Calculating the vertical column density of O_4 from surface values of pressure, temperature and relative humidity

Steffen Beirle[1], Christian Borger[1], Steffen Dörner[1], Vinod Kumar[1], and Thomas Wagner[1]

[1]Max-Planck-Institut für Chemie

Correspondence: Steffen Beirle (steffen.beirle@mpic.de)

Abstract. We present a formalism that relates the vertical column density (VCD) of the oxygen collision complex O_2-O_2 (denoted as O_4 below) to surface ($2\,m$) values of temperature and pressure, based on physical laws. In addition, we propose an empirical modification which also accounts for surface relative humidity (RH). This allows for simple and quick calculation of the O_4 VCD without the need for constructing full vertical profiles. The parameterization reproduces the real O_4 VCD, as

5   derived from vertically integrated profiles, within $-0.9\,\% \pm 1.0\,\%$ for WRF simulations around Germany, $0.1\,\% \pm 1.2\,\%$ for global reanalysis data (ERA5), and $-0.4\,\% \pm 1.4\,\%$ for GRUAN radiosonde measurements around the world. When applied to measured surface values, uncertainties of $1\,K$, $1\,hPa$, and $16\,\%$ for temperature, pressure, and RH correspond to relative uncertainties of the O_4 VCD of $0.3\,\%$, $0.2\,\%$, and $1\,\%$, respectively. The proposed parameterization thus provides a simple and accurate formula for the calculation of the O_4 VCD which is expected to be useful in particular for MAX-DOAS applications.

10   # 1 Introduction

In the atmosphere, two oxygen molecules can build collision pairs and dimers, which are often denoted as O_4 (Greenblatt et al., 1990; Thalman and Volkamer, 2013, and references therein). O_4 has absorption bands in the UV/visible spectral range, thus O_4 can be retrieved from atmospheric absorption spectra, e.g. by applying Differential Optical Absorption Spectroscopy (DOAS) (Platt and Stutz, 2008). Measurements of the O_4 absorption in scattered light provide information about light path

15   distributions in the atmosphere, for instance allowing to investigate light path increase within clouds (Wagner et al., 1998) or the retrieval of cloud heights from satellite measurements (Acarreta et al., 2004; Veefkind et al., 2016).

For Multi-Axis (MAX) DOAS, i.e. ground-based instruments measuring scattered light at different elevation angles, O_4 measurements provide information on vertical profiles of aerosol extinction (Heckel et al., 2005). Prerequisite for MAX-DOAS profile inversions is knowledge about the O_4 vertical column density (VCD) which provides the link between the

20   measured slant column densities (SCDs) at different viewing angles and the forward modelled SCDs based on radiative transfer calculations. Thus, a wrong input of the O_4 VCD directly affects the resulting aerosol profiles. For the profile inversion algorithm MAPA (Beirle et al., 2019) applied to measurements taken during the CINDI-2 campaign (Kreher et al., 2020), for instance, a change of the input O_4 VCD of $2\,\%$, $5\,\%$, or $10\,\%$ causes changes of the resulting median aerosol optical depth of $6\,\%$, $13\,\%$, or $20\,\%$, respectively. Thus, the O_4 VCD should be determined with accuracy and precision better than about

Atmospheric
Measurement
Techniques
Discussions

25 3%, leaving other sources of uncertainty, i.e. the spectral analysis ($\approx 5\%$) as well as radiative transfer modeling ($\approx 4\%$) (see Wagner et al., 2021, table 3 therein) as the limiting factors in MAXDOAS profile inversions.

The O_4 VCD can be calculated by vertical integration of the O_2 number density profile squared. This requires knowledge of vertical profiles of temperature, pressure, and humidity, e.g. as derived from radiosonde measurements or meteorological models. However, measured profiles are only available for few stations and do not provide continuous temporal coverage,

30 while modelled profiles might not be available in some cases (e.g. during measurement campaigns in remote regions and poor internet connection), or might not reflect the conditions at the measurement site appropriately, in particular in mountainous terrain not resolved by the model.

Measurements of surface air (at $2\,\mathrm{m}$) temperature, pressure, and humidity, on the other hand, are routinely performed by meteorological stations, and could be added to any MAX-DOAS measurement site with relatively low costs and efforts. Wagner

35 et al. (2019) proposed a procedure how to construct full temperature and pressure profiles from the respective surface values by assuming (a) a constant lapse rate of $-6.5\,\mathrm{K}\,\mathrm{km}^{-1}$ from ground up to $12\,\mathrm{km}$, and constant temperature above, and (b) applying the barometric formula. Wagner et al. (2019) estimate the uncertainty of the calculated O_4 VCD to 3% and list the diurnal variation of the surface temperature and the limited representativeness of the surface temperature for the temperature profile above the boundary layer as main source of uncertainty.

40 The method proposed by Wagner et al. (2019) reproduces the true O_4 VCD within about 2% (mean bias) $\pm 2\%$ standard deviation (SD) globally when compared to ECMWF profiles, as shown below. Locally, however, large deviations up to 7% could be found. Main reason for systematic deviations to the true O_4 VCD turned out to be the assumption of a fixed lapse rate of $-6.5\,\mathrm{K}\,\mathrm{km}^{-1}$. While this value reflects typical continental conditions quite well, it is not appropriate in particular over deserts, where lapse rates are stronger (closer to the dry adiabatic lapse rate), and large parts of the oceans with weaker lapse

45 rates (closer to 0 due to condensation).

In this paper we present a simpler approach for the calculation of O_4 VCD just from surface values of temperature and pressure and an a-priori lapse rate based on physical laws, without the need of constructing full profiles. In addition, we provide an empirical parameterization involving surface relative humidity that also accounts for variations of the atmospheric lapse rate. The final equation allows for simple and quick calculation of the O_4 VCD with high accuracy and precision just

50 from surface measurements of temperature, pressure, and relative humidity.

The manuscript derives the formalism of the parameterizations of the O_4 VCD in section 2. In section 3, the datasets used for illustration and quantification of uncertainties are introduced, followed by applications of the O_4 parameterizations in section 4. Important aspects like accuracy/precision, diurnal cycle, or the dependency on surface altitude, are discussed in section 5, followed by conclusions.

2 Formalism

In this section, we provide the formalism for the calculation of O_4 VCDs from surface values of pressure, temperature, and relative humidity.

2.1 Notation

Basic quantities of the derivation below are (a) the number density n, and (b) the vertical column density (VCD) V, i.e. the vertically integrated number density.

The O_4 number density is just defined as the O_2 number density squared. Consequently, the O_4 number density has the unit $\mathrm{molecules}^2\,\mathrm{cm}^{-6}$, and the O_4 VCD has the unit $\mathrm{molecules}^2\,\mathrm{cm}^{-5}$. This matches the common procedure in the DOAS community; the O_4 cross section is given in $\mathrm{cm}^5\,\mathrm{molecules}^{-2}$ accordingly (Greenblatt et al., 1990; Thalman and Volkamer, 2013).

Pressure is denoted by p, temperature by T, and the altitude above sea level by z, while altitude above ground level is denoted by z'. For relative humidity, RH is used in the text as well as in formulas. Surface values are indicated by the subscript "0". Quantities related to O_2 and O_4 are indicated by a respective subscript. For a full list of quantities and symbols see tables 1 and 2.

Table 1. Variables used in this study. A subscript of 0 indicates surface values for n, p, T, z, or RH.

