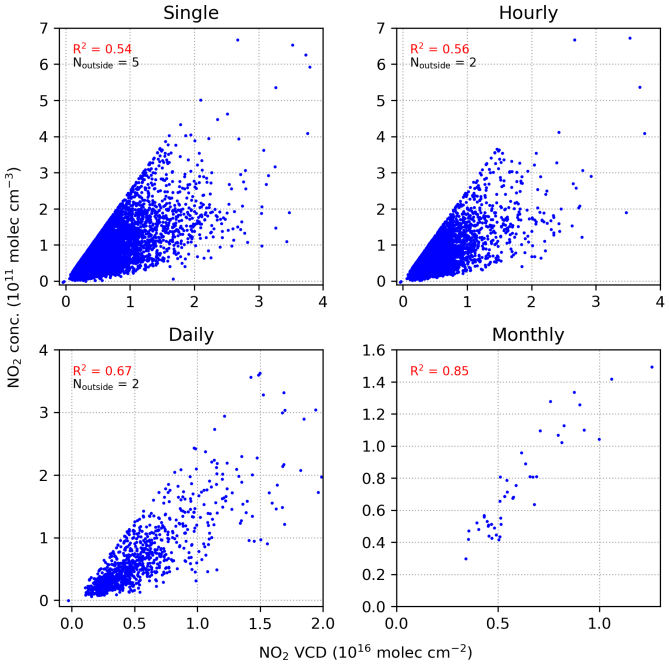


Figure B.8: Correlation of NO_2 near-surface concentrations and corresponding VCDs measured at Bayfordbury. Only valid results for a variable O_4 SF are shown. The four panels present individual measurements as well as hourly, daily and monthly means.

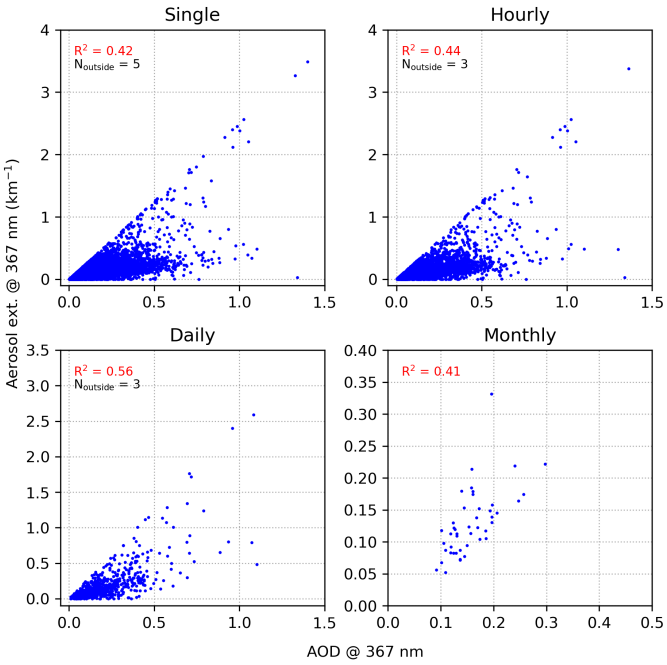


Figure B.9: Correlation of near-surface aerosol extinctions at 367 nm and corresponding AODs measured at Bayfordbury. Only valid results for a variable O_4 SF are shown. The four panels present individual measurements as well as hourly, daily and monthly means.

Figures B.10 to B.13 show comparisons between the AODs and water vapour VCDs retrieved with MAPA and the corresponding AERONET measurements at Mainz, Mohali and ATTO (both instruments). The results obtained with the fixed O_4 SF of 0.8 and the variable SF are depicted in the left and right columns, respectively.

The bad agreement between the AODs retrieved with MAPA and the ones measured by AERONET at Mohali (compare Fig. B.11) can be explained by the fact that no collocated AERONET station is available (compare Sect. 5.2). Hence, AERONET data are taken from the station at the Amity University Gurugram in Haryana located approximately 45 km south-west of New Delhi and around 250 km south of Mohali itself. Nevertheless, reasonable agreement is found for water vapour. Also for the ATTO 298 m instrument rather bad agreement between both data sets is found for both species (compare Fig. B.13) which can be explained by a substantial fraction of the total tropospheric column being located below the altitude of the MAX-DOAS instrument. Reasonable agreements between the MAPA retrievals and the AERONET measurements are found at Mainz and for the ATTO 80 m instrument. For further descriptions, refer to the text in Sect. 9.3 in the main part.

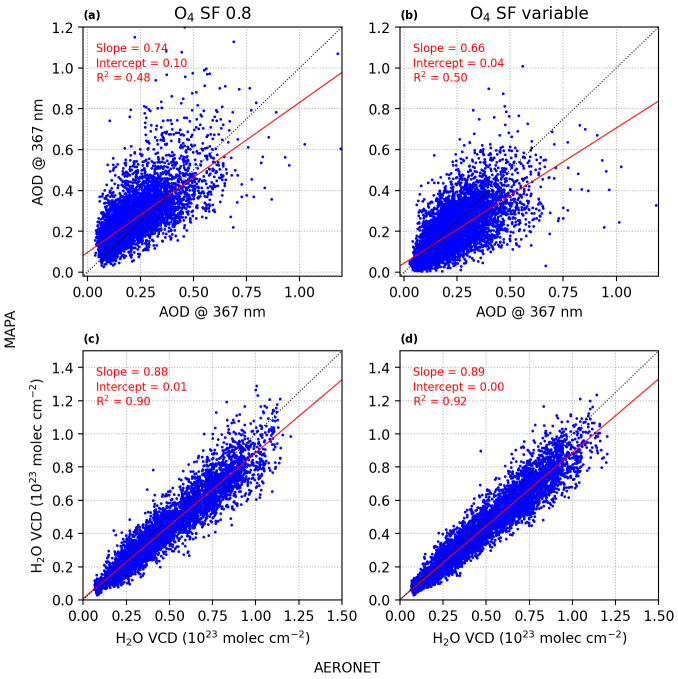


Figure B.10: Correlations between MAPA and AERONET AODs at 367 nm applying fixed (0.8) and variable O_4 scaling factors for the Mainz measurements are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

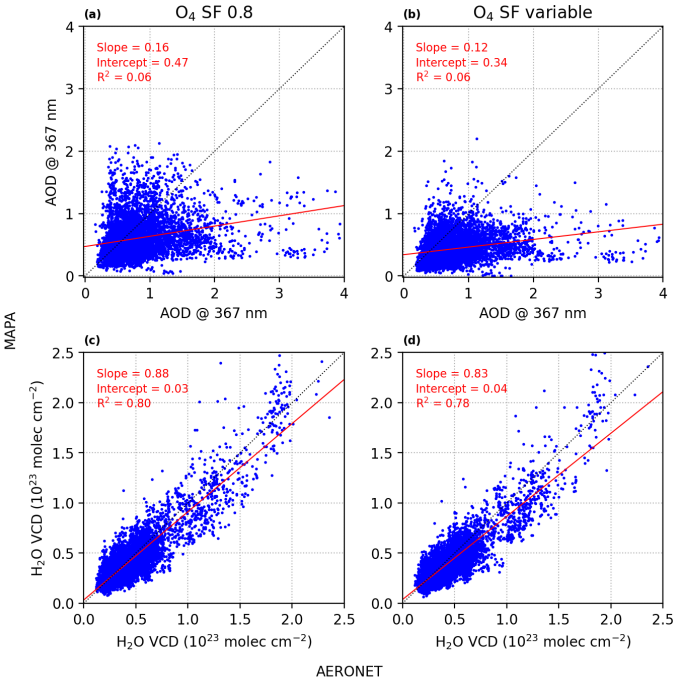


Figure B.11: Correlations between MAPA and AERONET AODs at 367 nm applying fixed (0.8) and variable O_4 scaling factors for the Mohali measurements are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

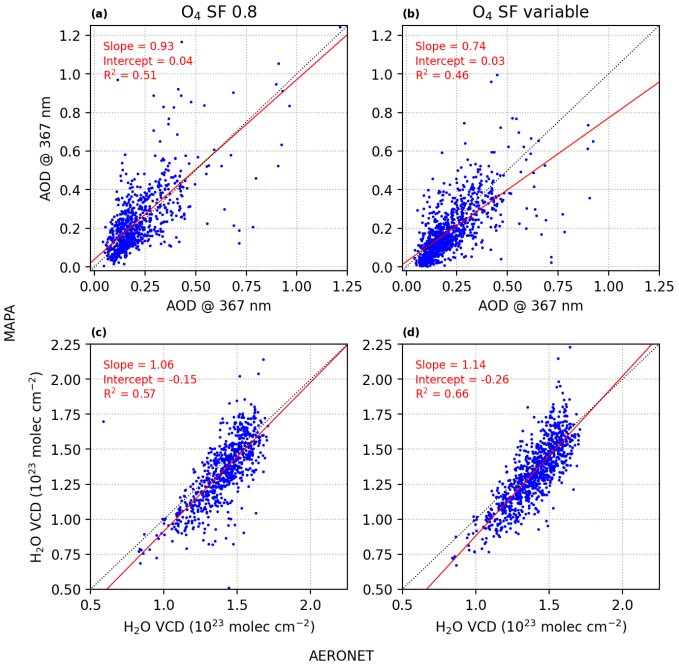


Figure B.12: Correlations between MAPA and AERONET AODs at 367 nm applying fixed (0.8) and variable O_4 scaling factors for the ATTO (80 m) measurements are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

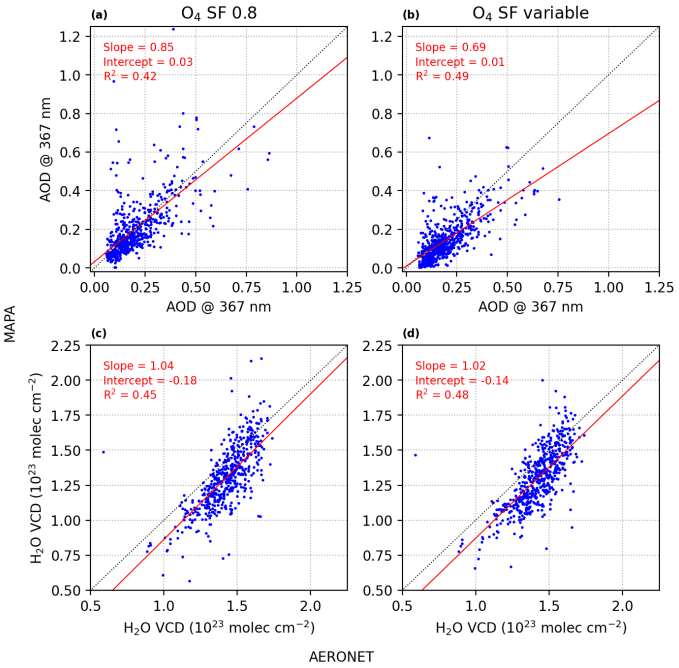


Figure B.13: Correlations between MAPA and AERONET AODs at 367 nm applying fixed (0.8) and variable O_4 scaling factors for the ATTO (298 m) measurements are shown in panels (a) and (b), respectively. Panels (c) and (d) show the same for water vapour VCDs. Results of simple linear regressions are presented in red.

B.4 BAYFORDBURY

Figure B.14 shows seasonal median profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction at 367 nm at Bayfordbury obtained using a fixed O_4 SF of 0.8 in the MAPA retrieval. The profiles of the different seasons are indicated by different colours. A comparison with the profiles obtained with a variable O_4 SF (Fig. 10.20 in Sect. 10.5 in the main part) shows that in principle similar profile shapes and heights are obtained regardless of the applied scaling factor. Nevertheless, a detailed look exhibits that slightly higher near-surface concentrations are obtained for all species using the fixed SF of 0.8, while in the upper parts of the profiles slightly steeper decreases of the concentrations with increasing altitude towards 0 can be observed. This effect is particularly pronounced for NO_2 whereas it does not occur for the aerosol retrieval. For a more detailed discussion, refer to Sect. 10.5 in the main part.

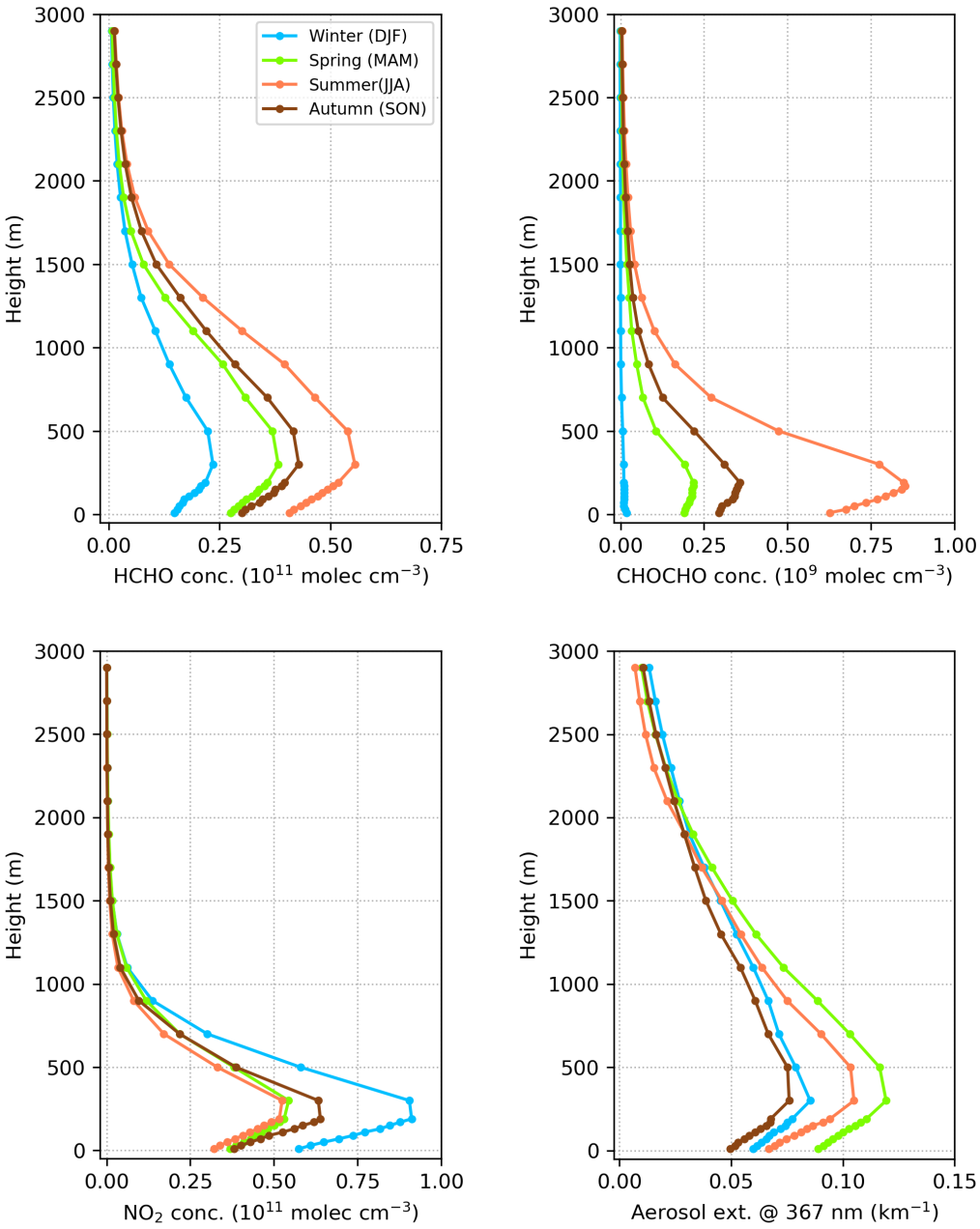


Figure B.14: Seasonal median profiles of formaldehyde (upper left), glyoxal (upper right) and NO_2 (lower left) concentrations as well as aerosol extinction at 367 nm (lower right) obtained for the Bayfordbury measurements applying a fixed O_4 SF of 0.8 in the MAPA retrieval. Profiles for the different seasons are indicated by different colours.

B.5 MOHALI

Figure B.15 presents the annual cycles of the aerosol results at Mohali obtained using a fixed O_4 SF of 0.8 in the MAPA retrieval. Compared to the aerosol results obtained with a variable SF (compare Fig. 12.7 in Sect. 12.2 in the main part), slightly different annual patterns are found for the AODs. In particular, highest AODs are found during the monsoon and in the early post-monsoon season with a fixed SF while their minimum is still found during early summer. Further, especially in the monsoon season the scatter of the results is largely increased for the fixed SF. Contrary to these differences, the annual patterns of the near-surface aerosol extinctions are almost unchanged. A more detailed discussion is given in Sect. 12.2 in the main part.