Quantity	Abbreviation	Symbol	Unit
Number density	-	n_{O_2}	$\mathrm{molecules\,cm}^{-3}$
		n_{O_4}	$\mathrm{molecules}^2\,\mathrm{cm}^{-6}$
Vertical column density	VCD	V_{O_2}	$\mathrm{molecules\,cm}^{-2}$
		V_{O_4}	$\mathrm{molecules}^2\,\mathrm{cm}^{-5}$
Pressure	-	p	hPa
Temperature	-	T	K
Altitude above sea level	-	z	m
Altitude above surface	-	z'	m
Effective height	-	h	m
Scale height	-	H	m
Relative humidity	RH	RH	
Effective tropospheric lapse rate	-	Γ	$\mathrm{K\,km}^{-1}$
Relative deviation between of parameterized and true O_4 VCD	-	δ	%
Top of atmosphere (here: highest available profile layer)	TOA	z_{TOA}	m
Total column water vapor	TCWV	V_{H_2O}	$\mathrm{molecules\,cm}^{-2}$

Atmospheric
Measurement
Techniques
Discussions

Table 2. Constants used in this study. Numbers are listed with 6 digits.

Quantity	Symbol	Value	Unit
Gravitational acceleration on Earth	g	9.80665^a	$\mathrm{m\,s^{-2}}$
Molar mass of dry air	M	0.0289655^a	$\mathrm{kg\,mol^{-1}}$
Universal gas constant	R	8.31446^a	$\mathrm{J\,K^{-1}\,mol^{-1}}$
O_2 volume mixing ratio in dry air	ν_{O_2}	0.209392^b	
Combined constants (eq. 10)	C	0.0185646	$\mathrm{K\,Pa^{-2}\,mol^2\,m^{-5}}$
		$6.73266e+39$	$\mathrm{K\,hPa^{-2}\,molecules^2\,cm^{-5}}$

[a] from the Python module MetPy (May et al., 2021)

[b] from Tohjima et al. (2005)

2.2 General approach

70 The VCD V is the vertically integrated number density n:

$$V = \int_{z_0}^{\infty} n(z)\,dz \tag{1}$$

This integral can be re-written as

$$V = n_0 \cdot h, \tag{2}$$

with

75 $$h = \int_{z_0}^{\infty} \frac{n(z)}{n_0}\,dz \tag{3}$$

This effective height h can be understood as the height of the gas column if the gas would be in a homogenous box under surface conditions p_0 and T_0. Note that the effective height equals the scale height H only in case of exponential profiles, i.e. an isothermal atmosphere (see Appendix A).

Thus, the VCDs for O_2 and O_4 can be written as

80 $$V_{O_2} = n_{O_2,0} \cdot h_{O_2} \tag{4}$$

and

$$V_{O_4} = n_{O_4,0} \cdot h_{O_4} \quad = n_{O_2,0}^2 \cdot h_{O_4} \tag{5}$$

Re-arranging eq. 4 for $n_{O_2,0}$ and replacing one $n_{O_2,0}$ term in eq. 5 yields

$$V_{O_4} = V_{O_2} \cdot n_{O_2,0} \cdot \frac{h_{O_4}}{h_{O_2}} \tag{6}$$

85 Hence the O_4 VCD can be expressed as the product of the O_2 VCD, the O_2 surface number density, and the ratio of effective heights of O_2 and O_4 profiles. So far no simplifications or approximations were made.

2.3 O_4 VCD as function of surface pressure, surface temperature, and lapse rate

Based on eq. 6, the O_4 VCD can be related to surface pressure, surface temperature, and lapse rate, if some further assumptions are made:

90 1. Assuming a hydrostatic atmosphere, the O_2 VCD, i.e. the vertically integrated column, is directly related to the surface pressure:

$$V_{O_2} = \frac{\nu_{O_2}}{g \cdot M} \cdot p_0, \tag{7}$$

with ν_{O_2} being the volume mixing ratio of O_2 in dry air, g being the gravitational acceleration on Earth, and M being the molar mass of dry air.

95 2. According to the ideal gas law, the surface number density of O_2 can be expressed as

$$n_{O_2,0} = \frac{\nu_{O_2}}{R} \cdot \frac{p_0}{T_0}, \tag{8}$$

with the universal gas constant R.

3. The ratio of effective heights for O_2 and O_4 depends on the actual profile shape for O_2. For some specific cases, the integral (eq. 3) can be solved analytically, as shown in Appendix A. For an isothermal atmosphere, i.e. an exponential

100 profile of n_{O_2}, the ratio $\frac{h_{O_2}}{h_{O_4}}$ is just 2. For the more realistic assumption of a constant lapse rate Γ, the ratio becomes $2 + \frac{R}{g \cdot M} \Gamma$.

Replacing these terms in eq. 6 yields

$$V_{O_4,\Gamma} = \frac{\nu_{O_2}^2}{R \cdot g \cdot M} \Big/ \left(2 + \frac{R}{g \cdot M} \Gamma\right) \cdot \frac{p_0^2}{T_0}$$

$$= \frac{C}{2 + \frac{R}{g \cdot M} \Gamma} \cdot \frac{p_0^2}{T_0} \tag{9}$$

with

105 $$C = \frac{\nu_{O_2}^2}{R \cdot g \cdot M} \tag{10}$$

combining the constant factors.

Thus with the assumptions specified above, the O_4 VCD is proportional to p_0^2/T_0, with the lapse rate Γ determining the slope.

2.4 O_4 VCD as function of surface pressure, surface temperature, and surface humidity

110 So far, the formalism was based on dry air. Humid air is lighter than dry air, and contains less O_2. Thus, humidity affects the vertical O_2 profile and hence all factors of eq. 6, i.e. the O_2 VCD, the O_2 surface number density, and the effective heights of O_2 and O_4.

Atmospheric
Measurement
Techniques
Discussions
EGU
Open Access

As the vertical humidity profile is generally not well known, these effects cannot be described analytically. In order to still have a simple parameterization of the O_4 VCD based on surface measurements, we follow an empirical approach and introduce a modification of eq. 9 involving surface humidity.

As shown in section 4.2, the O_4 VCD is closely related to the *relative* humidity at ground, while no correlation to *specific* humidity was found. This was surprising on first glance, as the effect of humidity on O_2 number density should be better described by specific humidity. However, RH_0 is closely related to the effective lapse rate of the lower troposphere, which has a strong impact in eq. 9. This will be discussed in more detail in section 4.2.

The parameterization of the O_4 VCD from surface values p_0, T_0, and RH_0 was thus chosen such that a linear function of RH_0 replaces the linear function of Γ in the denominator of eq. 9:

$$V_{O_4,RH} = \frac{C}{a + b \cdot RH_0} \cdot \frac{p_0^2}{T_0},$$ (11)

The parameters were derived as $a = 1.769$ and $b = 0.1257$ by a least squares fit based on ECMWF profiles for 18 June 2018 (see section 4.2). This allows for simple calculation of the O_4 VCD as

$$V_{O_4,RH} = \frac{6.733 \cdot 10^{39}}{1.769 + 0.1257 \cdot RH_0} \cdot \frac{p_0^2}{T_0} \, \text{molec}^2 \, \text{cm}^{-5}.$$ (12)

for RH as dimensionless number (i.e., 0.5 for 50% RH), p_0 in hPa, and T_0 in K.

Note that while this empirical approach basically parameterizes the effective lapse rate by RH_0, also the effect of humid air being lighter is, at least partly, implicitly accounted for by the empirical fit.

2.5 Calculation of the "true" O_4 VCD

In section 4, we investigate the performance of the different parameterizations for the O_4 VCD for modelled and measured profiles. For this purpose, we compare the results of eq. 9 and eq. 12 to the "true" O_4 VCD, which is derived by

(a) calculating the profile of n_{O_2} from profiles of T, p and RH, fully considering the effects of humidity, and

(b) performing the numerical integration (using Simpson's rule) of $n_{O_2}^2$ from surface to top of atmosphere (TOA).

The integration has to be performed up to sufficiently high altitudes (Wagner et al. (2019) recommend $z_{TOA} \geq 30\,\text{km}$) as otherwise the integrated VCD would be biased low due to the missing column above. As not all datasets considered below cover this altitude range, we estimate and correct for the missing O_4 column above the highest profile level by applying eq. 9 for the highest available layer, assuming a lapse rate of zero above. Note that the temperature increase in the upper stratosphere is not relevant here as the contribution to the O_4 VCD above 30 km is negligible. Thus, the "true" O_4 VCD is calculated as

$$V_{O_4,true} = \int_{z_0}^{z_{TOA}} n_{O_2}^2(z)dz + \frac{C}{2} \cdot \frac{p_{TOA}^2}{T_{TOA}}$$ (13)

For z_{TOA} of 20 km, the correction term is of the order of 0.3% of the total O_4 column.

2.6 Comparison of parameterized to "true" O_4 VCD

In order to assess accuracy and precision of the proposed calculation of the O_4 VCD from surface measurements of T_0, p_0 and RH$_0$, we define the relative deviation δ of parameterized O_4 VCDs to the true value:

$$\delta_\Gamma = \frac{V_{O_4,\Gamma} - V_{O_4,\text{true}}}{V_{O_4,\text{true}}} \tag{14}$$

and

$$\delta_{\text{RH}} = \frac{V_{O_4,\text{RH}} - V_{O_4,\text{true}}}{V_{O_4,\text{true}}} \tag{15}$$

Deviations δ_Γ and δ_{RH} are presented below (section 4) as frequency distributions or as mean μ and SD σ.

3 Datasets

We apply the derived formalism to atmospheric datasets for illustration and uncertainty estimates below. For this purpose, we use different datasets:

- Global model data, in order to check for the performance of the parameterizations globally, covering the full range of the input parameter space for surface values of pressure, temperature, humidity, and altitude.

- Regional model data with high spatial resolution, which is also compared to surface stations and allows to investigate diurnal cycles.

- Balloon-borne radiosonde measurements, in order to apply the formalism to high-resolved profile measurements.

Nighttime profiles of T can be considerably different from daytime, in particular in case of temperature inversions (i.e. positive lapse rates) often occurring within the nocturnal boundary layer. For MAX-DOAS measurements, however, these cases are irrelevant. Thus, we consider all atmospheric datasets for daytime conditions only. This is done by selecting data for SZA $< 85°$.

3.1 Global model (ECMWF)

We use global model data as provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) for two purposes:

- In order to investigate global patterns, we use ERA5 reanalysis data (Hersbach et al., 2020) with a truncation at T639 and Gaussian grid N320, corresponding to $\approx0.3°$ resolution. Model output is provided hourly. Here we focus on ERA5 data for 18 June and 18 December 2018, covering the full globe (note that for each day, the polar region of hemispheric winter is not covered due to the SZA selection).

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

- For comparison with radiosonde profiles (see below), we use ERA-Interim reanalysis data with a truncation at T255, corresponding to $\approx 0.7°$ resolution. A preprocessed dataset was created where the 6 hourly model output (0:00, 6:00, 12:00, 18:00 UTC) was interpolated to a regular horizontal grid with a resolution of $1°$. From this, profile data is interpolated to the radiosonde launch in space and time. The reason for also including this rather coarsely resolved model data was that we also use the same dataset and interpolation procedure as default for the extraction of ECMWF profiles at our MAX-DOAS instruments and the calculation of the O_4 VCD within profile inversions with MAPA (Beirle et al., 2019).

3.2 Regional model (WRF-Chem)

We use the Weather Research and Forecasting (WRF) model version 4.2 (Skamarock et al., 2019) for high resolution simulation of meteorological parameters (including T, p and RH) around Germany. A nested domain centred at 49.12°N, 10.20°E was set up in Lambert conformal conic (LCC) projection with coarser domain (d01) at $15 \times 15 \, \text{km}^2$ horizontal resolution and finer domain (d02) at $3 \times 3 \, \text{km}^2$ resolution (fig. B1). Vertically the model extends from surface until $50 \, \text{hPa}$ with 42 terrain following layers in between. The spatial extent of the d01 domain is $4800 \times 3416 \, \text{km}^2$ while that for d02 is $1578 \times 1473 \, \text{km}^2$. The model simulations were set up for a two months period (May&June) in 2018.

We use the ERA5 reanalysis dataset with a horizontal resolution $0.25° \times 0.25°$ and a temporal resolution of 3 hours, downloaded at pressure levels and at the surface for constraining the meteorological initial and lateral boundary conditions. The soil classification, terrain height, and land use patterns were taken from the 21 category Noah-modified IGBP-MODIS land use data.

Here we focus on WRF data for 1-9 May 2018 in the domain d02. The selection of SZA$< 85°$ results in a daily coverage from $6:00 \, \text{h}$ to $17:00 \, \text{h}$ UTC for each day. The vertical profiles reach up to a pressure level of $50 \, \text{hPa}$, corresponding to an altitude of about 20 km. The missing part of the atmosphere contributes about 0.3% to the total O_4 VCD. This effect is considered accordingly in the calculation of the true O_4 VCD (see section 2.5).

3.3 Surface measurements

Germany's National Meteorological Service (Deutscher Wetterdienst, DWD) provides hourly measurements of surface temperature, pressure and relative humidity for a network of ground stations in Germany (Kaspar et al., 2013). Data are provided via the climate data center web interface (CDC-v2.1; https://cdc.dwd.de/portal/). The meteorological measurements are performed in accordance to the guidelines of the world meteorological organization (WMO) to minimize local effects. Additionally we have applied quality control filters such that the parameters QUALITAETS_BYTE (QB) <4 and QUALITAETS_NIVEAU (QN) is either 3 (automatic control and correction) or 7 (second control done, before correction) to only retain the measurements of highest quality.

For this study, we only consider DWD stations providing T_0, p_0, and RH_0 simultaneously, resulting in 206 stations which are displayed in fig. C1. We select measurements for the time period covered by the WRF simulations in order to quantify accuracy and precision of the WRF simulations of surface values. In particular, we investigate how far WRF reflects the diurnal pattern of surface properties.

200   In Appendix D, a comparison of surface values from WRF to the station network is shown, revealing that the surface temperatures modeled by WRF are biased low by $1 \, \mathrm{K}$ on average, while RH surface values are biased high by 7%.

3.4 Radiosonde measurements (GRUAN)

The Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN) is an international reference observing network of sites measuring essential climate variables above Earth's surface (Sommer et al., 2012; Bodeker et al.,
205   2016). Atmospheric profiles of temperature, pressure, and humidity are measured by regular balloon soundings equipped with radiosondes and water vapor measurements (Dirksen et al., 2014). Here we use the RS92 GRUAN Data Product Version 2 (RS92-GDP.2), focusing on certified stations. Vertical profiles and surface values of pressure, temperature and relative humidity are taken directly from the level-2 files for each launch. Further information on the GRUAN stations used in this study are provided in Appendix E.

4   Application to atmospheric datasets

In this section, we apply the parameterizations if the O_4 VCD derived in section 2 to modeled and measured atmospheric datasets. We first apply eq. 9 in section 4.1, discuss the impact of humidity in section 4.2, and apply eq. 12 involving also RH_0 in section 4.3.