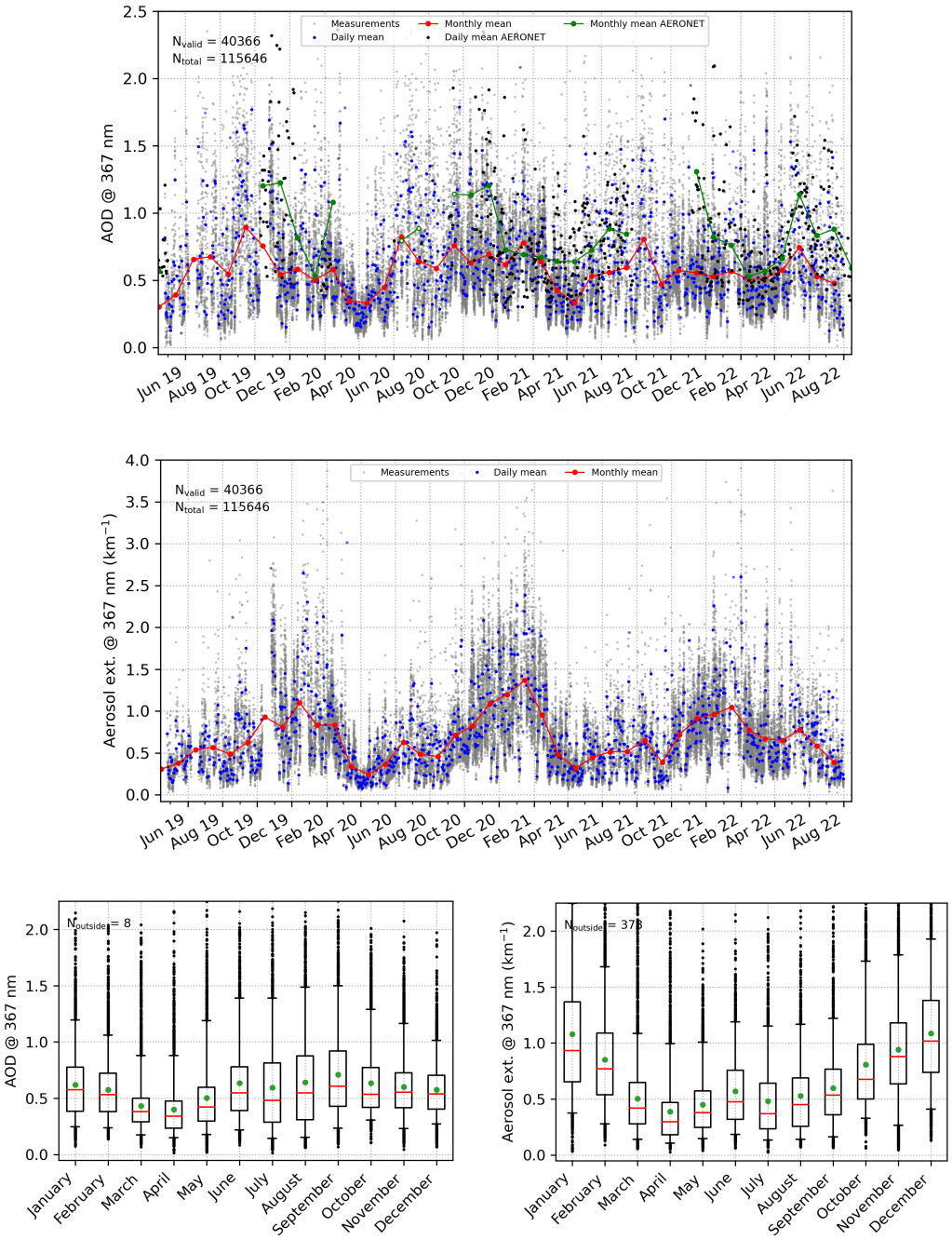


Figure B.15: The two upper panels show time series of AOD and surface aerosol extinction at 367 nm for Mohali obtained with a fixed O_4 SF of 0.8. Individual measurements are displayed in light grey, blue dots mark daily means and red lines and dots indicate monthly means. AERONET daily means are shown in black, while monthly means are marked in green. Open circles label months with less than 60 valid data points. The lower panels depict box-whisker-plots of the annual cycles of AOD (left) and surface aerosol extinction at 367 nm (right). Green dots indicate the means of the respective months, while red lines mark the corresponding medians. 5th and 95th percentiles are marked by the whiskers. Outliers are displayed as black circles.

LIST OF FIGURES

Figure 1.1 Global distribution of HCHO measured by TROPOMI. . . 6
Figure 2.1 Illustration of the NO_x cycle. 14
Figure 2.2 Illustration of methane oxidation and its link to the NO_x cycle. 17
Figure 2.3 Illustration of the HO_x cycle and its link to the NO_x cycle. 19
Figure 3.1 Sketch to illustrate the Beer-Lambert law. 28
Figure 3.2 The basic idea of the classical DOAS approach. 30
Figure 3.3 Viewing geometry of MAX-DOAS measurements. 40
Figure 3.4 Visualisation of the profile parametrisation used in MAPA. 46
Figure 4.1 Overview map of all stations and the track of the research ship cruise. 51
Figure 4.2 Overview map of Mainz. 52
Figure 4.3 Weather statistics for Mainz. 53
Figure 4.4 Telescope unit of the 4-azimuth MAX-DOAS instrument. . 55
Figure 4.5 Image of the four fibre bundles on the CCD. 56
Figure 4.6 Overview map of the Bayfordbury Observatory. 57
Figure 4.7 Weather statistics for Bayfordbury. 58
Figure 4.8 Telescope unit mounted at Bayfordbury. 60
Figure 4.9 Spectrometer unit of the laptop-based TubeMAX-DOAS instrument. 60
Figure 4.10 Spectrometer unit of the RaspberryPi-based TubeMAX-DOAS instrument at ATTO. 61
Figure 4.11 Illustration of a symmetrical Czerny-Turner system. 62
Figure 4.12 Typical offset and dark current spectra of the used AVANTES spectrometers. 63
Figure 4.13 Overview map of Mohali. 65
Figure 4.14 Weather statistics for Mohali. 66
Figure 4.15 Overview map of ATTO. 67
Figure 4.16 Weather statistics for ATTO. 68
Figure 4.17 Overview map of research cruise MSM 68/2. 70
Figure 6.1 Retrieved HCHO dSCDs and fit RMS for a sequence in the case of high NO_2, an SZA of 45° and an RAA of 180° for different ways of calculating the linear Puḳīte term. 89
Figure 6.2 Fit examples for different approaches to calculate the linear Puḳīte term. 90
Figure 6.3 Retrieved HCHO dSCDs and fit RMS for a sequence in the case of high NO_2, an SZA of 45° and an RAA of 180° for different ways of convolving the NO_2 cross-section and its Puḳīte terms. 91
Figure 6.4 Example sequence for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. 94

Figure 6.5 Example sequence for the case of high NO_2, high HCHO, an SZA of 45° and an RAA of 180°. 95
Figure 6.6 Mean RMS and mean sums of the absolute deviations from the true reference dSCDs for all 16 combinations of NO_2 cross-sections in case of low NO_2. 97
Figure 6.7 Mean RMS and mean sums of the absolute deviations from the true reference dSCDs for all 16 combinations of NO_2 cross-sections in case of high NO_2. 98
Figure 6.8 Absolute deviations from reference dSCDs and corresponding *CQ*-values as function of elevation for the case of high NO_2 and HCHO, an SZA of 45° and an RAA of 180°. 100
Figure 6.9 Mean RMS_{CQ} and mean Σ_{CQ} obtained for all 16 combinations of NO_2 cross-sections. 101
Figure 6.10 Higher order absorptions of NO_2: Decision tree 1. 103
Figure 6.11 Higher order absorptions of NO_2: Decision tree 2. 105
Figure 6.12 Time series of retrieved HCHO dSCDs and fit RMS including different sets of NO_2 cross-sections in the DOAS fit on 29 March 2019 at 3° elevation at Bayfordbury. 107
Figure 6.13 Time series of retrieved HCHO dSCDs and fit RMS including different sets of NO_2 cross-sections in the DOAS fit on 12 November 2019 at 3° elevation during the MSM 68/2 ship cruise. 109
Figure 6.14 Time series of retrieved HCHO dSCDs with and without including additional NO_2 cross-sections in the DOAS fit and corresponding fit RMS for the year 2020 at 3° elevation at ATTO. 111
Figure 6.15 Example sequence for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. 114
Figure 6.16 Example sequence for the case of low NO_2, high HCHO, an SZA of 60° and a RAA of 180°. 116
Figure 6.17 Mean RMS and mean sums of the absolute deviations from the true reference dSCDs for all 16 combinations of O_3 cross-sections over all scenarios. 117
Figure 6.18 Time series of retrieved HCHO dSCDs and fit RMS including different sets of O_3 cross-sections in the DOAS fit on 29 March 2019 at 10° elevation at Bayfordbury. 119
Figure 6.19 Time series of retrieved HCHO dSCDs and fit RMS including different sets of O_3 cross-sections in the DOAS fit on 12 November 2017 at 10° elevation during the MSM68/2 ship cruise. 122
Figure 6.20 Mean fit RMS at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the boundaries of the tested fitting intervals. 127
Figure 6.21 Mean fit RMS at 1° elevation for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017. 128

Figure 6.22 Mean fit error at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the boundaries of the tested fitting intervals. 130
Figure 6.23 Mean fit error at 1° elevation for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017. 131
Figure 6.24 Standard deviation of the 100 retrieved HCHO dSCDs at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as function of the boundaries of the tested fitting intervals. 132
Figure 6.25 Standard deviation of the HCHO dSCDs at 1° elevation retrieved for spectra measured during the MSM 68/2 ship cruise between 11 a.m and 3 p.m. (UT) on 12 November 2017. 133
Figure 6.26 Relative and absolute deviations of the retrieved HCHO dSCDs from the true reference values at 1° elevation for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as functions of the boundaries of the tested fitting intervals. 135
Figure 6.27 RMS of the absolute deviations from the true reference dSCDs and corresponding standard deviations for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180° as functions of the boundaries of the tested fitting intervals. 137
Figure 6.28 Mean RMS_{abs} and corresponding standard deviations as functions of the boundaries of the tested fitting intervals. . 139
Figure 6.29 Example sequence for the influence of the Ring effect for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. 144
Figure 6.30 Example sequence for the effects of intensity weighted convolutions for HCHO, BrO and O_4 on the HCHO results for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. 147
Figure 6.31 Example sequence for the effects of further Puḳīte terms of NO_2, O_3 and O_4 on the HCHO results for the case of low NO_2, high HCHO, an SZA of 60° and an RAA of 180°. . . 150
Figure 6.32 Comparison of AMFs obtained with different accuracy and precision settings for the case of high NO_2, an SZA of 45° and an RAA of 180° at 10° elevation. 154
Figure 6.33 Example sequence for the effect of the wavelength dependence of the AMFs on the comparison results for the case of low NO_2, high HCHO, an SZA of 45° and an RAA of 180°. 156
Figure 6.34 Same as Fig. 6.33 but for the case of high AOD. 158

Figure 6.35 Logarithmic ratios of spectra obtained with improved simulations settings and standard spectrum for the case of high HCHO and NO_2, low AOD, an SZA of 60° and an RAA of 30° at 3° elevation. 160
Figure 6.36 Example sequence for the effects of enhanced accuracy and precision settings for the simulations of the synthetic spectra on the HCHO results for the case of high HCHO and NO_2, an SZA of 60° and an RAA of 30°. 161
Figure 6.37 Example sequence for the HCHO results obtained for synthetic spectra simulated at low spectral resolution for the case of high HCHO and NO_2, an SZA of 60° and an RAA of 30°. 164
Figure 6.38 Same as Fig. 6.37 but for spectra without Ring effect. . . . 166
Figure 6.39 Example sequence for the HCHO results obtained with and without intensity offset as well as spectral shift and stretch included in the fit for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. 171
Figure 6.40 Fit residuals and fitted HCHO cross-section for the 1° spectrum of the example sequence shown in Fig. 6.39. . . . 172
Figure 6.41 Example sequence for the HCHO results obtained with different polynomial degrees used in the fit for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. 175
Figure 6.42 Normalised absorption cross-sections of HCHO and BrO. 177
Figure 6.43 Example sequence for the HCHO results obtained with and without BrO cross-section included in the fit for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. 178
Figure 6.44 Example sequence for the HCHO results obtained with the fit settings suggested in this work and used at IUP-Bremen and BIRA for the case of high HCHO, low NO_2, an SZA of 45° and an RAA of 180°. 181
Figure 6.45 Same as Fig. 6.44 but for the case of high NO_2. 183
Figure 6.46 Same as Fig. 6.44 but for spectra without Ring effect. . . . 184
Figure 7.1 Residuals and fitted H_2O cross-section of an O_4 fit for a spectrum recorded by the ATTO 80 m instrument. 194
Figure 7.2 Time series of O_4 dSCDs, differences and ratios of the retrieved O_4 dSCDs and the fit RMS for different fit scenarios for the ATTO 80 m instrument at 3° elevation. 195
Figure 7.3 Time series of O_4 dSCDs, differences and ratios of the retrieved O_4 dSCDs and the fit RMS for different fit scenarios for the Bayfordbury instrument at 3° elevation. 197
Figure 7.4 Correlation plots comparing water vapour dSCDs retrieved in the UV and visible spectral ranges. 198
Figure 7.5 Residual as well as NO_2 and glyoxal fits for a spectrum recorded by the ATTO 80 m instrument. 200

Figure 7.6 Residual and glyoxal fit for a spectrum recorded by the ATTO 298 m instrument. . . . 201
Figure 7.7 Difference of H_2O cross-sections convolved with Gaussian and Super Gaussian SLFs. . . . 202
Figure 7.8 Measured SLF and corresponding fitted parametrisation. . 203
Figure 7.9 Partial derivatives with respect to the different parameters of the parametrised SLF. . . . 204
Figure 7.10 Pseudo absorbers for variations of the slit function parameters. . . . 205
Figure 7.11 Fit residuals without and with PSAs for SLF changes included in the DOAS fit. . . . 205
Figure 7.12 Residual spectra, mean residual spectrum (PSA) and residual spectrum for example fit. . . . 206
Figure 7.13 Residuals and glyoxal fits for a spectrum recorded by the Bayfordbury instrument. . . . 207
Figure 7.14 Time series of glyoxal dSCDs, differences and ratios of the retrieved glyoxal dSCDs and the fit RMS for different fit scenarios for the ATTO 298 m instrument at 3° elevation. . 208
Figure 7.15 Absorption cross-sections of glyoxal and water vapour between 360 and 460 nm. . . . 209
Figure 7.16 Time series of glyoxal dSCDs, differences and ratios of the retrieved glyoxal dSCDs and the fit RMS for two different fit intervals for the ATTO 298 m instrument at 3° elevation. 210
Figure 7.17 Examples for the residual structure at around 430 nm for the ATTO 298 m instrument and solar spectrum in this spectral range. . . . 212
Figure 7.18 Absorption cross-section of glyoxal and tested gaps between 410 and 460 nm. . . . 213
Figure 7.19 Time series of glyoxal dSCDs, differences and ratios of the retrieved glyoxal dSCDs and the fit RMS obtained with different gaps included in the fit for the ATTO 298 m instrument at 3° elevation. . . . 214
Figure 7.20 Correlation between glyoxal dSCDs retrieved with different gap scenarios and the standard scenario. . . . 215
Figure 8.1 Comparison of Kurucz-Fit results for Gaussian and Super Gaussian SLFs. . . . 218
Figure 8.2 Comparison of NO_2 results for Gaussian and Super Gaussian SLFs. . . . 219
Figure 8.3 Comparison of the spectral calibrations for the three temperature regimes at Bayfordbury. . . . 220
Figure 8.4 Time series of spectral pre-shifts and shifts of spectra of HCHO and NO_2 fits for Bayfordbury. . . . 222
Figure 8.5 Time series of HCHO results obtained with correct and wrong calibrations. . . . 223
Figure 9.1 Fit examples for HCHO, glyoxal, NO_2, O_4 and HONO using the standard fit settings. . . . 231