4.1 O_4 VCD as function of p_0, T_0, and lapse rate Γ

215   According to eq. 9, the O_4 VCD is proportional to p_0^2/T_0, with the lapse rate Γ determining the slope. We illustrate this correlation for the investigated datasets as shown in Fig. 1.
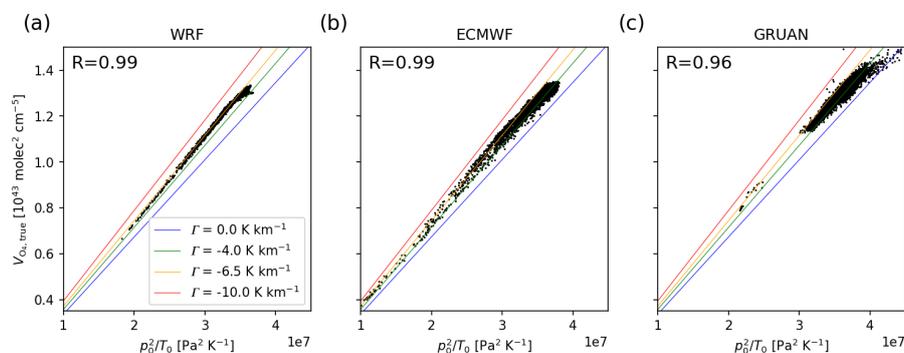

Figure 1. Relation of the O_4 VCD to p_0^2/T_0 and the expected dependency according to eq. 9 for different lapse rates (colored lines) for (a) WRF data on 1 May, 12:00 UTC, (b) ECMWF data on 18 June, 12:00 UTC, and (c) all available GRUAN profiles. For (a) and (b), only 1% of the data is plotted in order to keep the figure readable. Low values correspond to mountaineous sites with low surface pressure. The very high values for GRUAN are observed for the station Barrow (Alaska) for very low temperatures (down to $< 240 \, \mathrm{K}$) in spring.

Atmospheric
Measurement
Techniques
Discussions

Open Access

For all datasets, a very good correlation between p_0^2/T_0 and $V_{O_4, \text{true}}$ (see sect. 2.5) is found, with most datapoints matching to plausible lapse rates in the range of -4 to $-6.5\,\text{K km}^{-1}$. For the WRF simulations for Germany, highest correlation is found, with most data points matching to a lapse rate close to $-6.5\,\text{K km}^{-1}$. ECMWF and GRUAN data show higher variability,

220 as they also cover a wider range of atmospheric conditions. For all datasets, the low values are caused by mountains due to reduced pressure. For the GRUAN measurements, the highest values are observed for Barrow (71.32°N), associated with very cold temperatures in spring.

Wagner et al. (2019) proposed to determine the O_4 VCD based on vertical profiles of T and p constructed from the respective surface values by assuming a constant tropospheric lapse rate of $-6.5\,\text{K km}^{-1}$. We can use eq. 9 for the same purpose, but

225 without the need for constructing full vertical profiles. Note that both methods yield almost the same results, as also the physical assumptions are the same (hydrostatic pressure, ideal gas, dry air, adiabatic lapse rate). The only difference is that Wagner et al. (2019) assumed a tropopause at $12\,\text{km}$, with constant T above, while in the calculation of the ratio of effective heights (see Appendix A), Γ is assumed to be constant throughout the atmosphere, resulting in a small overestimation of eq. 9 of about $0.5\,\%$ compared to the procedure described in Wagner et al. (2019).

230 Figures 2, 3 and 4 display the deviation δ_Γ between parameterized and true O_4 VCD for WRF and ECMWF, respectively, assuming a constant a-priori lapse rate of $-6.5\,\text{K km}^{-1}$.
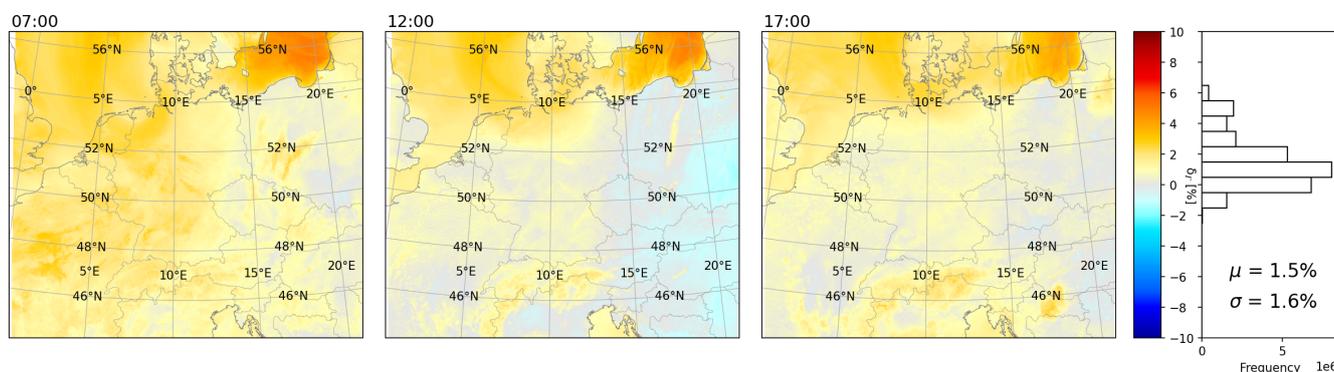

Figure 2. Deviation δ_Γ according to eq. 14 for WRF simulations at 7:00, 12:00, and 17:00 UTC on 1 May 2018. On the right, the frequency distribution of δ_Γ and its mean and SD are given for the WRF simulation period from 1 to 9 May 2018.

Within the WRF domain d02, a generally good agreement between $V_{O_4, \Gamma}$ and $V_{O_4, \text{true}}$ is found (Fig. 2). On average, δ_Γ is $1.5\,\%$, i.e. $V_{O_4, \Gamma}$ are higher than $V_{O_4, \text{true}}$ by $1.5\,\%$. Over land around noon, δ_Γ is close to 0. Over ocean, however, δ_Γ is generally higher (up to $6\,\%$).

235 Also for ECMWF data on 18 June 2018, δ_Γ over Germany is close to 0 (Fig. 3). On global scale, however, only moderate agreement is found between $V_{O_4, \Gamma}$ and $V_{O_4, \text{true}}$, with a mean value of $2.6\,\%$ and $2.8\,\%$ for δ_Γ in June and December, respectively. High values for δ_Γ are found generally over ocean. For continents, δ_Γ is closer to 0, but particularly over deserts, negative values are observed. If the O_4 VCD is calculated as proposed by Wagner et al. (2019), the deviations show the same
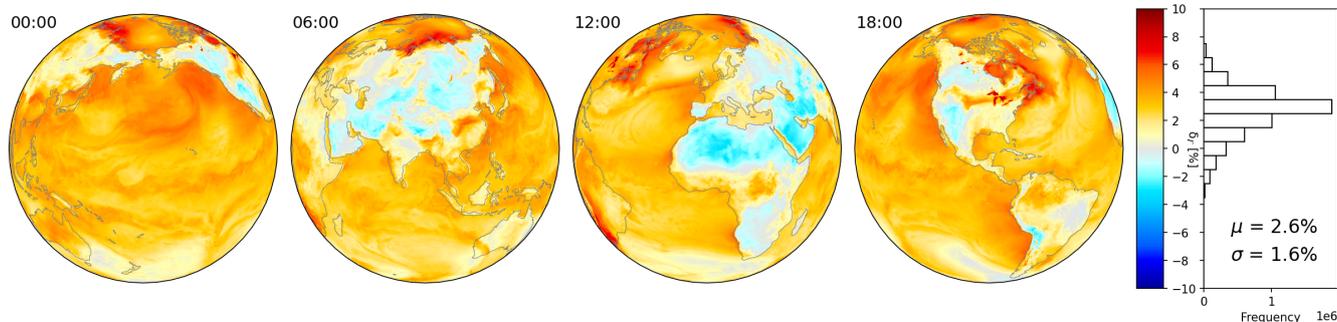
Figure 3. Deviation δ_Γ according to eq. 14 for ECMWF at 0:00, 6:00, 12:00 and 18:00 UTC on 18 June 2018. The projection focuses on daytime for each timestep. On the right, the frequency distribution and mean and SD are given for all hourly outputs from 18 June.
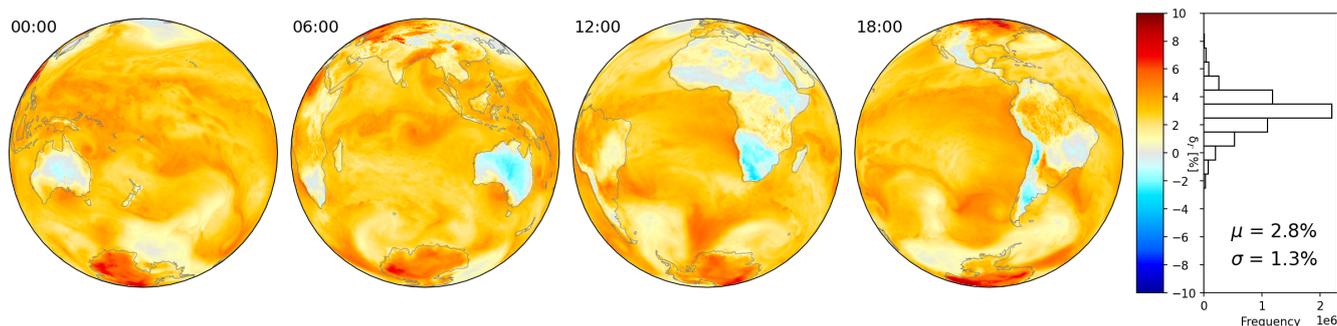

Figure 4. Deviation δ_Γ according to eq. 14 for ECMWF at 0:00, 6:00, 12:00 and 18:00 UTC on 18 December 2018. The projection focuses on daytime for each timestep. On the right, the frequency distribution and mean and SD are given for all hourly outputs from 18 December.

patterns, with slightly lower means (due to consideration of the tropopause) but same SD. These general patterns of systematic

240 deviations from 0 are mainly caused by the simple assumption of a globally constant lapse rate in the calculation of δ_Γ.

4.2 Effects of humidity

Humid air is lighter than dry air. This is considered in the calculation of $V_{O_4,\,true}$. For $V_{O_4,\,\Gamma}$, however, dry air is assumed in the derivation in section 2. One would thus expect that the observed deviation δ_Γ is affected by humidity. However, when comparing δ_Γ to *specific* humidity at surface, we found no correlation (Fig. 5 (a)). We also compared δ_Γ to the total column

245 of water vapor (TCWV), i.e. the vertically integrated water vapor number density. The reason for choosing this quantity was that it (a) represents the total amount of water vapor rather than just the surface value and (b) could also be derived from MAX-DOAS measurements directly. But again, we found no correlation (Fig. 5 (b)).

Instead, we observed a very good correlation between *relative* humidity and δ_Γ (Fig. 5 (c)). This cannot be explained by the impact of humidity on air density, as this is a direct function of specific humidity. But RH at surface is closely related to the

11

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

250   effective lapse rate: for ascending air, RH_0 determines the lifted condensation level (LCL) (Lawrence, 2005; Romps, 2017): the lower RH_0, the higher the LCL, with dry adiabatic lapse rates below. For tropical deserts, on the other hand, which are affected by large-scale subsidence, no condensation takes place (i.e. the dry adiabatic lapse rate applies), and RH at ground is very low due to the adiabatic heating of the descending air masses. Thus, the main systematic deviations seen in δ_Γ are caused by the simple assumption of a constant lapse rate of -6.5 $\mathrm{K\,km^{-1}}$, while actual effective lapse rates are far stronger (more negative)

255   over deserts, with low RH_0. Over most parts of the ocean, on the other hand, RH_0 is high, and the effective lapse rate is weaker (closer to zero).
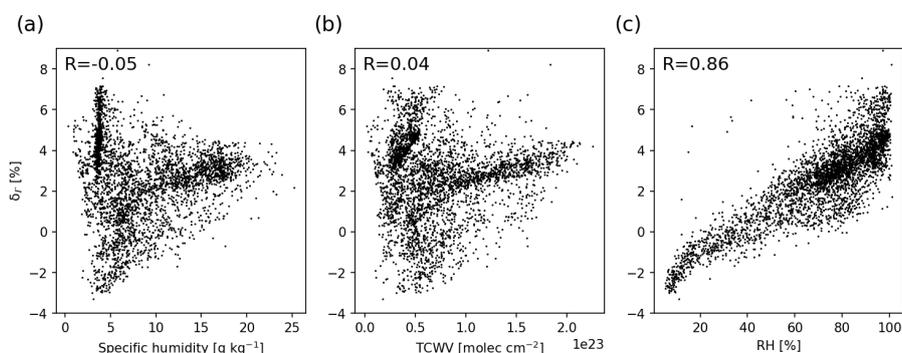

Figure 5. Deviation δ_Γ according to eq. 14 as function of (a) specific humidity at surface, (b) total column water vapor, and (c) relative humidity at surface for ECMWF data on 18 June 2018, 12:00 UTC. Only 1% of the data points are plotted in order to keep the figure readable. Correlation coefficients are given in the respective subplots.

We make use of the good correlation of δ_Γ to RH_0 in order to construct an empirical parameterization according to eq. 12. For this purpose, the parameters a and b were determined by a linear least squares fit (after re-arranging eq. 11 for $a + b \cdot RH_0$) to global ECMWF data for 18 June 2018.

260   ## 4.3 O_4 VCD as function of p_0, T_0, and RH_0

With eq. 12, an empirical parameterization of the O_4 VCD was derived based on surface values of temperature, pressure, and relative humidity. We applied this parameterization to all investigated datasets. Figures 6, 7 and 8 display δ_{RH} for WRF and ECMWF, respectively. GRUAN results are listed in table 3.

For the WRF domain d02, δ_Γ was already quite close to 0 (mean $\delta_\Gamma = 1.5\%$). δ_{RH} is closer to 0, but now showing a slight

265   negative bias (mean $\delta_{RH} = -0.9\%$). Variability has reduced considerably (SD of δ_{RH} is 1.0%, compared to 1.6% for δ_Γ). δ_{RH} shows a weaker land-ocean contrast. Over the Alps, δ_{RH} is biased low (down to -3%).