Figure 9.2 Influence of an elevation bias on the AODs retrieved with MAPA. . . . 235
Figure 9.3 Histograms of O_4 SFs obtained with and without elevation bias. . . . 236
Figure 9.4 Correlations between MAPA AODs and water vapour VCDs and their corresponding AERONET data. . . . 237
Figure 9.5 Diurnal evolution of water vapour concentration profiles on 23 April 2020. . . . 238
Figure 9.6 Mean water vapour concentration profiles of ERA5 and MAPA retrievals. . . . 239
Figure 9.7 Mean formaldehyde and NO_2 concentration profiles for April 2020. . . . 240
Figure 9.8 Correlation of HCHO near-surface concentrations and corresponding VCDs at Bayfordbury. . . . 243
Figure 9.9 Histograms of variable O_4 SFs obtained for the five instruments. . . . 246
Figure 9.10 Correlation between MAPA AODs and water vapour VCDs and their corresponding AERONET data for both applied SFs at Bayfordbury. . . . 248
Figure 9.11 Correlations between ERA5 and locally measured temperatures and wind speeds at Bayfordbury. . . . 249
Figure 9.12 Angular differences between ERA5 and locally measured wind directions. . . . 250
Figure 10.1 Time series and annual cycles of HCHO VCDs and surface concentrations for Bayfordbury. . . . 255
Figure 10.2 Diurnal cycles of HCHO VCDs and surface concentrations for Bayfordbury. . . . 256
Figure 10.3 Time series and annual cycles of glyoxal VCDs and surface concentrations for Bayfordbury. . . . 258
Figure 10.4 Diurnal cycles of glyoxal VCDs and surface concentrations for Bayfordbury. . . . 259
Figure 10.5 Time series and annual cycles of NO_2 VCDs and surface concentrations for Bayfordbury. . . . 260
Figure 10.6 Diurnal cycles of NO_2 VCDs and surface concentrations for Bayfordbury. . . . 261
Figure 10.7 Time series and annual cycles of AOD and surface aerosol extinction at 367 nm for Bayfordbury. . . . 263
Figure 10.8 Diurnal cycles of AOD and surface aerosol extinction at 367 nm for Bayfordbury. . . . 264
Figure 10.9 Temperature and wind dependence of HCHO VCDs and surface concentrations at Bayfordbury. . . . 266
Figure 10.10 Temperature and wind dependence of glyoxal VCDs and surface concentrations at Bayfordbury. . . . 269
Figure 10.11 Temperature and wind dependence of NO_2 VCDs and surface concentrations at Bayfordbury. . . . 272

Figure 10.12 Temperature and wind dependence of AODs and surface aerosol extinctions at 367 nm at Bayfordbury. 274
Figure 10.13 Comparison of glyoxal and formaldehyde VCDs and surface concentrations at Bayfordbury. 276
Figure 10.14 Comparison of formaldehyde and NO_2 VCDs and surface concentrations at Bayfordbury. 277
Figure 10.15 Comparison of glyoxal and NO_2 VCDs and surface concentrations at Bayfordbury. 278
Figure 10.16 Comparison of formaldehyde and aerosol results at Bayfordbury. 279
Figure 10.17 Comparison of glyoxal and aerosol results at Bayfordbury. 280
Figure 10.18 Comparison of NO_2 and aerosol results at Bayfordbury. . 281
Figure 10.19 Annual cycles of R_{GF} for VCDs and surface concentrations at Bayfordbury. 283
Figure 10.20 Seasonal median profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction at 367 nm for Bayfordbury. 284
Figure 10.21 Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Bayfordbury. 287
Figure 11.1 Time series and annual cycles of HCHO VCDs and surface concentrations for Mainz. 294
Figure 11.2 Diurnal cycles of HCHO VCDs and surface concentrations for Mainz. 295
Figure 11.3 Time series and annual cycles of glyoxal VCDs and surface concentrations for Mainz. 297
Figure 11.4 Diurnal cycles of glyoxal VCDs and surface concentrations for Mainz. 298
Figure 11.5 Time series and annual cycles of NO_2 VCDs and surface concentrations for Mainz. 299
Figure 11.6 Diurnal cycles of NO_2 VCDs and surface concentrations for Mainz. 300
Figure 11.7 Time series and annual cycles of AOD and surface aerosol extinction at 367 nm for Mainz. 302
Figure 11.8 Diurnal cycles of AOD and surface aerosol extinction at 367 nm for Mainz. 303
Figure 11.9 Temperature and wind dependence of HCHO VCDs and surface concentrations at Mainz. 304
Figure 11.10 Temperature and wind dependence of NO_2 VCDs and surface concentrations at Mainz. 307
Figure 11.11 Temperature and wind dependence of AODs and surface aerosol extinctions at 367 nm at Mainz. 310
Figure 11.12 Comparison of formaldehyde and NO_2 VCDs and surface concentrations at Mainz. 312
Figure 11.13 Comparison of formaldehyde and aerosol results at Mainz. 313
Figure 11.14 Comparison of NO_2 and aerosol results at Mainz. 314

Figure 11.15 Seasonal median profiles of formaldehyde and NO_2 concentrations as well as aerosol extinction at 367 nm for Mainz. 315
Figure 11.16 Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Mainz. . . . 318
Figure 11.17 Annual cycles of HONO VCDs and surface concentrations for Mainz. . . . 320
Figure 11.18 Diurnal cycles of HONO VCDs and surface concentrations for Mainz. . . . 321
Figure 11.19 Temperature and wind dependence of HONO VCDs and surface concentrations at Mainz. . . . 322
Figure 11.20 Seasonal median profiles of HONO concentrations for Mainz. 324
Figure 12.1 Time series and annual cycles of HCHO VCDs and surface concentrations for Mohali. . . . 327
Figure 12.2 Diurnal cycles of HCHO VCDs and surface concentrations for Mohali. . . . 329
Figure 12.3 Time series and annual cycles of glyoxal VCDs and surface concentrations for Mohali. . . . 332
Figure 12.4 Diurnal cycles of glyoxal VCDs and surface concentrations for Mohali. . . . 333
Figure 12.5 Time series and annual cycles of NO_2 VCDs and surface concentrations for Mohali. . . . 335
Figure 12.6 Diurnal cycles of NO_2 VCDs and surface concentrations for Mohali. . . . 336
Figure 12.7 Time series and annual cycles of AOD and surface aerosol extinction at 367 nm for Mohali. . . . 339
Figure 12.8 Diurnal cycles of AOD and surface aerosol extinction at 367 nm for Mohali. . . . 341
Figure 12.9 Diurnal cycles of AOD and surface aerosol extinction at 367 nm for Mohali using a fixed SF of 0.8. . . . 342
Figure 12.10 Temperature and wind dependence of HCHO VCDs and surface concentrations at Mohali. . . . 344
Figure 12.11 Temperature and wind dependence of glyoxal VCDs and surface concentrations at Mohali. . . . 346
Figure 12.12 Temperature and wind dependence of NO_2 VCDs and surface concentrations at Mohali. . . . 349
Figure 12.13 Temperature and wind dependence of AODs and surface aerosol extinctions at 367 nm at Mohali. . . . 351
Figure 12.14 Comparison of glyoxal and formaldehyde VCDs and surface concentrations at Mohali. . . . 354
Figure 12.15 Comparison of formaldehyde and NO_2 VCDs and surface concentrations at Mohali. . . . 355
Figure 12.16 Comparison of glyoxal and NO_2 VCDs and surface concentrations at Mohali. . . . 356
Figure 12.17 Comparison of formaldehyde and aerosol results at Mohali. 357
Figure 12.18 Comparison of glyoxal and aerosol results at Mohali. . . . 358
Figure 12.19 Comparison of NO_2 and aerosol results at Mohali. . . . 359

Figure 12.20 Annual cycles of R_{GF} for VCDs and surface concentrations at Mohali. 361
Figure 12.21 Seasonal median profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction at 367 nm for Mohali. 362
Figure 12.22 Comparison of MAX-DOAS and TROPOMI HCHO VCDs at Mohali. 366
Figure 12.23 Annual cycles of HONO VCDs and surface concentrations for Mohali. 369
Figure 12.24 Diurnal cycles of HONO VCDs and surface concentrations for Mohali. 370
Figure 12.25 Temperature and wind dependence of HONO VCDs and surface concentrations at Mohali. 371
Figure 12.26 Seasonal median profiles of HONO concentrations for Mohali. 373
Figure 13.1 Time series and annual cycles of HCHO VCDs and concentrations at instrument altitude for ATTO (80 m). 378
Figure 13.2 Time series and annual cycles of HCHO VCDs and concentrations at instrument altitude for ATTO (298 m). 379
Figure 13.3 Diurnal cycles of HCHO VCDs and concentrations at instrument altitude for ATTO. 381
Figure 13.4 Time series and annual cycles of glyoxal VCDs and concentrations at instrument altitude for ATTO (80 m). 383
Figure 13.5 Time series and annual cycles of glyoxal VCDs and concentrations at instrument altitude for ATTO (298 m). 384
Figure 13.6 Diurnal cycles of glyoxal VCDs and concentrations at instrument altitude for ATTO. 386
Figure 13.7 Time series and annual cycles of NO_2 VCDs and concentrations at instrument altitude for ATTO (80 m). 388
Figure 13.8 Time series and annual cycles of NO_2 VCDs and concentrations at instrument altitude for ATTO (298 m). 389
Figure 13.9 Diurnal cycles of NO_2 VCDs and concentrations at instrument altitude for ATTO. 391
Figure 13.10 Time series and annual cycles of AOD and aerosol extinction at 367 nm at instrument altitude for ATTO (80 m). . . 393
Figure 13.11 Time series and annual cycles of AOD and aerosol extinction at instrument altitude at 367 nm for ATTO (298 m). . . 394
Figure 13.12 Diurnal cycles of AOD and aerosol extinction at instrument altitude at 367 nm for ATTO. 396
Figure 13.13 Diurnal cycles of AOD and aerosol extinction at instrument altitude at 367 nm for ATTO using a fixed SF of 0.8. 397
Figure 13.14 Temperature and wind dependence of HCHO VCDs and concentrations at instrument altitude at ATTO (80 m). . . . 399
Figure 13.15 Temperature and wind dependence of HCHO VCDs and concentrations at instrument altitude at ATTO (298 m). . . 400

Figure 13.16 Temperature and wind dependence of glyoxal VCDs and concentrations at instrument altitude at ATTO (80 m). . . . 403
Figure 13.17 Temperature and wind dependence of glyoxal VCDs and concentrations at instrument altitude at ATTO (298 m). . . 404
Figure 13.18 Temperature and wind dependence of NO_2 VCDs and concentrations at instrument altitude at ATTO (80 m). . . . 406
Figure 13.19 Temperature and wind dependence of NO_2 VCDs and concentrations at instrument altitude at ATTO (298 m). . . 407
Figure 13.20 Temperature and wind dependence of AODs and aerosol extinctions at instrument altitude at 367 nm at ATTO (80 m).409
Figure 13.21 Temperature and wind dependence of AODs and aerosol extinctions at instrument altitude at 367 nm at ATTO (298 m).410
Figure 13.22 Comparison of glyoxal and formaldehyde VCDs and concentrations at instrument altitude at ATTO (80 m). 412
Figure 13.23 Comparison of glyoxal and formaldehyde VCDs and concentrations at instrument altitude at ATTO (298 m). 413
Figure 13.24 Comparison of formaldehyde and NO_2 VCDs and concentrations at instrument altitude at ATTO (80 m). 414
Figure 13.25 Comparison of formaldehyde and NO_2 VCDs and concentrations at instrument altitude at ATTO (298 m). 415
Figure 13.26 Comparison of glyoxal and NO_2 VCDs and concentrations at instrument altitude at ATTO (80 m). 416
Figure 13.27 Comparison of glyoxal and NO_2 VCDs and concentrations at instrument altitude at ATTO (298 m). 417
Figure 13.28 Comparison of formaldehyde and aerosol results at ATTO (80 m). 418
Figure 13.29 Comparison of formaldehyde and aerosol results at ATTO (298 m). 419
Figure 13.30 Comparison of glyoxal and aerosol results at ATTO (80 m). 420
Figure 13.31 Comparison of glyoxal and aerosol results at ATTO (298 m).421
Figure 13.32 Comparison of NO_2 and aerosol results at ATTO (80 m). . 422
Figure 13.33 Comparison of NO_2 and aerosol results at ATTO (298 m). . 423
Figure 13.34 Annual cycles of R_{GF} for VCDs and concentrations at instrument altitude at ATTO. 424
Figure 13.35 Seasonal median profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction at 367 nm for ATTO (80 m). 426
Figure 13.36 Seasonal median profiles of formaldehyde, glyoxal and NO_2 concentrations as well as aerosol extinction at 367 nm for ATTO (298 m). 427
Figure 13.37 Comparison of MAX-DOAS and TROPOMI HCHO VCDs at ATTO (80 m). 431
Figure 13.38 Comparison of MAX-DOAS and TROPOMI HCHO VCDs at ATTO (298 m). 432
Figure 13.39 Annual cycles of H_2O VCDs and concentrations at instrument altitude for ATTO (80 m). 435

Figure 13.40 Temperature dependence of H_2O VCDs and concentrations at instrument altitude at ATTO (80 m). 435
Figure 13.41 Seasonal median profiles of H_2O concentrations for ATTO (80 m). 437
Figure 13.42 Comparison of the formaldehyde results of both ATTO instruments. 438
Figure 13.43 Comparison of the glyoxal results of both ATTO instruments. 439
Figure 13.44 Comparison of the NO_2 results of both ATTO instruments. 440
Figure 13.45 Comparison of the aerosol results of both ATTO instruments. 442

LIST OF TABLES

Table 4.1 Overview of the spectral characteristics of the spectrometers used at the different stations. 62
Table 6.1 Overview of the scenarios for the synthetic spectra. 82
Table 6.2 List of reference trace gas absorption cross-sections and solar reference spectrum used for the simulations. 83
Table 6.3 Overview of the HCHO standard fit settings (sensitivity study). 86
Table 6.4 Comparison of the HCHO DOAS fit settings suggested in this work with the ones used by the DOAS groups at IUP-Bremen and BIRA. 180
Table 9.1 Overview of the HCHO standard fit settings. 228
Table 9.2 Overview of the glyoxal standard fit settings. 228
Table 9.3 Overview of the NO_2 standard fit settings. 229
Table 9.4 Overview of the O_4 standard fit settings. 229
Table 9.5 Overview of the HONO standard fit settings. 230
Table 9.6 Overview of wavelengths and aerosol scenarios used for MAPA. 233
Table 9.7 Overview of applied elevation biases and lowest elevation angles used in the MAPA analyses. 241
Table 9.8 Fractions of valid aerosol, HCHO and NO_2 retrievals for fixed and variable SFs. 245
Table 9.9 Overview of mean O_4 SFs for the five instruments. 246
Table 10.1 Comparison results between daily TROPOMI and MAX-DOAS data at Bayfordbury. 288
Table 11.1 Overview of the spectral characteristics of the Mainz instrument. 292
Table 11.2 Comparison results between daily TROPOMI and MAX-DOAS data at Mainz. 319
Table 12.1 Comparison results between daily TROPOMI and MAX-DOAS data at Mohali. 368
Table 13.1 Overview of the spectral gaps introduced to the DOAS fits at ATTO. 376
Table 13.2 Overview of representative monthly mean R_{GF}-values at ATTO. 424
Table 13.3 Comparison results between daily TROPOMI and MAX-DOAS data at ATTO (80 m). 433
Table 13.4 Comparison results between daily TROPOMI and MAX-DOAS data at ATTO (298 m). 434

AVANTES BV (2023). *Optical Spectrometers: an introduction*. https://www.avantes.com/support/theoretical-background/introduction-to-spectrometers/. Accessed: 2023-08-08.

Aliwell, S. R., Van Roozendael, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P., Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J.-C., Pfeilsticker, K. and Pundt, I. (2002). 'Analysis for BrO in zenith-sky spectra: An intercomparison exercise for analysis improvement'. In: *Journal of Geophysical Research: Atmospheres* 107.D14, ACH 10–1–ACH 10–20. DOI: 10.1029/2001JD000329.

Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Kalisz Hedegaard, A. B., Schneising, O. and Burrows, J. P. (2020). 'Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus Sentinel-5 Precursor satellite during the 2018 Canadian wildfires'. In: *Atmospheric Chemistry and Physics* 20.4, pp. 2057–2072. DOI: 10.5194/acp-20-2057-2020.

Alvarado, L. (2020). unpublished personal communication. IUP-Bremen.

Andreae, M. O. (2019). 'Emission of trace gases and aerosols from biomass burning – an updated assessment'. In: *Atmospheric Chemistry and Physics* 19.13, pp. 8523–8546. DOI: 10.5194/acp-19-8523-2019.

Andreae, M. O., Acevedo, O. C., Araùjo, A., Artaxo, P., Barbosa, C. G. G., Barbosa, H. M. J., Brito, J., Carbone, S., Chi, X., Cintra, B. B. L., Silva, N. F. da, Dias, N. L., Dias-Júnior, C. Q., Ditas, F., Ditz, R., Godoi, A. F. L., Godoi, R. H. M., Heimann, M., Hoffmann, T., Kesselmeier, J., Könemann, T., Krüger, M. L., Lavric, J. V., Manzi, A. O., Lopes, A. P., Martins, D. L., Mikhailov, E. F., Moran-Zuloaga, D., Nelson, B. W., Nölscher, A. C., Santos Nogueira, D., Piedade, M. T. F., Pöhlker, C., Pöschl, U., Quesada, C. A., Rizzo, L. V., Ro, C.-U., Ruckteschler, N., Sá, L. D. A., Oliveira Sá, M. de, Sales, C. B., Santos, R. M. N. dos, Saturno, J., Schöngart, J., Sörgel, M., Souza, C. M. de, Souza, R. A. F. de, Su, H., Targhetta, N., Tóta, J., Trebs, I., Trumbore, S., Eijck, A. van, Walter, D., Wang, Z., Weber, B., Williams, J., Winderlich, J., Wittmann, F., Wolff, S. and Serrano, A. M. Yáñez (2015). 'The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols'. In: *Atmospheric Chemistry and Physics* 15.18, pp. 10723–10776. DOI: 10.5194/acp-15-10723-2015.