For ECMWF, the parameterization involving RH is a substantial improvement compared to the results for δ_Γ. For 18 June 2018, the mean of $\delta_{RH} = 0.0\%$ is of course a consequence of the fit optimizing a and b which is based on the same ECMWF dataset. But there is also a considerable reduction of SD from 1.6% for δ_Γ to 1.0% for δ_{RH}. Land-ocean contrasts are sup-
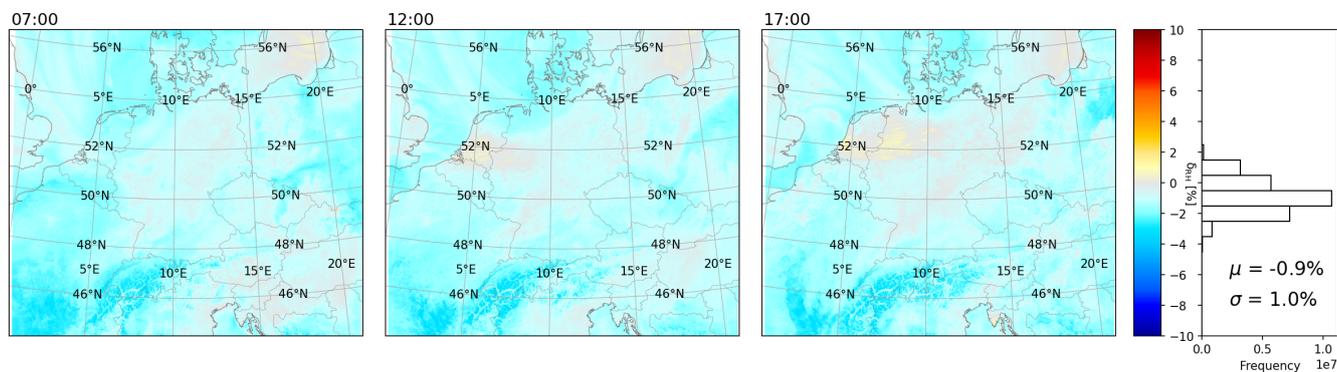
Figure 6. Deviation δ_{RH} according to eq. 15 for WRF simulations at 7:00, 12:00, and 17:00 UTC on 1 May 2018. On the right, the frequency distribution and mean and SD are given for the WRF simulation period from 1 to 9 May.
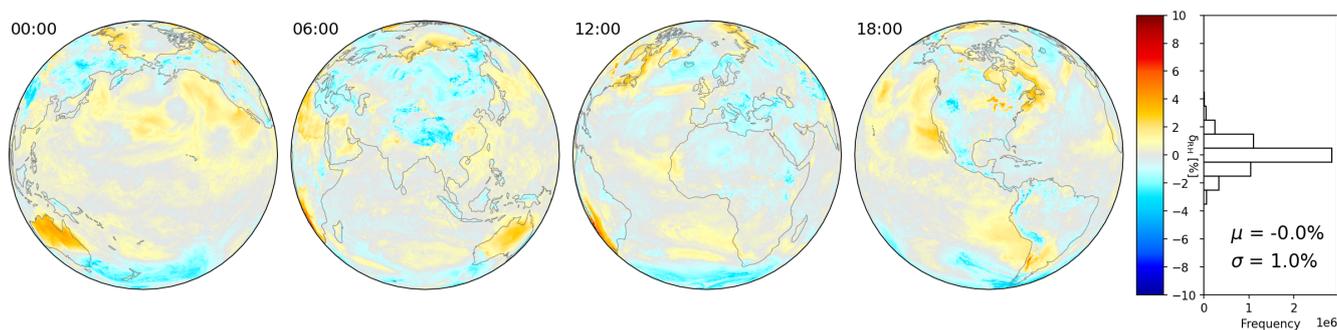

Figure 7. Deviation δ_{RH} according to eq. 15 for ECMWF at 0:00, 6:00, 12:00 and 18:00 UTC on 18 June 2018. The projection focuses on daytime for each timestep. On the right, the frequency distribution and mean and SD are given for all hourly outputs from 18 June.
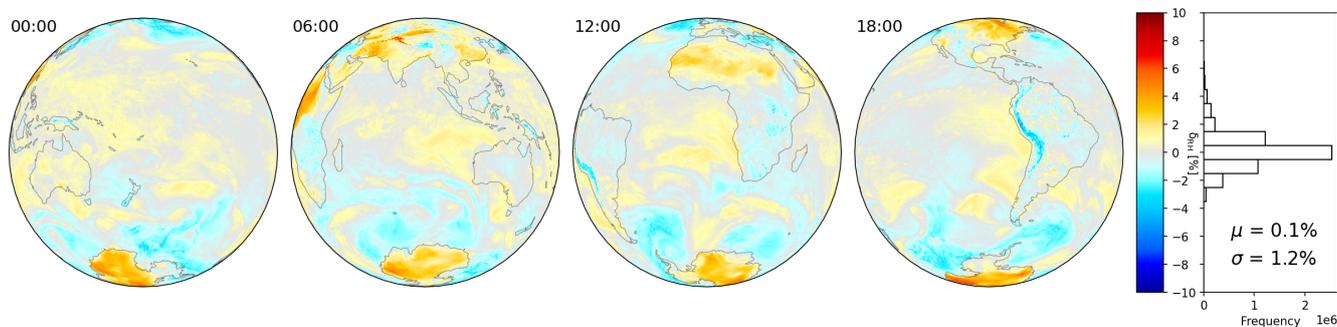

Figure 8. Deviation δ_{RH} according to eq. 15 for ECMWF at 0:00, 6:00, 12:00 and 18:00 UTC on 18 December 2018. The projection focuses on daytime for each timestep. On the right, the frequency distribution and mean and SD are given for all hourly outputs from 18 December.

270   pressed, but are still visible for some coastlines like the West coasts of North and South America. The low values over deserts observed for δ_Γ are largely improved for δ_{RH}.

For December, mean δ_{RH} (based on the parameterization optimized for June) is close to 0 as well (0.1%). SD improved slightly (SD = 1.3% for δ_Γ vs. 1.2% for δ_{RH}).

Remaining systematic deviations in the maps of δ_{RH} are basically due to

275   – weather, for instance associated with low pressure or frontal systems. This reflects the simplifying assumptions made, in particular assuming hydrostatic conditions in section 2. Note, however, that weather conditions associated with rain and clouds are usually not considered in MAX-DOAS retrievals.

– mountains, which tend to show systematic deviations δ_{RH} that are mostly negative (e.g. over the Andes), but sometimes also positive (e.g. over Antarctica). For further discussion see sect. 5.3.

280   – patterns of enhanced δ_{RH} at the edge of the maps for ECMWF, corresponding to high solar zenith angles.

So far, the formalism derived in section 2 was applied to data from meterological models. Now we test it for radiosonde profiles as well. Application of eq. 12 to GRUAN data yields low deviations between parameterized and true O_4 VCDs close to 0 for all stations, as listed in table 3. Overall, the mean deviation δ_{RH} of all considered GRUAN profiles is -0.4%, with a SD of 1.4%. For 11 out of the 17 stations, the mean agreement is within 1%.

285   Largest deviations are found for La Reunion, where $V_{O_4, RH}$ is biased low by -2.5%. This is probably related to the altitude of this station of more than $2\,km$ on a remote island in the Indian ocean.

Highest positive deviation of 1.0% is found for Barrow, with also highest SD of 1.8%. Closer inspection revealed that for Barrow, the high SD is mainly caused by some very high values during spring where surface temperatures are very low ($< 240\,K$) and temperature inversions occur (i.e. lapse rates are positive in the boundary layer).

5   Discussion

5.1   Accuracy and precision

In eq. 12, we provide a formula for the calculation of the O_4 VCD. Accuracy and precision of the resulting V_{O_4} thereby depend on accuracy and precision of (1) the chosen parameterization and (2) surface values p_0, T_0 and RH_0.

1. We estimate overall accuracy and precision of eq. 12 to $< 1\%$ and $< 2\%$ based on mean and SD of deviations between
295   parameterized and true O_4 VCD for WRF, ECMWF and GRUAN data as presented above. For high SZA as well as for mountaineous regions (see also section 5.3), uncertainties can be larger up to about 3%.

2. Application of eq. 12 requires surface measurements of p_0, T_0, and RH_0. Uncertainties of temperature and pressure are rather uncritical, as an error of $1\,K$ and $1\,hPa$ for T_0 and p_0 would correspond to an error of 0.3% and 0.2% in $V_{O_4, RH}$, respectively. In order to reach an accuracy/precision of 1%, the corresponding errors of RH_0 have to be lower

Table 3. Deviations δ_{RH} for GRUAN stations. Additional information on GRUAN stations is given in Appendix E. In the last column, also the deviation δ_{ECMWF}, i.e. the relative difference between integrated O_4 VCDs based on ECMWF profiles (interpolated in space and time to the radiosonde data) and GRUAN profiles is shown. The last row (all) is the mean of all available profiles and is thus reflecting conditions of the stations with high number of radiosonde measurements.

Station	δ_{RH} [%]	δ_{ECMWF} [%]
Barrow	1.0 ± 1.8	-0.1 ± 4.8
Beltsville	-0.3 ± 0.5	-0.6 ± 2.5
Boulder	-1.4 ± 0.8	-9.5 ± 2.7
Cabauw	-1.2 ± 1.0	-0.3 ± 3.9
Darwin	0.3 ± 0.2	-0.8 ± 0.6
Graciosa	-1.2 ± 0.8	-0.6 ± 3.5
Lauder	-1.3 ± 1.0	-6.9 ± 3.0
Lindenberg	-1.0 ± 1.0	-0.6 ± 3.7
Manus	-0.2 ± 1.1	-1.1 ± 0.8
Nauru	-0.9 ± 0.3	-1.0 ± 0.5
NyAlesund	-0.0 ± 1.0	-3.5 ± 3.2
Payerne	-0.7 ± 1.4	-10.0 ± 2.8
LaReunion	-2.5 ± 0.5	-3.4 ± 5.6
Lamont	-0.1 ± 1.3	-1.2 ± 4.3
Sodankyla	-0.9 ± 1.2	-1.3 ± 4.4
Tateno	-0.1 ± 0.9	-2.5 ± 3.9
Tenerife	-0.6 ± 0.7	-0.5 ± 2.7
all	-0.4 ± 1.4	-1.2 ± 4.2

300 than 16%. These limits should be achievable for adequate meteorological instrumentation and a measurement procedure following WMO guidelines. In particular, surface temperature should be measured at about $1.25 - 2\,\mathrm{m}$ above ground using a radiation shield (WMO, 2018).

The proposed parameterization thus allows to calculate O_4 VCD with accuracy and precision sufficient for applications in MAX-DOAS profile inversions. Interestingly, the default procedure used in MAPA, i.e. integrating the O_4 VCD based 305 on ERA-Interim profiles pre-gridded on $1°$ resolution, results in larger deviations (in particular larger SD) when applied to GRUAN profiles (see table 3). We thus consider the proposed parameterization as useful approach for determining the O_4 VCD even for cases where model profiles are available.

Atmospheric
Measurement
Techniques
Discussions

Open Access
EGU

5.2 Diurnal cycles

Surface conditions can change rapidly, e.g. in case of passing frontal systems or storm tracks. For such rapid changes, the change of the true O_4 VCD might not be adequately represented by the change of $V_{O_4, \text{RH}}$. These effects are reflected in the SD of deviations δ_{RH} for ECMWF, WRF, and GRUAN.

In addition, surface values could change *systematically* during the day in case of strong solar irradiation, causing a diurnal cycle of the O_4 VCD (Wagner et al., 2019). Thus we investigate the diurnal cycles of T_0, p_0, RH_0, and the respective O_4 VCDs $V_{O_4, \text{RH}}$ and $V_{O_4, \text{true}}$ in more detail, and investigate how far (a) the WRF simulations reflect the actual diurnal cycles and (b) the parameterized O_4 VCD based on surface values reflect the diurnal cycle of the true O_4 VCD. For this we extract the WRF simulations at the locations of the DWD ground station network. In order to focus on strong diurnal patterns, we select days where the change of surface temperature exceeds $10\,\text{K}$ for each station.

Fig. 9 displays the diurnal cycles of surface properties and O_4 VCDs for WRF and DWD station data. Overall, the diurnal cycle simulated by WRF matches the patterns measured by the surface stations quite well, and $V_{O_4, \text{RH}}$ is almost the same for WRF and ground stations. Surface pressure changes only slightly over the day; the systematic decrease is of the same magnitude as the respective standard mean error for each hour of the day of about $2.5\,\text{hPa}$. But surface temperature increases by $10.5\,\text{K}$ from morning to evening due to the selection of days with strong diurnal cycle in T_0. As $V_{O_4, \text{RH}}$ is reciprocal to T, this alone would correspond to a change of $V_{O_4, \text{RH}}$ of $3.5\,\%$. However, at the same time, RH decreases by about $30\,\%$, which has an opposite effect on $V_{O_4, \text{RH}}$. Consequently, the diurnal cycle of $V_{O_4, \text{RH}}$ is only moderate (about $2\,\%$ decrease from morning to evening).

The true O_4 VCD, as derived from the integrated WRF profiles, also decreases over the day, and $V_{O_4, \text{true}}$ follows nicely $V_{O_4, \text{RH}}$ in the afternoon. In the morning, however, $V_{O_4, \text{RH}}$ is higher compared to noon by $1.4\,\%$, while $V_{O_4, \text{true}}$ is only $0.7\,\%$ higher. This deviation between parameterized and true O_4 VCD indicates that in the early morning, surface measurements are not as useful for determining the full column, which is probably related to remainders of the nocturnal boundary layer which often has atypical lapse rates due to temperature inversions.

But even during morning hours, the systematic error made by $V_{O_4, \text{RH}}$ is relatively small, at least for the investigated time period for Germany. But also for the global ECMWF analysis, the impact of diurnal cycles on the calculation of the O_4 VCD is only moderate; otherwise, Figures 7 and 8 would show systematic East-West gradients.

Thus, the parameterization of eq. 12 also reflects the diurnal cycle of the O_4 VCD sufficiently.

5.3 Dependency on surface altitude

The empirical parameterization eq. 12 works generally well, but is of course not perfect. Remaining patterns in the maps of δ_{RH} show weather patterns like low pressure systems, but also some systematic effects. In particular mountains can be recognized in Figures 6, 7 and 8. We thus investigate a possible relation between surface altitude and δ_{RH} for all investigated datasets (Fig. 10).

Atmospheric
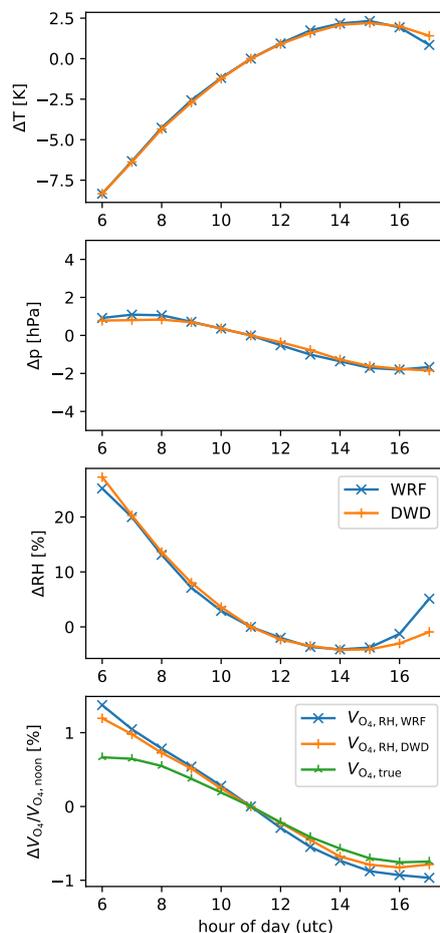Measurement
Techniques
Discussions

Open Access

EGU

Figure 9. Diurnal cycles of surface temperature (a), pressure (b), RH (c), and the O_4 VCD (d). Data points show the mean values for all stations for the considered time period 1-9 May 2018 for days where increase in T_0 over the day is larger than $10\,\mathrm{K}$. For better comparison, all cycles are referred to the mean value at 11 UTC (around solar noon for Germany). For the O_4 VCD, the relative change is shown.