Arbeitsgemeinschaft Deutscher Verkehrsflughäfen (June 2023). *Arbeitsgemeinschaft Deutscher Verkehrsflughäfen*. https://de.statista.com/statistik/daten/studie/5852/umfrage/entwicklung-der-flugzeugbewegungen-am-flughafen-frankfurt/. Accessed: 2023-07-25.

Atkinson, R. (2000). 'Atmospheric chemistry of VOCs and NO_x'. In: *Atmospheric Environment* 34.12, pp. 2063–2101. DOI: 10.1016/S1352-2310(99)00460-4.

Barkley, M. P., De Smedt, I., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E. and Mao, J. (2013). 'Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns'. In: *Journal of Geophysical Research: Atmospheres* 118.12, pp. 6849–6868. DOI: 10.1002/jgrd.50552.

Bassett, M. and Seinfeld, J.H. (1983). 'Atmospheric equilibrium model of sulfate and nitrate aerosols'. In: *Atmospheric Environment (1967)* 17.11, pp. 2237–2252. DOI: 10.1016/0004-6981(83)90221-4.

Behrens, L. K., Hilboll, A., Richter, A., Peters, E., Alvarado, L. M. A., Kalisz Hedegaard, A. B., Wittrock, F., Burrows, J. P. and Vrekoussis, M. (2019). 'Detection of outflow of formaldehyde and glyoxal from the African continent to the Atlantic Ocean with a MAX-DOAS instrument'. In: *Atmospheric Chemistry and Physics* 19.15, pp. 10257–10278. DOI: 10.5194/acp-19-10257-2019.

Beirle, S. (2020). *Impact of an instrument tilt on MAX-DOAS profile retrievals*. unpublished seminar presentation. MPIC.

Beirle, S., Borger, C., Dörner, S., Kumar, V. and Wagner, T. (2022). 'Calculating the vertical column density of O_4 during daytime from surface values of pressure, temperature, and relative humidity'. In: *Atmospheric Measurement Techniques* 15.4, pp. 987–1006. DOI: 10.5194/amt-15-987-2022.

Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y. and Wagner, T. (2019). 'The Mainz profile algorithm (MAPA)'. In: *Atmospheric Measurement Techniques* 12.3, pp. 1785–1806. DOI: 10.5194/amt-12-1785-2019.

Beirle, S., Lampel, J., Lerot, C., Sihler, H. and Wagner, T. (2017). 'Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives'. In: *Atmospheric Measurement Techniques* 10.2, pp. 581–598. DOI: 10.5194/amt-10-581-2017.

Benavent, N., Garcia-Nieto, D., Wang, S. and Saiz-Lopez, A. (2019). 'MAX-DOAS measurements and vertical profiles of glyoxal and formaldehyde in Madrid, Spain'. In: *Atmospheric Environment* 199, pp. 357–367. DOI: 10.1016/j.atmosenv.2018.11.047.

Bobrowski, N., Hönninger, G., Galle, B. and Platt, U. (2003). 'Detection of bromine monoxide in a volcanic plume'. In: *Nature* 423, pp. 273–276. DOI: 10.1038/nature01625.

Borger, C., Beirle, S., Dörner, S., Sihler, H. and Wagner, T. (2020). 'Total column water vapour retrieval from S-5P/TROPOMI in the visible blue spectral range'. In: *Atmospheric Measurement Techniques* 13.5, pp. 2751–2783. DOI: 10.5194/amt-13-2751-2020.

Borovski, A. N., Dzhola, A. V., Grechko, E. I., Postylyakov, O. V., Ivanov, V. A., Kanaya, Y., and Obukhov, A. M. (2015). 'Measurements of formaldehyde total content in the troposphere using DOAS technique in Moscow region'. In: *PROCEEDINGS OF SPIE*. Ed. by G. G. Matvienko and O. A. Romanovskii. Vol. 9680. doi: 10.1117/12.2205933. 21st International symposium on atmospheric and ocean optics. SPIE.

Bussemer, M. (1993). 'Der Ring-Effekt: Ursachen und Einfluß auf die Messung stratosphärischer Spurenstoffe'. Diploma thesis. University of Heidelberg.

Calvert, J. G., Atkinson, R., Becker, K. H., Kamens, R. M., Seinfeld, J. H., Wallington, T. J. and Yarwood, G. (2002). *The Mechanisms of Atmospheric Oxidation of Aromatic Hydrocarbons*. Oxford University Press.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J. and Chance, K. (2014). 'Glyoxal retrieval from the Ozone Monitoring Instrument'. In: *Atmospheric Measurement Techniques* 7.11, pp. 3891–3907. DOI: 10.5194/amt-7-3891-2014.

Chan Miller, C., Jacob, D. J., González Abdad, G. and Chance, K. (2016). 'Hotspot of glyoxal over the Pearl River delta seen from the OMI satellite instrument: implications for emissions of aromatic hydrocarbons'. In: *Atmospheric Chemistry and Physics* 16.7, pp. 4631–4639. DOI: 10.5194/acp-16-4631-2016.

Chan Miller, C., Jacob, D. J., Marais, E. A., Yu, K., Travis, K. R., Kim, P. S., Fisher, J. A., Zhu, L., Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Kaiser, J., Min, K.-E., Brown, S. S., Washenfelder, R. A., González Abad, G. and Chance, K. (2017). 'Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data'. In: *Atmospheric Chemistry and Physics* 17.14, pp. 8725–8738. DOI: 10.5194/acp-17-8725-2017.

Chance, K. and Kurucz, R.L. (2010). 'An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 111.9. Special Issue Dedicated to Laurence S. Rothman on the Occasion of his 70th Birthday., pp. 1289–1295. DOI: 10.1016/j.jqsrt.2010.01.036.

Chance, K., Palmer, P. I., Spurr, R. J. D., Martin, R. V., Kurosu, T. P. and Jacob, D. J. (2000). 'Satellite observations of formaldehyde over North America from GOME'. In: *Geophysical Research Letters* 27.21, pp. 3461–3464. DOI: 10.1029/2000GL011857.

City Population (July 2023). *United Kingdom*. https://www.citypopulation.de/en/uk/eastofengland/. Accessed: 2023-07-19.

Clémer, K., Van Roozendael, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Spurr, R., Wang, P. and De Mazière, M. (2010). 'Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing'. In: *Atmospheric Measurement Techniques* 3.4, pp. 863–878. DOI: 10.5194/amt-3-863-2010.

Danckaert, T., Fayt, C., Van Roozendael, M., De Smedt, I., Letocart, V., Merlaud, A. and Pinardi, G. (2017). *QDOAS – Software user manual*. BIRA-IASB. Brussels, Belgium.

De Smedt, I. (2021). unpublished personal communication. BIRA.

De Smedt, I. (2023). unpublished personal communication. BIRA.

De Smedt, I., Müller, J.-F., Stavrakou, T., A, R. van der, Eskes, H. and Van Roozendael, M. (2008). 'Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors'. In: *Atmospheric Chemistry and Physics* 8.16, pp. 4947–4963. DOI: 10.5194/acp-8-4947-2008.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H. and Van Roozendael, M. (2021). 'Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements'. In: *Atmospheric Chemistry and Physics* 21.16, pp. 12561–12593. DOI: 10.5194/acp-21-12561-2021.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F. and Van Roozendael, M. (2015). 'Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations'. In: *Atmospheric Chemistry and Physics* 15.21, pp. 12519–12545. DOI: 10.5194/acp-15-12519-2015.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., Geffen, J. van, Boersma, K. F. and Veefkind, P. (2018). 'Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project'. In: *Atmospheric Measurement Techniques* 11.4, pp. 2395–2426. DOI: 10.5194/amt-11-2395-2018.

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., Valks, P., Hao, N. and A, R. van der (2012). 'Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues'. In: *Atmospheric Measurement Techniques* 5.11, pp. 2933–2949. DOI: 10.5194/amt-5-2933-2012.

Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., Platt, U., Prados-Román, C., Puķīte, J., Wagner, T., Werner, B. and Pfeilsticker, K. (2011). 'The Monte Carlo atmospheric radiative transfer model McArtim: Introduction and validation of Jacobians and 3D features'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 112.6, pp. 1119–1137. DOI: 10.1016/j.jqsrt.2010.12.009.

DiGangi, J. P., Henry, S. B., Kammrath, A., Boyle, E. S., Kaser, L., Schnitzhofer, R., Graus, M., Turnipseed, A., Park, J.-H., Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin III, R. L., Kim, S., Nakashima, Y., Wolfe, G. M., Kajii, Y., Apel, E. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A. and Keutsch, F. N. (2012). 'Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry'. In: *Atmospheric Chemistry and Physics* 12.20, pp. 9529–9543. DOI: 10.5194/acp-12-9529-2012.

Dimitropoulou, E., Hendrick, F., Pinardi, G., Friedrich, M. M., Merlaud, A., Tack, F., De Longueville, H., Fayt, C., Hermans, C., Laffineur, Q., Fierens, F.

and Van Roozendael, M. (2020). 'Validation of TROPOMI tropospheric NO_2 columns using dual-scan multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements in Uccle, Brussels'. In: *Atmospheric Measurement Techniques* 13.10, pp. 5165–5191. DOI: 10.5194/amt-13-5165-2020.

Donner, S. (2014). 'Aufbau und Charakterisierung eines MAX-DOAS-Instruments'. Bachelor thesis. Johannes Gutenberg University Mainz.

Donner, S. (2016). 'Mobile MAX-DOAS measurements of the tropospheric formaldehyde column in the Rhein-Main region'. Master thesis. Johannes Gutenberg University Mainz. URL: http://hdl.handle.net/11858/00-001M-0000-002C-EB17-2.

Donner, S., Kuhn, J., Van Roozendael, M., Bais, A., Beirle, S., Bösch, T., Bognar, K., Bruchkouski, I., Chan, K. L., Dörner, S., Drosoglou, T., Fayt, C., Frieß, U., Hendrick, F., Hermans, C., Jin, J., Li, A., Ma, J., Peters, E., Pinardi, G., Richter, A., Schreier, S. F., Seyler, A., Strong, K., Tirpitz, J.-L., Wang, Y., Xie, P., Xu, J., Zhao, X. and Wagner, T. (2020). 'Evaluating different methods for elevation calibration of MAX-DOAS (Multi AXis Differential Optical Absorption Spectroscopy) instruments during the CINDI-2 campaign'. In: *Atmospheric Measurement Techniques* 13.2, pp. 685–712. DOI: 10.5194/amt-13-685-2020.

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J. and Crawford, J. H. (2010). 'Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation'. In: *Atmospheric Environment* 44.18, pp. 2213–2223. DOI: 10.1016/j.atmosenv.2010.03.010.

ECMWF (Feb. 2024). *ERA5: data documentation*. https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation. Accessed: 2024-02-14.

ESA (2020). *TROPOMI Level 2 Formaldehyde Total Column products – Version 02*. https://doi.org/10.5270/S5P-vg1i7t0. Accessed: 2024-02-15.

ESA (2023a). *S5P Mission Performance Centre – Formaldehyde [L2__HCHO___] Readme*. https://sentinels.copernicus.eu/documents/247904/3541451/Sentinel-5P-Formaldehyde-Readme.pdf. Accessed: 2024-02-15.

ESA (2023b). *S5P Mission Performance Centre – Level 1b Readme*. https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Level-1b-Product-Readme-File. Accessed: 2024-02-15.

Feierabend, K. J., Zhu, L., Talukdar, R. K. and Burkholder, J. B. (2008). 'Rate Coefficients for the OH + HC(O)C(O)H (Glyoxal) Reaction between 210 and 390 K'. In: *The Journal of Physical Chemistry A* 112.1, pp. 73–82. DOI: 10.1021/jp0768571.

Fiedler, S. (2018). *Expedition to the North Atlantic with RV MARIA S. MERIAN*. Tech. rep. Hamburg, Germany: Max-Planck-Institut für Meteorologie. DOI: 10.17617/2.3006588.

Finlayson-Pitts, B.J. and Pitts Jr., J.N. (1999). *Chemistry of the upper and lower atmosphere: theory, experiments, and applications*. Academic press.

Fleischmann, C. O., Hartmann, M., Burrows, J. P. and Orphal, J. (2004). 'New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy'. In: *Journal of Photochemistry and Photobiology A: Chemistry* 168.1, pp. 117–132. DOI: 10.1016/j.jphotochem.2004.03.026.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K. and Fried, A. (2012). 'The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system'. In: *Atmospheric Chemistry and Physics* 12.15, pp. 6699–6721. DOI: 10.5194/acp-12-6699-2012.

Frieß, U., Monks, P. S., Remedios, J. J., Rozanov, A., Sinreich, R., Wagner, T. and Platt, U. (2006). 'MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies'. In: *Journal of Geophysical Research: Atmospheres* 111.D14. DOI: 10.1029/2005JD006618.

Fu, T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M. and Henze, D. K. (2008). 'Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols'. In: *Journal of Geophysical Research: Atmospheres* 113.D15. DOI: 10.1029/2007JD009505.

Goldstein, A. H. and Galbally, I. E. (2007). 'Known and Unexplored Organic Constituents in the Earth's Atmosphere'. In: *Environmental Science & Technology* 41.5. PMID: 17396635, pp. 1514–1521. DOI: 10.1021/es072476p.

Gomes Alves, E., Aquino Santana, R., Quaresma Dias-Júnior, C., Botía, S., Taylor, T., Serrano, A. M. Yáñez, Kesselmeier, J., Bourtsoukidis, E., Williams, J., Assis, P. I. Lembo Silveira de, Martins, G., Souza, R. de, Duvoisin Júnior, S., Guenther, A., Gu, D., Tsokankunku, A., Sörgel, M., Nelson, B., Pinto, D., Komiya, S., Martins Rosa, D., Weber, B., Barbosa, C., Robin, M., Feeley, K. J., Duque, A., Lemos, V. Londoño, Contreras, M. P., Idarraga, A., López, N., Husby, C., Jestrow, B. and Cely Toro, I. M. (2023). 'Intra- and interannual changes in isoprene emission from central Amazonia'. In: *Atmospheric Chemistry and Physics* 23.14, pp. 8149–8168. DOI: 10.5194/acp-23-8149-2023.

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P. and Suleiman, R. (2015). 'Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval'. In: *Atmospheric Measurement Techniques* 8.1, pp. 19–32. DOI: 10.5194/amt-8-19-2015.

Gordon, I.E., Rothman, L.S., Hargreaves, R.J., Hashemi, R., Karlovets, E.V., Skinner, F.M., Conway, E.K., Hill, C., Kochanov, R.V., Tan, Y., Wcisło, P., Finenko, A.A., Nelson, K., Bernath, P.F., Birk, M., Boudon, V., Campargue, A., Chance, K.V., Coustenis, A., Drouin, B.J., Flaud, J.–M., Gamache, R.R., Hodges, J.T., Jacquemart, D., Mlawer, E.J., Nikitin, A.V., Perevalov, V.I., Rotger, M., Tennyson, J., Toon, G.C., Tran, H., Tyuterev, V.G., Adkins, E.M., Baker, A., Barbe, A., Canè, E., Császár, A.G., Dudaryonok, A., Egorov, O., Fleisher, A.J., Fleurbaey, H., Foltynowicz, A., Furtenbacher, T., Harrison, J.J., Hartmann, J.–M., Horneman, V.–M., Huang, X., Karman, T., Karns, J., Kassi, S., Kleiner, I., Kofman, V., Kwabia–Tchana, F., Lavrentieva, N.N., Lee, T.J., Long, D.A., Lukashevskaya, A.A.,

Lyulin, O.M., Makhnev, V.Yu., Matt, W., Massie, S.T., Melosso, M., Mikhailenko, S.N., Mondelain, D., Müller, H.S.P., Naumenko, O.V., Perrin, A., Polyansky, O.L., Raddaoui, E., Raston, P.L., Reed, Z.D., Rey, M., Richard, C., Tóbiás, R., Sadiek, I., Schwenke, D.W., Starikova, E., Sung, K., Tamassia, F., Tashkun, S.A., Vander Auwera, J., Vasilenko, I.A., Vigasin, A.A., Villanueva, G.L., Vispoel, B., Wagner, G., Yachmenev, A. and Yurchenko, S.N. (2022). 'The HITRAN2020 molecular spectroscopic database'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 277, p. 107949. DOI: 10.1016/j.jqsrt.2021.107949.

Grainger, J. F. and Ring, J. (1962). 'Anomalous Fraunhofer Line Profiles'. In: *Nature* 193.762. DOI: 10.1038/193762a0.

Greenblatt, G. D., Orlando, J. J., Burkholder, J. B. and Ravishankara, A. R. (1990). 'Absorption measurements of oxygen between 330 and 1140 nm'. In: *Journal of Geophysical Research: Atmospheres* 95.D11, pp. 18577–18582. DOI: 10.1029/JD095iD11p18577.

Haagen-Smit, A. J. (1952). 'Chemistry and Physiology of Los Angeles Smog'. In: *Industrial & Engineering Chemistry* 44.6, pp. 1342–1346. DOI: 10.1021/ie50510a045.

Hak, C., Pundt, I., Trick, S., Kern, C., Platt, U., Dommen, J., Ordóñez, C., Prévôt, A. S. H., Junkermann, W., Astorga-Lloréns, C., Larsen, B. R., Mellqvist, J., Strandberg, A., Yu, Y., Galle, B., Kleffmann, J., Lörzer, J. C., Braathen, G. O. and Volkamer, R. (2005). 'Intercomparison of four different in-situ techniques for ambient formaldehyde measurements in urban air'. In: *Atmospheric Chemistry and Physics* 5.11, pp. 2881–2900. DOI: 10.5194/acp-5-2881-2005.

Harder, J. W., Jakoubek, R. O. and Mount, G. H. (1997). 'Measurement of tropospheric trace gases by long-path differential absorption spectroscopy during the 1993 OH Photochemistry Experiment'. In: *Journal of Geophysical Research: Atmospheres* 102.D5, pp. 6215–6226. DOI: 10.1029/96JD03608.

Heckel, A., Richter, A., Tarsu, T., Wittrock, F., Hak, C., Pundt, I., Junkermann, W. and Burrows, J. P. (2005). 'MAX-DOAS measurements of formaldehyde in the Po-Valley'. In: *Atmospheric Chemistry and Physics* 5.4, pp. 909–918. DOI: 10.5194/acp-5-909-2005.

Hendrick, F., Müller, J.-F., Clémer, K., Wang, P., De Mazière, M., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G., Stavrakou, T., Vlemmix, T. and Van Roozendael, M. (2014). 'Four years of ground-based MAX-DOAS observations of HONO and NO_2 in the Beijing area'. In: *Atmospheric Chemistry and Physics* 14.2, pp. 765–781. DOI: 10.5194/acp-14-765-2014.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M. and Abuhassan, N. (2009). 'NO_2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation'. In: *Journal of Geophysical Research: Atmospheres* 114.D13. DOI: 10.1029/2009JD011848.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D. and Thépaut, J-N. (2023). 'ERA5 hourly data on single levels from 1940

to present'. In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. Accessed: 2024-02-12. DOI: 10.24381/cds.adbb2d47.

Holben, B. N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y.J., Nakajima, T., Lavenu, F., Jankowiak, I. and Smirnov, A. (1998). 'AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization'. In: *Remote Sensing of Environment* 66.1, pp. 1–16. DOI: 10.1016/S0034-4257(98)00031-5.

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. Vande, Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K. and Zibordi, G. (2001). 'An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET'. In: *Journal of Geophysical Research: Atmospheres* 106.D11, pp. 12067–12097. DOI: 10.1029/2001JD900014.

Hong, Qianqian, Liu, Cheng, Hu, Qihou, Zhang, Yanli, Xing, Chengzhi, Ou, Jinping, Tan, Wei, Liu, Haoran, Huang, Xiaoqing and Wu, Zhenfeng (2022). 'Vertical distribution and temporal evolution of formaldehyde and glyoxal derived from MAX-DOAS observations: The indicative role of VOC sources'. In: *Journal of Environmental Sciences* 122, pp. 92–104. DOI: 10.1016/j.jes.2021.09.025.

Hönninger, G., Friedeburg, C. von and Platt, U. (2004). 'Multi axis differential optical absorption spectroscopy (MAX-DOAS)'. In: *Atmospheric Chemistry and Physics* 4.1, pp. 231–254. DOI: 10.5194/acp-4-231-2004.

Hönninger, G. and Platt, U. (2002). 'Observations of BrO and its vertical distribution during surface ozone depletion at Alert'. In: *Atmospheric Environment* 36.15. Air/Snow/Ice Interactions in the Arctic: Results from ALERT 2000 and SUMMIT 2000, pp. 2481–2489. DOI: 10.1016/S1352-2310(02)00104-8.

Horbanski, Martin (2015). 'Emissions and Distribution of Reactive Iodine from Seaweed in Coastal Regions - Investigations using new mobile and in-situ DOAS techniques'. Dissertation. Ruprecht-Karls-Universität Heidelberg.

Hulst, H. C. van de (1981). 'Light scattering by small particles'. In: *Courier Corporation*.

Ibrahim, O., Shaiganfar, R., Sinreich, R., Stein, T., Platt, U. and Wagner, T. (2010). 'Car MAX-DOAS measurements around entire cities: quantification of NO_x emissions from the cities of Mannheim and Ludwigshafen (Germany)'. In: *Atmospheric Measurement Techniques* 3.3, pp. 709–721. DOI: 10.5194/amt-3-709-2010.

Iglesias-Suarez, F., Kinnison, D. E., Rap, A., Maycock, A. C., Wild, O. and Young, P. J. (2018). 'Key drivers of ozone change and its radiative forcing over the 21st century'. In: *Atmospheric Chemistry and Physics* 18.9, pp. 6121–6139. DOI: 10.5194/acp-18-6121-2018.

Indian Census (2011). *Population Census 2011*. https://www.census2011.co.in/. Accessed: 2023-07-26.

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H. and Courrèges-Lacoste, G. B. (2012). 'Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and -5p'. In: *Remote Sensing of Environment* 120. The Sentinel Missions - New Opportunities for Science, pp. 58–69. DOI: 10.1016/j.rse.2012.01.023.

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., Zhang, L., Zhang, X., Fried, A., Crounse, J. D., St. Clair, J. M. and Wisthaler, A. (2018). 'High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US'. In: *Atmospheric Chemistry and Physics* 18.8, pp. 5483–5497. DOI: 10.5194/acp-18-5483-2018.

Kaiser, J., Wolfe, G. M., Min, K. E., Brown, S. S., Miller, C. C., Jacob, D. J., deGouw, J. A., Graus, M., Hanisco, T. F., Holloway, J., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Washenfelder, R. A. and Keutsch, F. N. (2015). 'Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation'. In: *Atmospheric Chemistry and Physics* 15.13, pp. 7571–7583. DOI: 10.5194/acp-15-7571-2015.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Chan Miller, C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., González Abad, G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K.-J., Cho, A., Yoon, J., Kim, S.-K., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B. R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T. and Choi, Y. (2020). 'New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS)'. In: *Bulletin of the American Meteorological Society* 101.1, E1 –E22. DOI: 10.1175/BAMS-D-18-0013.1.

Kluge, F., Hüneke, T., Knecht, M., Lichtenstern, M., Rotermund, M., Schlager, H., Schreiner, B. and Pfeilsticker, K. (2020). 'Profiling of formaldehyde, glyoxal, methylglyoxal, and CO over the Amazon: normalized excess mixing ratios and related emission factors in biomass burning plumes'. In: *Atmospheric Chemistry and Physics* 20.20, pp. 12363–12389. DOI: 10.5194/acp-20-12363-2020.

Koppmann, R. and Wildt, J. (2007). 'Oxygenated Volatile Organic Compounds'. In: *Volatile Organic Compounds in the Atmosphere*. Ed. by R. Koppmann. Oxford: Blackwell Publishing, pp. 129–172. ISBN: 978-1-4051-3115-5.

Kraus, S. (2006). 'DOASIS – A Framework Design for DOAS'. PhD thesis. University of Mannheim.

Krautwurst, Sven (2010). 'Charakterisierung eines neu aufgebauten MAX-DOAS-Systems und Interpretation von ersten Messergebnissen zu dem Spurenstoff NO_2'. Diploma thesis. Fachhochschule Coburg.

Kreher, K., Johnston, P. V., Wood, S. W., Nardi, B. and Platt, U. (1997). 'Ground-based measurements of tropospheric and stratospheric BrO at Arrival Heights, Antarctica'. In: *Geophysical Research Letters* 24.23, pp. 3021–3024. DOI: 10.1029/97GL02997.

Kreher, K., Van Roozendael, M., Hendrick, F., Apituley, A., Dimitropoulou, E., Frieß, U., Richter, A., Wagner, T., Lampel, J., Abuhassan, N., Ang, L., Anguas, M., Bais, A., Benavent, N., Bösch, T., Bognar, K., Borovski, A., Bruchkouski, I., Cede, A., Chan, K. L., Donner, S., Drosoglou, T., Fayt, C., Finkenzeller, H., Garcia-Nieto, D., Gielen, C., Gómez-Martín, L., Hao, N., Henzing, B., Herman, J. R., Hermans, C., Hoque, S., Irie, H., Jin, J., Johnston, P., Khayyam Butt, J., Khokhar, F., Koenig, T. K., Kuhn, J., Kumar, V., Liu, C., Ma, J., Merlaud, A., Mishra, A. K., Müller, M., Navarro-Comas, M., Ostendorf, M., Pazmino, A., Peters, E., Pinardi, G., Pinharanda, M., Piters, A., Platt, U., Postylyakov, O., Prados-Roman, C., Puentedura, O., Querel, R., Saiz-Lopez, A., Schönhardt, A., Schreier, S. F., Seyler, A., Sinha, V., Spinei, E., Strong, K., Tack, F., Tian, X., Tiefengraber, M., Tirpitz, J.-L., Gent, J. van, Volkamer, R., Vrekoussis, M., Wang, S., Wang, Z., Wenig, M., Wittrock, F., Xie, P. H., Xu, J., Yela, M., Zhang, C. and Zhao, X. (2020). 'Intercomparison of NO_2, O_4, O_3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2'. In: *Atmospheric Measurement Techniques* 13.5, pp. 2169–2208. DOI: 10.5194/amt-13-2169-2020.

Kumar, V., Beirle, S., Dörner, S., Mishra, A. K., Donner, S., Wang, Y., Sinha, V. and Wagner, T. (2020). 'Long-term MAX-DOAS measurements of NO_2, HCHO, and aerosols and evaluation of corresponding satellite data products over Mohali in the Indo-Gangetic Plain'. In: *Atmospheric Chemistry and Physics* 20.22, pp. 14183–14235. DOI: 10.5194/acp-20-14183-2020.

Kumar, V., Sarkar, C. and Sinha, V. (2016). 'Influence of post-harvest crop residue fires on surface ozone mixing ratios in the N.W. IGP analyzed using 2 years of continuous in situ trace gas measurements'. In: *Journal of Geophysical Research: Atmospheres* 121.7, pp. 3619–3633. DOI: 10.1002/2015JD024308.

Kurucz, R. L., Furenlid, I., Brault, J. and Testermann, L. (1984). 'Solar flux atlas from 296 to 1300 nm'. In: *National Solar Observatory, Sunspot, New Mexico*. URL: http://kurucz.harvard.edu/sun/fluxatlas/.

Lampel, J., Pöhler, D., Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., Tennyson, J., Lodi, L., Frieß, U., Wang, Y., Beirle, S., Platt, U. and Wagner, T. (2017a). 'Detection of water vapour absorption around 363 nm in measured atmospheric absorption spectra and its effect on DOAS evaluations'. In: *Atmospheric Chemistry and Physics* 17.2, pp. 1271–1295. DOI: 10.5194/acp-17-1271-2017.

Lampel, J., Pöhler, D., Tschritter, J., Frieß, U. and Platt, U. (2015). 'On the relative absorption strengths of water vapour in the blue wavelength range'. In: *Atmo-*

spheric Measurement Techniques 8.10, pp. 4329–4346. DOI: 10.5194/amt-8-4329-2015.

Lampel, J., Wang, Y., Hilboll, A., Beirle, S., Sihler, H., Puķīte, J., Platt, U. and Wagner, T. (2017b). 'The tilt effect in DOAS observations'. In: *Atmospheric Measurement Techniques* 10.12, pp. 4819–4831. DOI: 10.5194/amt-10-4819-2017.

Landeshauptstadt Mainz (June 2023). *Statistische Informationen zur Stadtentwicklung*. https://www.mainz.de/medien/internet/downloads/statistik/Einwohner_nach_Stadtteilen_am_30_04_2023.pdf. Accessed: 2023-07-25.

Lange, K., Richter, A., Schönhardt, A., Meier, A. C., Bösch, T., Seyler, A., Krause, K., Behrens, L. K., Wittrock, F., Merlaud, A., Tack, F., Fayt, C., Friedrich, M. M., Dimitropoulou, E., Van Roozendael, M., Kumar, V., Donner, S., Dörner, S., Lauster, B., Razi, M., Borger, C., Uhlmannsiek, K., Wagner, T., Ruhtz, T., Eskes, H., Bohn, B., Santana Diaz, D., Abuhassan, N., Schüttemeyer, D. and Burrows, J. P. (2023). 'Validation of Sentinel-5P TROPOMI tropospheric NO_2 products by comparison with NO_2 measurements from airborne imaging DOAS, ground-based stationary DOAS, and mobile car DOAS measurements during the S5P-VAL-DE-Ruhr campaign'. In: *Atmospheric Measurement Techniques* 16.5, pp. 1357–1389. DOI: 10.5194/amt-16-1357-2023.

Lauster, B., Dörner, S., Beirle, S., Donner, S., Gromov, S., Uhlmannsiek, K. and Wagner, T. (2021). 'Estimating real driving emissions from multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements at the A60 motorway near Mainz, Germany'. In: *Atmospheric Measurement Techniques* 14.1, pp. 769–783. DOI: 10.5194/amt-14-769-2021.

Lauster, B., Uhlmannsiek, K., Donner, S., Dörner, S. and Wagner, T. (2019). 'Characterisation of the newly developed Tube MAX-DOAS instrument.' In: *Geophysical Research Abstracts*. Vol. 21.

Lawson, S. J., Selleck, P. W., Galbally, I. E., Keywood, M. D., Harvey, M. J., Lerot, C., Helmig, D. and Ristovski, Z. (2015). 'Seasonal in situ observations of glyoxal and methylglyoxal over the temperate oceans of the Southern Hemisphere'. In: *Atmospheric Chemistry and Physics* 15.1, pp. 223–240. DOI: 10.5194/acp-15-223-2015.

Leighton, P. (1961). *Photochemistry of Air Pollution*. Oxford: Elsevier Science.

Lelieveld, J. and Dorland, R. van (1995). 'Ozone Chemistry Changes in the Troposphere and Consequent Radiative Forcing of Climate'. In: *Atmospheric Ozone as a Climate Gas*. Ed. by W.-C. Wang and I. S. A. Isaksen. Berlin, Heidelberg: Springer, pp. 227–258. ISBN: 978-3-642-79869-6.

Lerot, C., Hendrick, F., Van Roozendael, M., Alvarado, L. M. A., Richter, A., De Smedt, I., Theys, N., Vlietinck, J., Yu, H., Van Gent, J., Stavrakou, T., Müller, J.-F., Valks, P., Loyola, D., Irie, H., Kumar, V., Wagner, T., Schreier, S. F., Sinha, V., Wang, T., Wang, P. and Retscher, C. (2021). 'Glyoxal tropospheric column retrievals from TROPOMI – multi-satellite intercomparison and ground-based validation'. In: *Atmospheric Measurement Techniques* 14.12, pp. 7775–7807. DOI: 10.5194/amt-14-7775-2021.

Lerot, C., Stavrakou, T., De Smedt, I., Müller, J.-F. and Van Roozendael, M. (2010). 'Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model'. In: *Atmospheric Chemistry and Physics* 10.24, pp. 12059–12072. DOI: 10.5194/acp-10-12059-2010.

Levelt, P.F., Oord, G.H.J. van den, Dobber, M.R., Malkki, A., Visser, Huib, Vries, Johan de, Stammes, P., Lundell, J.O.V. and Saari, H. (2006). 'The ozone monitoring instrument'. In: *IEEE Transactions on Geoscience and Remote Sensing* 44.5, pp. 1093–1101. DOI: 10.1109/TGRS.2006.872333.

Levenberg, K. (1944). 'A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES'. In: *Quarterly of Applied Mathematics* 2.2, pp. 164–168. URL: http://www.jstor.org/stable/43633451 (visited on 22/02/2024).

Levitt, S.B. and Chock, D.P. (1976). 'Weekday-Weekend Pollutant Studies of the Los Angeles Basin'. In: *Journal of the Air Pollution Control Association* 26.11, pp. 1091–1092. DOI: 10.1080/00022470.1976.10470368.

Li, X., Brauers, T., Shao, M., Garland, R. M., Wagner, T., Deutschmann, T. and Wahner, A. (2010). 'MAX-DOAS measurements in southern China: retrieval of aerosol extinctions and validation using ground-based in-situ data'. In: *Atmospheric Chemistry and Physics* 10.5, pp. 2079–2089. DOI: 10.5194/acp-10-2079-2010.

Liao, J., Sihler, H., Huey, L. G., Neuman, J. A., Tanner, D. J., Friess, U., Platt, U., Flocke, F. M., Orlando, J. J., Shepson, P. B., Beine, H. J., Weinheimer, A. J., Sjostedt, S. J., Nowak, J. B., Knapp, D. J., Staebler, R. M., Zheng, W., Sander, R., Hall, S. R. and Ullmann, K. (2011). 'A comparison of Arctic BrO measurements by chemical ionization mass spectrometry and long path-differential optical absorption spectroscopy'. In: *Journal of Geophysical Research: Atmospheres* 116.D14. DOI: 10.1029/2010JD014788.

Lockwood, David J. (2016). 'Rayleigh and Mie Scattering'. In: *Encyclopedia of Color Science and Technology*. Ed. by Ming Ronnier Luo. New York, NY: Springer New York, pp. 1097–1107. DOI: 10.1007/978-1-4419-8071-7_218.

Lukošiūnaitė, Simona (2021). 'Deriving Nitrogen Oxide emissions from inland waterway vessels using MAX-DOAS measurements'. Master thesis. Johannes Gutenberg University Mainz.

Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P. and Wagner, T. (2013). 'Tropospheric NO_2 vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008 – 2011) and satellite validation'. In: *Atmospheric Chemistry and Physics* 13.3, pp. 1547–1567. DOI: 10.5194/acp-13-1547-2013.

MacDonald, S. M., Oetjen, H., Mahajan, A. S., Whalley, L. K., Edwards, P. M., Heard, D. E., Jones, C. E. and Plane, J. M. C. (2012). 'DOAS measurements of formaldehyde and glyoxal above a south-east Asian tropical rainforest'. In: *Atmospheric Chemistry and Physics* 12.13, pp. 5949–5962. DOI: 10.5194/acp-12-5949-2012.

Mahajan, A. S., Prados-Roman, C., Hay, T. D., Lampel, J., Pöhler, D., Großmann, K., Tschritter, J., Frieß, U., Platt, U., Johnston, P., Kreher, K., Wittrock, F., Burrows, J. P., Plane, J. M.C. and Saiz-Lopez, A. (2014). 'Glyoxal observations in the global marine boundary layer'. In: *Journal of Geophysical Research: Atmospheres* 119.10, pp. 6160–6169. DOI: 10.1002/2013JD021388.

Marquardt, D. W. (1963). 'An Algorithm for Least-Squares Estimation of Nonlinear Parameters'. In: *Journal of the Society for Industrial and Applied Mathematics* 11.2, pp. 431–441. URL: http://www.jstor.org/stable/2098941 (visited on 22/02/2024).

Meller, R. (1992). unpublished personal communication. MPIC.

Meller, R. and Moortgat, G. K. (2000). 'Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm'. In: *Journal of Geophysical Research: Atmospheres* 105.D6, pp. 7089–7101. DOI: 10.1029/1999JD901074.

Mellouki, A., Wallington, T. J. and Chen, J. (2015). 'Atmospheric Chemistry of Oxygenated Volatile Organic Compounds: Impacts on Air Quality and Climate'. In: *Chemical Reviews* 115.10. PMID: 25828273, pp. 3984–4014. DOI: 10.1021/cr500549n.

Mie, Gustav (1908). 'Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen'. In: *Annalen der Physik* 330.3, pp. 377–445. DOI: 10.1002/andp.19083300302.

Miller, H. L., Weaver, A., Sanders, R. W., Arpag, K. and Solomon, S. (1997). 'Measurements of arctic sunrise surface ozone depletion events at Kangerlussuaq, Greenland (67°N, 51°W)'. In: *Tellus B: Chemical and Physical Meteorology*. DOI: 10.3402/tellusb.v49i5.15988.

Mount, G. H. (1992). 'The measurement of tropospheric OH by long path absorption 1. Instrumentation'. In: *Journal of Geophysical Research: Atmospheres* 97.D2, pp. 2427–2444. DOI: 10.1029/91JD02971.

Myriokefalitakis, S., Vrekoussis, M., Tsigaridis, K., Wittrock, F., Richter, A., Brühl, C., Volkamer, R., Burrows, J. P. and Kanakidou, M. (2008). 'The influence of natural and anthropogenic secondary sources on the glyoxal global distribution'. In: *Atmospheric Chemistry and Physics* 8.16, pp. 4965–4981. DOI: 10.5194/acp-8-4965-2008.

NASA-GSFC (Feb. 2024a). *AERONET-Aerosol optical depth*. https://aeronet.gsfc.nasa.gov/new_web/aot_levels_versions.html. Accessed: 2024-02-01.

NASA-GSFC (Feb. 2024b). *AERONET-Aerosol robotic network*. https://aeronet.gsfc.nasa.gov/. Accessed: 2024-02-01.

NOAA – Global Monitoring Laboratory (Mar. 2024). *Carbon Cycle Greenhouse Gases*. https://gml.noaa.gov/ccgg/. Accessed: 2024-03-20.

Nasse, J.-M., Eger, P. G., Pöhler, D., Schmitt, S., Frieß, U. and Platt, U. (2019). 'Recent improvements of long-path DOAS measurements: impact on accuracy and

stability of short-term and automated long-term observations'. In: *Atmospheric Measurement Techniques* 12.8, pp. 4149–4169. DOI: 10.5194/amt-12-4149-2019.

Nishino, N., Arey, J. and Atkinson, R. (2010). 'Formation Yields of Glyoxal and Methylglyoxal from the Gas-Phase OH Radical-Initiated Reactions of Toluene, Xylenes, and Trimethylbenzenes as a Function of NO_2 Concentration'. In: *The Journal of Physical Chemistry A* 114.37, pp. 10140–10147. DOI: 10.1021/jp105112h.

Noxon, J. F. (1975). 'Nitrogen Dioxide in the Stratosphere and Troposphere Measured by Ground-Based Absorption Spectroscopy'. In: *Science* 189.4202, pp. 547–549. DOI: 10.1126/science.189.4202.547.

Noxon, J. F., Norton, R. B. and Henderson, W. R. (1978). 'Observation of atmospheric NO_3'. In: *Geophysical Research Letters* 5.8, pp. 675–678. DOI: 10.1029/GL005i008p00675.

Noxon, J. F., Whipple Jr., E. C. and Hyde, R. S. (1979). 'Stratospheric NO_2: 1. Observational method and behavior at mid-latitude'. In: *Journal of Geophysical Research: Oceans* 84.C8, pp. 5047–5065. DOI: 10.1029/JC084iC08p05047.

Nussbaumer, C. M., Crowley, J. N., Schuladen, J., Williams, J., Hafermann, S., Reiffs, A., Axinte, R., Harder, H., Ernest, C., Novelli, A., Sala, K., Martinez, M., Mallik, C., Tomsche, L., Plass-Dülmer, C., Bohn, B., Lelieveld, J. and Fischer, H. (2021). 'Measurement report: Photochemical production and loss rates of formaldehyde and ozone across Europe'. In: *Atmospheric Chemistry and Physics* 21.24, pp. 18413–18432. DOI: 10.5194/acp-21-18413-2021.

Nussbaumer, C. M., Fischer, H., Lelieveld, J. and Pozzer, A. (2023). 'What controls ozone sensitivity in the upper tropical troposphere?' In: *Atmospheric Chemistry and Physics* 23.19, pp. 12651–12669. DOI: 10.5194/acp-23-12651-2023.

Nussbaumer, C.M. and Cohen, R.C. (2020). 'The Role of Temperature and NOx in Ozone Trends in the Los Angeles Basin'. In: *Environmental Science & Technology* 54.24, pp. 15652–15659. DOI: 10.1021/acs.est.0c04910.

Ortega, I., Berg, L. K., Ferrare, R. A., Hair, J. W., Hostetler, C. H. and Volkamer, R. (2016). 'Elevated aerosol layers modify the O_2–O_2 absorption measured by ground-based MAX-DOAS'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 176, pp. 34–49. DOI: 10.1016/j.jqsrt.2016.02.021.

Perner, D., Ehhalt, D. H., Pätz, H. W., Platt, U., Röth, E. P. and Volz, A. (1976). 'OH - Radicals in the lower troposphere'. In: *Geophysical Research Letters* 3.8, pp. 466–468. DOI: 10.1029/GL003i008p00466.

Perner, D. and Platt, U. (1979). 'Detection of nitrous acid in the atmosphere by differential optical absorption'. In: *Geophysical Research Letters* 6.12, pp. 917–920. DOI: 10.1029/GL006i012p00917.

Pfannerstill, E. Y., Reijrink, N. G., Edtbauer, A., Ringsdorf, A., Zannoni, N., Araújo, A., Ditas, F., Holanda, B. A., Sá, M. O., Tsokankunku, A., Walter, D., Wolff, S., Lavrič, J. V., Pöhlker, C., Sörgel, M. and Williams, J. (2021). 'Total OH reactivity over the Amazon rainforest: variability with temperature, wind, rain, altitude,

time of day, season, and an overall budget closure'. In: *Atmospheric Chemistry and Physics* 21.8, pp. 6231–6256. DOI: 10.5194/acp-21-6231-2021.

Pinardi, G. (2020). unpublished personal communication. BIRA.

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., Fayt, C., Frieß, U., Gil, M., Herman, J., Hermans, C., Hendrick, F., Irie, H., Merlaud, A., Navarro Comas, M., Peters, E., Piters, A. J. M., Puentedura, O., Richter, A., Schönhardt, A., Shaiganfar, R., Spinei, E., Strong, K., Takashima, H., Vrekoussis, M., Wagner, T., Wittrock, F. and Yilmaz, S. (2013). 'MAX-DOAS formaldehyde slant column measurements during CINDI: intercomparison and analysis improvement'. In: *Atmospheric Measurement Techniques* 6.1, pp. 167–185. DOI: 10.5194/amt-6-167-2013.

Piters, A. J. M., Boersma, K. F., Kroon, M., Hains, J. C., Van Roozendael, M., Wittrock, F., Abuhassan, N., Adams, C., Akrami, M., Allaart, M. A. F., Apituley, A., Beirle, S., Bergwerff, J. B., Berkhout, A. J. C., Brunner, D., Cede, A., Chong, J., Clémer, K., Fayt, C., Frieß, U., Gast, L. F. L., Gil-Ojeda, M., Goutail, F., Graves, R., Griesfeller, A., Großmann, K., Hemerijckx, G., Hendrick, F., Henzing, B., Herman, J., Hermans, C., Hoexum, M., Hoff, G. R. van der, Irie, H., Johnston, P. V., Kanaya, Y., Kim, Y. J., Klein Baltink, H., Kreher, K., Leeuw, G. de, Leigh, R., Merlaud, A., Moerman, M. M., Monks, P. S., Mount, G. H., Navarro-Comas, M., Oetjen, H., Pazmino, A., Perez-Camacho, M., Peters, E., Piesanie, A. du, Pinardi, G., Puentedura, O., Richter, A., Roscoe, H. K., Schönhardt, A., Schwarzenbach, B., Shaiganfar, R., Sluis, W., Spinei, E., Stolk, A. P., Strong, K., Swart, D. P. J., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Whyte, C., Wilson, K. M., Yela, M., Yilmaz, S., Zieger, P. and Zhou, Y. (2012). 'The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results'. In: *Atmospheric Measurement Techniques* 5.2, pp. 457–485. DOI: 10.5194/amt-5-457-2012.

Platt, U., Perner, D. and Pätz, H. W. (1979). 'Simultaneous measurement of atmospheric CH_2O, O_3, and NO_2 by differential optical absorption'. In: *Journal of Geophysical Research: Oceans* 84.C10, pp. 6329–6335. DOI: 10.1029/JC084iC10p06329.

Platt, Ulrich and Stutz, Jochen (2008). 'Differential Absorption Spectroscopy'. In: *Differential Optical Absorption Spectroscopy: Principles and Applications*. Berlin, Heidelberg: Springer, pp. 135–174. DOI: 10.1007/978-3-540-75776-4_6.

Pöhlker, C., Walter, D., Paulsen, H., Könemann, T., Rodríguez-Caballero, E., Moran-Zuloaga, D., Brito, J., Carbone, S., Degrendele, C., Després, V. R., Ditas, F., Holanda, B. A., Kaiser, J. W., Lammel, G., Lavrič, J. V., Ming, J., Pickersgill, D., Pöhlker, M. L., Praß, M., Löbs, N., Saturno, J., Sörgel, M., Wang, Q., Weber, B., Wolff, S., Artaxo, P., Pöschl, U. and Andreae, M. O. (2019). 'Land cover and its transformation in the backward trajectory footprint region of the Amazon Tall Tower Observatory'. In: *Atmospheric Chemistry and Physics* 19.13, pp. 8425–8470. DOI: 10.5194/acp-19-8425-2019.

Pöhlker, M. L., Pöhlker, C., Ditas, F., Klimach, T., Angelis, I. Hrabe de, Araújo, A., Brito, J., Carbone, S., Cheng, Y., Chi, X., Ditz, R., Gunthe, S. S., Kesselmeier, J., Könemann, T., Lavrič, J. V., Martin, S. T., Mikhailov, E., Moran-Zuloaga,

D., Rose, D., Saturno, J., Su, H., Thalman, R., Walter, D., Wang, J., Wolff, S., Barbosa, H. M. J., Artaxo, P., Andreae, M. O. and Pöschl, U. (2016). 'Long-term observations of cloud condensation nuclei in the Amazon rain forest – Part 1: Aerosol size distribution, hygroscopicity, and new model parametrizations for CCN prediction'. In: *Atmospheric Chemistry and Physics* 16.24, pp. 15709–15740. DOI: 10.5194/acp-16-15709-2016.

Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., Tennyson, J., Yurchenko, S. N. and Lodi, L. (Aug. 2018). 'ExoMol molecular line lists XXX: a complete high-accuracy line list for water'. In: *Monthly Notices of the Royal Astronomical Society* 480.2, pp. 2597–2608. DOI: 10.1093/mnras/sty1877.

Puķīte, J., Kühl, S., Deutschmann, T., Platt, U. and Wagner, T. (2010). 'Extending differential optical absorption spectroscopy for limb measurements in the UV'. In: *Atmospheric Measurement Techniques* 3.3, pp. 631–653. DOI: 10.5194/amt-3-631-2010.

Puķīte, J. and Wagner, T. (2016). 'Quantification and parametrization of non-linearity effects by higher-order sensitivity terms in scattered light differential optical absorption spectroscopy'. In: *Atmospheric Measurement Techniques* 9.5, pp. 2147–2177. DOI: 10.5194/amt-9-2147-2016.

Punjab State Power Corporation Ltd. (2023). *Statistics*. https://www.pspcl.in/1Statistics. Accessed: 2023-07-26.

Pusede, S. E. and Cohen, R. C. (2012). 'On the observed response of ozone to NO_x and VOC reactivity reductions in San Joaquin Valley California 1995–present'. In: *Atmospheric Chemistry and Physics* 12.18, pp. 8323–8339. DOI: 10.5194/acp-12-8323-2012.

Pusede, S.E., Steiner, A.L. and Cohen, R.C. (2015). 'Temperature and Recent Trends in the Chemistry of Continental Surface Ozone'. In: *Chemical Reviews* 115.10, pp. 3898–3918. DOI: 10.1021/cr5006815.

Puķīte, J., Borger, C., Dörner, S., Gu, M., Frieß, U., Meier, A. C., Enell, C.-F., Raffalski, U., Richter, A. and Wagner, T. (2021). 'Retrieval algorithm for OClO from TROPOMI (TROPOspheric Monitoring Instrument) by differential optical absorption spectroscopy'. In: *Atmospheric Measurement Techniques* 14.12, pp. 7595–7625. DOI: 10.5194/amt-14-7595-2021.

Remmers, J., Beirle, S., Donner, S., Borger, C. and Wagner, T. (2019). *Investigating the vertical and horizontal distribution of trace gases and aerosols with the 4-Azimuth-MAX-DOAS in Mainz, Germany*. https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p/ExtraContent/ContentPage?page=5. Accessed: 2024-03-29. S5PVT Workshop, Frascati, Italy.

Remmers, Julia (2011). 'Untersuchung der räumlichen Verteilung troposphärischer Spurenstoffe und Aerosole mit der MAX-DOAS-Methode'. Master thesis. Technische Universität Darmstadt.

Richter, A., Eisinger, M., Ladstätter-Weißenmayer, A. and Burrows, J.P. (1999). 'DOAS Zenith Sky Observations: 2. Seasonal Variation of BrO Over Bremen

(53°N) 1994-1995'. In: *Journal of Atmospheric Chemistry* 32, pp. 83–99. DOI: 10.1023/A:1006077725894.

Richter, Andreas and Wagner, Thomas (2011). 'The use of UV, Visible and Near IR solar back scattered radiation to determine trace gases'. In: *The Remote Sensing of Tropospheric Composition from Space*. Springer, pp. 67–121.

Rodgers, C. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practic*. World Scientific, 238 pp.

Rothman, L.S., Gordon, I.E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P.F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L.R., Campargue, A., Chance, K., Cohen, E.A., Coudert, L.H., Devi, V.M., Drouin, B.J., Fayt, A., Flaud, J.-M., Gamache, R.R., Harrison, J.J., Hartmann, J.-M., Hill, C., Hodges, J.T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R.J., Li, G., Long, D.A., Lyulin, O.M., Mackie, C.J., Massie, S.T., Mikhailenko, S., Müller, H.S.P., Naumenko, O.V., Nikitin, A.V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E.R., Richard, C., Smith, M.A.H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G.C., Tyuterev, Vl.G. and Wagner, G. (2013). 'The HITRAN2012 molecular spectroscopic database'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 130. HITRAN2012 special issue, pp. 4–50. DOI: 10.1016/j.jqsrt.2013.07.002.

Rothman, L.S., Gordon, I.E., Barber, R.J., Dothe, H., Gamache, R.R., Goldman, A., Perevalov, V.I., Tashkun, S.A. and Tennyson, J. (2010). 'HITEMP, the high-temperature molecular spectroscopic database'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 111.15. XVIth Symposium on High Resolution Molecular Spectroscopy (HighRus-2009), pp. 2139–2150. DOI: 10.1016/j.jqsrt.2010.05.001.

Rozanov, V.V., Rozanov, A.V., Kokhanovsky, A.A. and Burrows, J.P. (2014). 'Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 133, pp. 13 –71. DOI: 10.1016/j.jqsrt.2013.07.004.

SCHOTT AG (2023). *Data Sheet - BG3*. https://www.schott.com/shop/advanced-optics/en/Matt-Filter-Plates/BG3/c/glass-BG3. Accessed: 2023-08-08.

Salthammer, T. (2013). 'Formaldehyde in the Ambient Atmosphere: From an Indoor Pollutant to an Outdoor Pollutant?' In: *Angewandte Chemie International Edition* 52.12, pp. 3320–3327. DOI: 10.1002/anie.201205984.

Sanders, R. W., Solomon, S., Smith, J. P., Perliski, L., Miller, H. L., Mount, G. H., Keys, J. G. and Schmeltekopf, A. L. (1993). 'Visible and near-ultraviolet spectroscopy at McMurdo Station, Antarctica: 9. Observations of OClO from April to October 1991'. In: *Journal of Geophysical Research: Atmospheres* 98.D4, pp. 7219–7228. DOI: 10.1029/93JD00042.

Schreier, S. F., Richter, A., Peters, E., Ostendorf, M., Schmalwieser, A. W., Weihs, P. and Burrows, J.P. (2020). 'Dual ground-based MAX-DOAS observations in Vienna, Austria: Evaluation of horizontal and temporal NO_2, HCHO, and CHOCHO distributions and comparison with independent data sets'. In: *Atmospheric Environment: X* 5, p. 100059. DOI: 10.1016/j.aeaoa.2019.100059.

Seco, R., Peñuelas, J. and Filella, I. (2007). 'Short-chain oxygenated VOCs: Emission and uptake by plants and atmospheric sources, sinks, and concentrations'. In: *Atmospheric Environment* 41.12, pp. 2477–2499. DOI: 10.1016/j.atmosenv.2006.11.029.

Seinfeld, J. H. and Pandis, S. N. (2016). *Atmospheric chemistry and physics: From Air Pollution to Climate Change*. Hoboken, New Jersey: John Wiley & Sons, pp. 175–264. ISBN: 978-1-118-94740-1.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W. and Burrows, J. P. (2014). 'High spectral resolution ozone absorption cross-sections - Part 2: Temperature dependence'. In: *Atmospheric Measurement Techniques* 7.2, pp. 625–636. DOI: 10.5194/amt-7-625-2014.

Shaiganfar, R., Beirle, S., Gon, H. Denier van der, Jonkers, S., Kuenen, J., Petetin, H., Zhang, Q., Beekmann, M. and Wagner, T. (2017). 'Estimation of the Paris NO_x emissions from mobile MAX-DOAS observations and CHIMERE model simulations during the MEGAPOLI campaign using the closed integral method'. In: *Atmospheric Chemistry and Physics* 17.12, pp. 7853–7890. DOI: 10.5194/acp-17-7853-2017.

Sinha, V., Kumar, V. and Sarkar, C. (2014). 'Chemical composition of pre-monsoon air in the Indo-Gangetic Plain measured using a new air quality facility and PTR-MS: high surface ozone and strong influence of biomass burning'. In: *Atmospheric Chemistry and Physics* 14.12, pp. 5921–5941. DOI: 10.5194/acp-14-5921-2014.

Sinreich, R., Coburn, S., Dix, B. and Volkamer, R. (2010). 'Ship-based detection of glyoxal over the remote tropical Pacific Ocean'. In: *Atmospheric Chemistry and Physics* 10.23, pp. 11359–11371. DOI: 10.5194/acp-10-11359-2010.

Sinreich, R., Volkamer, R., Filsinger, F., Frieß, U., Kern, C., Platt, U., Sebastián, O. and Wagner, T. (2007). 'MAX-DOAS detection of glyoxal during ICARTT 2004'. In: *Atmospheric Chemistry and Physics* 7.5, pp. 1293–1303. DOI: 10.5194/acp-7-1293-2007.

Solomon, S., Mount, G. H., Sanders, R. W. and Schmeltekopf, A. L. (1987a). 'Visible spectroscopy at McMurdo Station, Antarctica: 2. Observations of OClO'. In: *Journal of Geophysical Research: Atmospheres* 92.D7, pp. 8329–8338. DOI: 10.1029/JD092iD07p08329.

Solomon, S., Schmeltekopf, A. L. and Sanders, R. W. (1987b). 'On the interpretation of zenith sky absorption measurements'. In: *Journal of Geophysical Research: Atmospheres* 92.D7, pp. 8311–8319. DOI: 10.1029/JD092iD07p08311.

Spinei, E., Cede, A., Herman, J., Mount, G. H., Eloranta, E., Morley, B., Baidar, S., Dix, B., Ortega, I., Koenig, T. and Volkamer, R. (2015). 'Ground-based direct-sun DOAS and airborne MAX-DOAS measurements of the collision-induced oxygen complex, O_2O_2, absorption with significant pressure and temperature differences'. In: *Atmospheric Measurement Techniques* 8.2, pp. 793–809. DOI: 10.5194/amt-8-793-2015.

Stadt Frankfurt am Main (Feb. 2023). *Themen – Bevölkerung*. https://frankfurt.de/service-und-rathaus/zahlen-daten-fakten/themen/bevoelkerung. Accessed: 2023-07-25.

Stadt Wiesbaden (July 2023). *Standortporträt Wiesbaden*. https://www.wiesbaden.de/wirtschaft/wirtschaftsstandort/standortportraet.php. Accessed: 2023-07-25.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P. (2009a). 'The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling'. In: *Atmospheric Chemistry and Physics* 9.21, pp. 8431–8446. DOI: 10.5194/acp-9-8431-2009.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Werf, G. R. van der, Giglio, L. and Guenther, A. (2009b). 'Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006'. In: *Atmospheric Chemistry and Physics* 9.11, pp. 3663–3679. DOI: 10.5194/acp-9-3663-2009.

Steiner, A. H. and Goldstein, A. L. (2007). 'Biogenic VOCs'. In: *Volatile Organic Compounds in the Atmosphere*. Ed. by R. Koppmann. Oxford: Blackwell Publishing, pp. 82–128. ISBN: 978-1-4051-3115-5.

Strutt Lord Rayleigh, J. W. (1899). 'On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky'. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47.287, pp. 375–384. DOI: 10.1080/14786449908621276.

Stutz, J., Kim, E. S., Platt, U., Bruno, P., Perrino, C. and Febo, A. (2000). 'UV-visible absorption cross sections of nitrous acid'. In: *Journal of Geophysical Research: Atmospheres* 105.D11, pp. 14585–14592. DOI: 10.1029/2000JD900003.

Stutz, J., Thomas, J. L., Hurlock, S. C., Schneider, M., Glasow, R. von, Piot, M., Gorham, K., Burkhart, J. F., Ziemba, L., Dibb, J. E. and Lefer, B. L. (2011). 'Long-path DOAS observations of surface BrO at Summit, Greenland'. In: *Atmospheric Chemistry and Physics* 11.18, pp. 9899–9910. DOI: 10.5194/acp-11-9899-2011.

Stutz, Jochen and Platt, Ulrich (1996). 'Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods'. In: *Appl. Opt.* 35.30, pp. 6041–6053. DOI: 10.1364/AO.35.006041.

Thalman, R. and Volkamer, R. (2013). 'Temperature dependent absorption cross-sections of O_2–O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure'. In: *Physical Chemistry Chemical Physics* 15 (37), pp. 15371–15381. DOI: 10.1039/C3CP50968K.

Thomas, W., Hegels, E., Slijkhuis, S., Spurr, R. and Chance, K. (1998). 'Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME Spectrometer'. In: *Geophysical Research Letters* 25.9, pp. 1317–1320. DOI: 10.1029/98GL01087.

University of Hertfordshire (July 2023). *Bayfordbury Observatory*. https://www.herts.ac.uk/bayfordbury/bayfordbury-observatory. Accessed: 2023-07-18.

Vairavamurthy, A., Roberts, J. M. and Newman, L. (1992). 'Methods for determination of low molecular weight carbonyl compounds in the atmosphere: A review'. In: *Atmospheric Environment. Part A. General Topics* 26.11, pp. 1965–1993. DOI: 10.1016/0960-1686(92)90083-W.

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A. and Coquart, B. (1998). 'Measurements of the NO_2 absorption cross-section from 42 000 cm^{-1} to 10 000 cm^{-1} (238-1000 nm) at 220 K and 294 K'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 59.3. Atmospheric Spectroscopy Applications 96, pp. 171 –184. DOI: 10.1016/S0022-4073(97)00168-4.

Vandaele, A. C., Hermans, C., Simon, P. C., Van Roozendael, M., Guilmot, J. M., Carleer, M. and Colin, R. (1996). 'Fourier transform measurement of NO_2 absorption cross-section in the visible range at room temperature'. In: *Journal of Atmospheric Chemistry* 25, pp. 289–305. DOI: 10.1007/BF00053797.

Veefkind, J.P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H.J., de Haan, J.F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H. and Levelt, P.F. (2012). 'TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications'. In: *Remote Sensing of Environment* 120. The Sentinel Missions - New Opportunities for Science, pp. 70–83. DOI: 10.1016/j.rse.2011.09.027.

Vlemmix, T., Hendrick, F., Pinardi, G., De Smedt, I., Fayt, C., Hermans, C., Piters, A., Wang, P., Levelt, P. and Van Roozendael, M. (2015). 'MAX-DOAS observations of aerosols, formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches'. In: *Atmospheric Measurement Techniques* 8.2, pp. 941–963. DOI: 10.5194/amt-8-941-2015.

Vogel, L., Sihler, H., Lampel, J., Wagner, T. and Platt, U. (2013). 'Retrieval interval mapping: a tool to visualize the impact of the spectral retrieval range on differential optical absorption spectroscopy evaluations'. In: *Atmospheric Measurement Techniques* 6.2, pp. 275–299. DOI: 10.5194/amt-6-275-2013.

Volkamer, R., Molina, L. T., Molina, M. J., Shirley, T. and Brune, W. H. (2005a). 'DOAS measurement of glyoxal as an indicator for fast VOC chemistry in urban air'. In: *Geophysical Research Letters* 32.8. DOI: 10.1029/2005GL022616.

Volkamer, R., Platt, U. and Wirtz, K. (2001). 'Primary and Secondary Glyoxal Formation from Aromatics: Experimental Evidence for the Bicycloalkyl-Radical Pathway from Benzene, Toluene, and p–Xylene'. In: *The Journal of Physical Chemistry A* 105.33, pp. 7865–7874. DOI: 10.1021/jp010152w.

Volkamer, R., San Martini, F., Molina, L. T., Salcedo, D., Jimenez, J. L. and Molina, M. J. (2007). 'A missing sink for gas-phase glyoxal in Mexico City: Formation of

secondary organic aerosol'. In: *Geophysical Research Letters* 34.19. DOI: 10.1029/2007GL030752.

Volkamer, R., Spietz, P., Burrows, J. and Platt, U. (2005b). 'High-resolution absorption cross-section of glyoxal in the UV–vis and IR spectral ranges'. In: *Journal of Photochemistry and Photobiology A: Chemistry* 172.1, pp. 35–46. DOI: 10.1016/j.jphotochem.2004.11.011.

Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P. (2010). 'GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?' In: *Atmospheric Chemistry and Physics* 10.21, pp. 10145–10160. DOI: 10.5194/acp-10-10145-2010.

Wagner, T. (1999). 'Satellite Observations of Atmospheric Halogen Oxides'. PhD thesis. Universität Heidelberg.

Wagner, T. (2019). unpublished personal communication. MPIC.

Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J. and Shaiganfar, R. (2014). 'Cloud detection and classification based on MAX-DOAS observations'. In: *Atmospheric Measurement Techniques* 7.5, pp. 1289–1320. DOI: 10.5194/amt-7-1289-2014.

Wagner, T., Beirle, S., Benavent, N., Bösch, T., Chan, K. L., Donner, S., Dörner, S., Fayt, C., Frieß, U., García-Nieto, D., Gielen, C., González-Bartolome, D., Gomez, L., Hendrick, F., Henzing, B., Jin, J. L., Lampel, J., Ma, J., Mies, K., Navarro, M., Peters, E., Pinardi, G., Puentedura, O., Puķīte, J., Remmers, J., Richter, A., Saiz-Lopez, A., Shaiganfar, R., Sihler, H., Van Roozendael, M., Wang, Y. and Yela, M. (2019). 'Is a scaling factor required to obtain closure between measured and modelled atmospheric O_4 absorptions? An assessment of uncertainties of measurements and radiative transfer simulations for 2 selected days during the MAD-CAT campaign'. In: *Atmospheric Measurement Techniques* 12.5, pp. 2745–2817. DOI: 10.5194/amt-12-2745-2019.

Wagner, T., Beirle, S., Brauers, T., Deutschmann, T., Frieß, U., Hak, C., Halla, J. D., Heue, K. P., Junkermann, W., Li, X., Platt, U. and Pundt-Gruber, I. (2011). 'Inversion of tropospheric profiles of aerosol extinction and HCHO and NO_2 mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets'. In: *Atmospheric Measurement Techniques* 4.12, pp. 2685–2715. DOI: 10.5194/amt-4-2685-2011.

Wagner, T., Beirle, S. and Deutschmann, T. (2009). 'Three-dimensional simulation of the Ring effect in observations of scattered sun light using Monte Carlo radiative transfer models'. In: *Atmospheric Measurement Techniques* 2.1, pp. 113–124. DOI: 10.5194/amt-2-113-2009.

Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R. and Wang, Y. (2016). 'Absolute calibration of the colour index and O_4 absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm'. In: *Atmospheric Measurement Techniques* 9.9, pp. 4803–4823. DOI: 10.5194/amt-9-4803-2016.

Wagner, T., Burrows, J. P., Deutschmann, T., Dix, B., Friedeburg, C. von, Frieß, U., Hendrick, F., Heue, K.-P., Irie, H., Iwabuchi, H., Kanaya, Y., Keller, J., McLinden, C. A., Oetjen, H., Palazzi, E., Petritoli, A., Platt, U., Postylyakov, O., Pukite, J., Richter, A., Roozendael, M. van, Rozanov, A., Rozanov, V., Sinreich, R., Sanghavi, S. and Wittrock, F. (2007). 'Comparison of box-air-mass-factors and radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models'. In: *Atmospheric Chemistry and Physics* 7.7, pp. 1809–1833. DOI: 10.5194/acp-7-1809-2007.

Wagner, T., Chance, K., Frieß, U., Gil, M., Goutail, F., Hönninger, G., Johnston, P., Karlsen-Tørnkvist, K., Kostadinov, I., Leser, H., Petritoli, A., Richter, A., Van Roozendael, M. and Platt, U. (2002a). *Correction of the Ring effect and I_0-effect for DOAS observations of scattered sunlight*. Tech. rep. ESA Technical Report. The Netherlands: ESTEC.

Wagner, T., Dix, B., Friedeburg, C. v., Frieß, U., Sanghavi, S., Sinreich, R. and Platt, U. (2004). 'MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols—Principles and information content'. In: *Journal of Geophysical Research: Atmospheres* 109.D22. DOI: 10.1029/2004JD004904.

Wagner, T., Dörner, S., Beirle, S., Donner, S. and Kinne, S. (2021). 'Quantitative comparison of measured and simulated O_4 absorptions for one day with extremely low aerosol load over the tropical Atlantic'. In: *Atmospheric Measurement Techniques* 14.5, pp. 3871–3893. DOI: 10.5194/amt-14-3871-2021.

Wagner, T., Friedeburg, C. von, Wenig, M., Otten, C. and Platt, U. (2002b). 'UV-visible observations of atmospheric O_4 absorptions using direct moonlight and zenith-scattered sunlight for clear-sky and cloudy sky conditions'. In: *Journal of Geophysical Research: Atmospheres* 107.D20, AAC 3–1–AAC 3–15. DOI: 10.1029/2001JD001026.

Wagner, T., Ibrahim, O., Shaiganfar, R. and Platt, U. (2010). 'Mobile MAX-DOAS observations of tropospheric trace gases'. In: *Atmospheric Measurement Techniques* 3.1, pp. 129–140. DOI: 10.5194/amt-3-129-2010.

Wagner, T., Otten, C., Pfeilsticker, K., Pundt, I. and Platt, U. (2000). 'DOAS moonlight observation of atmospheric NO_3 in the Arctic winter'. In: *Geophysical Research Letters* 27.21, pp. 3441–3444. DOI: 10.1029/1999GL011153.

Wang, Y., Apituley, A., Bais, A., Beirle, S., Benavent, N., Borovski, A., Bruchkouski, I., Chan, K. L., Donner, S., Drosoglou, T., Finkenzeller, H., Friedrich, M. M., Frieß, U., Garcia-Nieto, D., Gómez-Martín, L., Hendrick, F., Hilboll, A., Jin, J., Johnston, P., Koenig, T. K., Kreher, K., Kumar, V., Kyuberis, A., Lampel, J., Liu, C., Liu, H., Ma, J., Polyansky, O. L., Postylyakov, O., Querel, R., Saiz-Lopez, A., Schmitt, S., Tian, X., Tirpitz, J.-L., Van Roozendael, M., Volkamer, R., Wang, Z., Xie, P., Xing, C., Xu, J., Yela, M., Zhang, C. and Wagner, T. (2020). 'Intercomparison of MAX-DOAS measurements of tropospheric HONO slant column densities and vertical profiles during the CINDI-2 campaign'. In: *Atmospheric Measurement Techniques* 13.9, pp. 5087–5116. DOI: 10.5194/amt-13-5087-2020.

Wang, Y., Beirle, S., Hendrick, F., Hilboll, A., Jin, J., Kyuberis, A. A., Lampel, J., Li, A., Luo, Y., Lodi, L., Ma, J., Navarro, M., Ortega, I., Peters, E., Polyansky, O. L., Remmers, J., Richter, A., Puentedura, O., Van Roozendael, M., Seyler, A., Tennyson, J., Volkamer, R., Xie, P., Zobov, N. F. and Wagner, T. (2017a). 'MAX-DOAS measurements of HONO slant column densities during the MAD-CAT campaign: inter-comparison, sensitivity studies on spectral analysis settings, and error budget'. In: *Atmospheric Measurement Techniques* 10.10, pp. 3719–3742. DOI: 10.5194/amt-10-3719-2017.

Wang, Y., Dörner, S., Donner, S., Böhnke, S., De Smedt, I., Dickerson, R. R., Dong, Z., He, H., Li, Z., Li, Z., Li, D., Liu, D., Ren, X., Theys, N., Wang, Y., Wang, Y., Wang, Z., Xu, H., Xu, J. and Wagner, T. (2019). 'Vertical profiles of NO_2, SO_2, HONO, HCHO, CHOCHO and aerosols derived from MAX-DOAS measurements at a rural site in the central western North China Plain and their relation to emission sources and effects of regional transport'. In: *Atmospheric Chemistry and Physics* 19.8, pp. 5417–5449. DOI: 10.5194/acp-19-5417-2019.

Wang, Y., Lampel, J., Xie, P., Beirle, S., Li, A., Wu, D. and Wagner, T. (2017b). 'Ground-based MAX-DOAS observations of tropospheric aerosols, NO_2, SO_2 and HCHO in Wuxi, China, from 2011 to 2014'. In: *Atmospheric Chemistry and Physics* 17.3, pp. 2189–2215. DOI: 10.5194/acp-17-2189-2017.

Williams, J. and Koppmann, R. (2007). 'Volatile Organic Compounds in the Atmosphere: An Overview'. In: *Volatile Organic Compounds in the Atmosphere*. Ed. by R. Koppmann. Oxford: Blackwell Publishing, pp. 1–32. ISBN: 978-1-4051-3115-5.

Wilmouth, D. M., Hanisco, T. F., Donahue, N. M. and Anderson, J. G. (1999). 'Fourier Transform Ultraviolet Spectroscopy of the A $^2\Pi_{3/2} \leftarrow$ X $^2\Pi_{3/2}$ Transition of BrO'. In: *The Journal of Physical Chemistry A* 103.45, pp. 8935–8945. DOI: 10.1021/jp9916510.

Wissenschaftsstadt Darmstadt (2021). *Bevölkerung*. https://www.darmstadt.de/fileadmin/Bilder-Rubriken/Standort/Statistik_und_Stadtforschung/pdf/jahrbuch/02Bevoelkerung/K02-1.pdf. Accessed: 2023-07-25.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U. and Wagner, T. (2006). 'Simultaneous global observations of glyoxal and formaldehyde from space'. In: *Geophysical Research Letters* 33.16. DOI: 10.1029/2006GL026310.

Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., Gouw, J. A. de, Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz, L. W., Lee, B. H., Lerner, B. M., Lopez-Hilifiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., Thornton, J. A., Veres, P. R. and Warneke, C. (2016). 'Formaldehyde production from isoprene oxidation across NO_x regimes'. In: *Atmospheric Chemistry and Physics* 16.4, pp. 2597–2610. DOI: 10.5194/acp-16-2597-2016.

World Meteorological Organization (2024). *State of the Global Climate 2023*. https://library.wmo.int/idurl/4/68835. Accessed: 2024-03-20. Geneva.

Zeldovich, Y.B. (1946). 'The Oxidation of Nitrogen in Combustion and Explosions'. In: *Acta Physicochimica U.S.S.R.* 21, pp. 577–628.

Zhang, R., Tie, X. and Bond, D.W. (2003). 'Impacts of anthropogenic and natural NO_x sources over the U.S. on tropospheric chemistry'. In: *Proceedings of the National Academy of Sciences* 100.4, pp. 1505–1509. DOI: 10.1073/pnas.252763799.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W.F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O'Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J. and Chance, K. (2017). 'Tropospheric emissions: Monitoring of pollution (TEMPO)'. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 186. Satellite Remote Sensing and Spectroscopy: Joint ACE-Odin Meeting, October 2015, pp. 17–39. DOI: 10.1016/j.jqsrt.2016.05.008.

Zörner, J., Vries, M. Penning de, Beirle, S., Sihler, H., Veres, P. R., Williams, J. and Wagner, T. (2016). 'Multi-satellite sensor study on precipitation-induced emission pulses of NO_x from soils in semi-arid ecosystems'. In: *Atmospheric Chemistry and Physics* 16.14, pp. 9457–9487. DOI: 10.5194/acp-16-9457-2016.

Ångström, Anders (1929). 'On the Atmospheric Transmission of Sun Radiation and on Dust in the Air'. In: *Geografiska Annaler* 11, pp. 156–166. DOI: 10.2307/519399. (Visited on 29/01/2024).

Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y. and Wagner, T. (2019). 'The Mainz profile algorithm (MAPA)'. In: *Atmospheric Measurement Techniques* 12.3, pp. 1785–1806. DOI: 10.5194/amt-12-1785-2019.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H. and Van Roozendael, M. (2021). 'Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements'. In: *Atmospheric Chemistry and Physics* 21.16, pp. 12561–12593. DOI: 10.5194/acp-21-12561-2021.

Donner, S. (2014). 'Aufbau und Charakterisierung eines MAX-DOAS-Instruments'. Bachelor thesis. Johannes Gutenberg University Mainz.

Donner, S. (2016). 'Mobile MAX-DOAS measurements of the tropospheric formaldehyde column in the Rhein-Main region'. Master thesis. Johannes Gutenberg University Mainz. URL: http://hdl.handle.net/11858/00-001M-0000-002C-EB17-2.

Donner, S., Kuhn, J., Van Roozendael, M., Bais, A., Beirle, S., Bösch, T., Bognar, K., Bruchkouski, I., Chan, K. L., Dörner, S., Drosoglou, T., Fayt, C., Frieß, U., Hendrick, F., Hermans, C., Jin, J., Li, A., Ma, J., Peters, E., Pinardi, G., Richter, A., Schreier, S. F., Seyler, A., Strong, K., Tirpitz, J.-L., Wang, Y., Xie, P., Xu, J., Zhao, X. and Wagner, T. (2020). 'Evaluating different methods for elevation calibration of MAX-DOAS (Multi AXis Differential Optical Absorption Spectroscopy) instruments during the CINDI-2 campaign'. In: *Atmospheric Measurement Techniques* 13.2, pp. 685–712. DOI: 10.5194/amt-13-685-2020.

Kreher, K., Van Roozendael, M., Hendrick, F., Apituley, A., Dimitropoulou, E., Frieß, U., Richter, A., Wagner, T., Lampel, J., Abuhassan, N., Ang, L., Anguas, M., Bais, A., Benavent, N., Bösch, T., Bognar, K., Borovski, A., Bruchkouski, I., Cede, A., Chan, K. L., Donner, S., Drosoglou, T., Fayt, C., Finkenzeller, H., Garcia-Nieto, D., Gielen, C., Gómez-Martín, L., Hao, N., Henzing, B., Herman, J. R., Hermans, C., Hoque, S., Irie, H., Jin, J., Johnston, P., Khayyam Butt, J., Khokhar, F., Koenig, T. K., Kuhn, J., Kumar, V., Liu, C., Ma, J., Merlaud, A., Mishra, A. K., Müller, M., Navarro-Comas, M., Ostendorf, M., Pazmino, A., Peters, E., Pinardi, G., Pinharanda, M., Piters, A., Platt, U., Postylyakov, O., Prados-Roman, C., Puentedura, O., Querel, R., Saiz-Lopez, A., Schönhardt, A., Schreier, S. F., Seyler, A., Sinha, V., Spinei, E., Strong, K., Tack, F., Tian, X., Tiefengraber, M., Tirpitz, J.-L., Gent, J. van, Volkamer, R., Vrekoussis, M., Wang, S., Wang, Z., Wenig, M., Wittrock, F., Xie, P. H., Xu, J., Yela, M., Zhang, C. and Zhao, X. (2020). 'Intercomparison of NO_2, O_4, O_3 and HCHO slant column

measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2'. In: *Atmospheric Measurement Techniques* 13.5, pp. 2169–2208. DOI: 10.5194/amt-13-2169-2020.

Kumar, V., Beirle, S., Dörner, S., Mishra, A. K., Donner, S., Wang, Y., Sinha, V. and Wagner, T. (2020). 'Long-term MAX-DOAS measurements of NO_2, HCHO, and aerosols and evaluation of corresponding satellite data products over Mohali in the Indo-Gangetic Plain'. In: *Atmospheric Chemistry and Physics* 20.22, pp. 14183–14235. DOI: 10.5194/acp-20-14183-2020.

Lange, K., Richter, A., Schönhardt, A., Meier, A. C., Bösch, T., Seyler, A., Krause, K., Behrens, L. K., Wittrock, F., Merlaud, A., Tack, F., Fayt, C., Friedrich, M. M., Dimitropoulou, E., Van Roozendael, M., Kumar, V., Donner, S., Dörner, S., Lauster, B., Razi, M., Borger, C., Uhlmannsiek, K., Wagner, T., Ruhtz, T., Eskes, H., Bohn, B., Santana Diaz, D., Abuhassan, N., Schüttemeyer, D. and Burrows, J. P. (2023). 'Validation of Sentinel-5P TROPOMI tropospheric NO_2 products by comparison with NO_2 measurements from airborne imaging DOAS, ground-based stationary DOAS, and mobile car DOAS measurements during the S5P-VAL-DE-Ruhr campaign'. In: *Atmospheric Measurement Techniques* 16.5, pp. 1357–1389. DOI: 10.5194/amt-16-1357-2023.

Lauster, B., Dörner, S., Beirle, S., Donner, S., Gromov, S., Uhlmannsiek, K. and Wagner, T. (2021). 'Estimating real driving emissions from multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements at the A60 motorway near Mainz, Germany'. In: *Atmospheric Measurement Techniques* 14.1, pp. 769–783. DOI: 10.5194/amt-14-769-2021.

Wagner, T., Beirle, S., Benavent, N., Bösch, T., Chan, K. L., Donner, S., Dörner, S., Fayt, C., Frieß, U., García-Nieto, D., Gielen, C., González-Bartolome, D., Gomez, L., Hendrick, F., Henzing, B., Jin, J. L., Lampel, J., Ma, J., Mies, K., Navarro, M., Peters, E., Pinardi, G., Puentedura, O., Puķīte, J., Remmers, J., Richter, A., Saiz-Lopez, A., Shaiganfar, R., Sihler, H., Van Roozendael, M., Wang, Y. and Yela, M. (2019). 'Is a scaling factor required to obtain closure between measured and modelled atmospheric O_4 absorptions? An assessment of uncertainties of measurements and radiative transfer simulations for 2 selected days during the MAD-CAT campaign'. In: *Atmospheric Measurement Techniques* 12.5, pp. 2745–2817. DOI: 10.5194/amt-12-2745-2019.

Wagner, T., Dörner, S., Beirle, S., Donner, S. and Kinne, S. (2021). 'Quantitative comparison of measured and simulated O_4 absorptions for one day with extremely low aerosol load over the tropical Atlantic'. In: *Atmospheric Measurement Techniques* 14.5, pp. 3871–3893. DOI: 10.5194/amt-14-3871-2021.

Wang, Y., Apituley, A., Bais, A., Beirle, S., Benavent, N., Borovski, A., Bruchkouski, I., Chan, K. L., Donner, S., Drosoglou, T., Finkenzeller, H., Friedrich, M. M., Frieß, U., Garcia-Nieto, D., Gómez-Martín, L., Hendrick, F., Hilboll, A., Jin, J., Johnston, P., Koenig, T. K., Kreher, K., Kumar, V., Kyuberis, A., Lampel, J., Liu, C., Liu, H., Ma, J., Polyansky, O. L., Postylyakov, O., Querel, R., Saiz-Lopez, A., Schmitt, S., Tian, X., Tirpitz, J.-L., Van Roozendael, M., Volkamer, R., Wang, Z., Xie, P., Xing, C., Xu, J., Yela, M., Zhang, C. and Wagner, T. (2020). 'Inter-

comparison of MAX-DOAS measurements of tropospheric HONO slant column densities and vertical profiles during the CINDI-2 campaign'. In: *Atmospheric Measurement Techniques* 13.9, pp. 5087–5116. DOI: 10.5194/amt-13-5087-2020.

Wang, Y., Dörner, S., Donner, S., Böhnke, S., De Smedt, I., Dickerson, R. R., Dong, Z., He, H., Li, Z., Li, Z., Li, D., Liu, D., Ren, X., Theys, N., Wang, Y., Wang, Y., Wang, Z., Xu, H., Xu, J. and Wagner, T. (2019). 'Vertical profiles of NO_2, SO_2, HONO, HCHO, CHOCHO and aerosols derived from MAX-DOAS measurements at a rural site in the central western North China Plain and their relation to emission sources and effects of regional transport'. In: *Atmospheric Chemistry and Physics* 19.8, pp. 5417–5449. DOI: 10.5194/acp-19-5417-2019.

ACKNOWLEDGEMENTS

At the end of this long journey, it is time to pause for a moment and thank the people who have travelled this path with me. The personal acknowledgements have been removed from the electronic version of the dissertation for data privacy reasons. Nevertheless, below some more general and institutional acknowledgements are given.

All members of the Satellite Remote Sensing Group deserve my thanks for the relaxed and friendly working atmosphere as well as scientific discussions and suggestions. Not to be forgotten are the non-work-related activities such as group excursions, coffee and lunch breaks, board game nights and the often interesting discussions.

I would also like to acknowledge the creators of the external data sets used in this work and the people without whom the measurements at all these wonderful locations would not have been possible.

In the first place, I would like to mention both the electronics and the mechanical workshops of the MPIC. Since the "birth" of the TubeMAX-DOAS instrument in 2015, they have put a lot of time and effort into maintaining and improving my "baby", often literally right up to the last second (without functioning instrument software travelling to ATTO was quite an experience).

Further, all institutions, international partners and other people involved at the different measurement sites are acknowledged not only for technical and organisational support but also for providing additional auxiliary data sets. Special credit goes to the global DOAS-family and all the people and international partners I worked with and met during campaigns and conferences.

This work used ERA5 data which were downloaded from the Copernicus Climate Change Service (2022) and thus contains modified Copernicus Climate Change Service information (2022). I acknowledge all involved people and institutions, in particular ECMWF.

The AERONET network is established by NASA and PHOTONS, I would like to acknowledge all involved people and institutions. In particular, I thank the local PIs for establishing and maintaining the AERONET sites at Bayfordbury, Mainz, Amity University Gurugram Haryana and ATTO whose data have been used in this study.

Modified Copernicus Sentinel-5P (TROPOMI) L2 HCHO data were used in this work. At this point, I would like to acknowledge all involved people and institutions, in particular BIRA-IASB, DLR and ESA.

Last but not least, I would like to thank all people who supported me privately. This includes my friends and in particular my family.

The curriculum vitae is not included in the electronic version.