340   For the WRF simulations for the domain d02, the Alps can be clearly recognized in Fig. 6, with mountains showing lower values of δ_{RH}. This can also be clearly seen in the scatter plot in Fig. 10 (a), where surface altitude and δ_{RH} are anticorrelated with R$=-0.53$, and a decrease of δ_{RH} of roughly $1\,\%$ per km. For GRUAN stations (c), results are similar, but statistics are poor, and the correlation coefficient is low, as only two stations (Boulder and La Reunion) are available with a surface altitude above $1\,\mathrm{km}$.

345   For ECMWF, however, results are not at all as clear as those for WRF. The correlation coefficient is close to zero. For altitudes between 2 and $3\,\mathrm{km}$, it looks like δ_{RH} is increasing rather than decreasing with altitude. And for very high surface altitudes as found over the Himalaya, δ_{RH} is still close to 0 and would not match the slope of $1\,\%$ per km derived for WRF.

Atmospheric
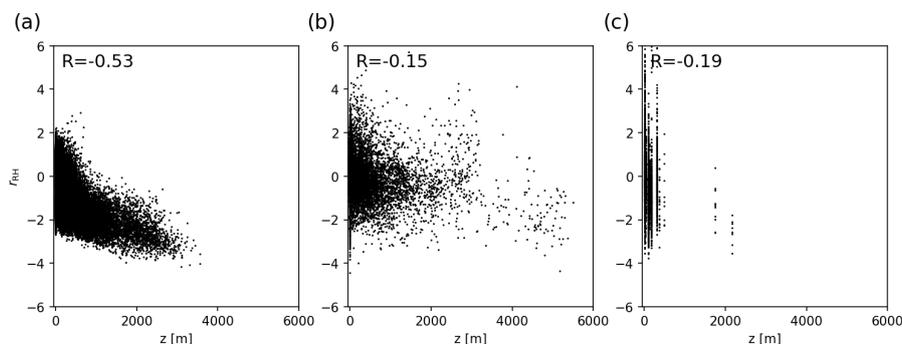Measurement
Techniques
Discussions

Open Access

Figure 10. Dependency of δ_{RH} on surface altitude for (a) WRF , (b) ECMWF, and (c) GRUAN data. For (a) and (b), only 0.2% of the data points are plotted in order to keep the figure readable.

Especially for 18 December 2018, it can clearly be seen that the impact of surface altitude is ambiguous (Fig. 8): While deviations over the Andes are strongly negative, they are positive over the Himalayas as well as over Antarctica.

350 The reason for the poor correlation between z_0 and δ_{RH} for ECMWF is not clear to us. Obviously, also other factors would probably have to be considered (season, SZA). But since there is no clear correlation, and a quantitative correction would rather worsen δ_{RH} instead of improving it for several mountain areas around the globe, we decided not to apply an explicit correction for surface altitude.

Consequently, the parameterization of eq. 12 has higher uncertainties up to about 3% when applied for mountainous sites.

355 **5.4 Application for MAX-DOAS profile inversions based on optimal estimation**

For profile inversion schemes based on profile parameterizations, like MAPA (Beirle et al., 2019), the O_4 VCD is needed in order to convert the measured SCDs to AMFs. For inversion schemes based on optimal estimation, however, vertical profiles of T and p are required for the online RTM calculations. For this case, we propose to extrapolate profiles of T and p from surface values as proposed in Wagner et al. (2019), but not with a constant lapse rate. Instead, the effective lapse rate should be 360 determined from surface RH according to equations 9 and 11:

$$2 + \frac{R}{g \cdot M}\Gamma = a + b \cdot \text{RH}_0 \tag{16}$$

and thus

$$\Gamma = (a - 2 + b \cdot \text{RH}_0) \cdot \frac{g \cdot M}{R}$$
$$= (-0.2308 + 0.1257 \cdot \text{RH}_0) \cdot 34.16 \, \text{K km}^{-1} \tag{17}$$

For RH_0 of 0%, 50%, and 100%, the corresponding effective lapse rate results in -7.89, -5.74, and $-3.59 \, \text{K km}^{-1}$, respec-365 tively.

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

5.5 Lapse rate from direct sun measurements of O_2 and O_4

Eq. 6 relates the O_4 VCD to the ratio of effective heights for O_2 and O_4, which can be expressed by the effective atmospheric lapse rate (see Appendix A). This formalism might also be used in the other direction: from total column measurements of O_2 and O_4 by direct sun observations, an effective atmospheric lapse rate can be derived:

370
$$2 + \frac{R}{g \cdot M} \cdot \Gamma = \frac{h_{O_2}}{h_{O_4}} = \frac{V_{O_2}}{V_{O_4}} \cdot n_{O_2,0} = \frac{S_{O_2}}{S_{O_4}} \cdot \frac{\nu_{O_2} \cdot p_0}{R \cdot T_0} \tag{18}$$

and thus

$$\Gamma = \left(\frac{S_{O_2}}{S_{O_4}} \cdot \frac{\nu_{O_2} \cdot p_0}{R \cdot T_0} - 2 \right) \cdot \frac{g \cdot M}{R} \tag{19}$$

with S being the slant column of the direct sun measurement. For direct sun measurements, the ratio between slant and vertical column is a simple function of the SZA and is the same for O_2 and O_4.

375 Even for limited accuracy of column measurements of O_2 and O_4, this would allow to derive time series of an effective lapse rate, reflecting the state of the lower atmosphere.

6 Conclusions

The O_4 VCD can be expressed in terms of surface pressure and temperature based on physical laws, if a constant lapse rate is assumed, without the need for constructing full vertical profiles. With an empirical correction which basically parameterizes

380 the effective lapse rate as linear function of surface RH, we could present a formula for simple and quick calculation of the O_4 VCD based on p_0, T_0, and RH_0. This parameterization reproduces the real O_4 VCD, as derived from vertically integrated profiles, within $-0.9\% \pm 1.0\%$ for WRF simulations around Germany, $0.1\% \pm 1.2\%$ for global reanalysis data (ERA5), and $-0.4\% \pm 1.4\%$ for radiosonde soundings around the world. Uncertainties over mountains are generally larger (up to about 3%). For applications to measured surface values, uncertainties of $1\,K$, $1\,hPa$, and 16% for temperature, pressure, and RH

385 correspond to relative uncertainties of the O_4 VCD of 0.3%, 0.2%, and 1%, respectively.

This accuracy and precision is sufficient for application in MAX-DOAS profile inversions. Moreover, the parameterization reflects the true O_4 VCD, as derived from radiosonde measurements, even better (in particular in terms of SD) than the standard approach we used so far for MAPA based on interpolated model data. We thus recommend to equip each MAX-DOAS measurement station with state-of-the-art thermometer, barometer, and hygrometer.

390 *Code availability.* A Python implementation of the derived functions for the calculation of the O_4 VCD is provided in the Supplementary material.

Atmospheric
Measurement
Techniques
Discussions

Appendix A: Ratio of effective heights

The ratio of the effective heights for O_2 and O_4 depends on the shape of the O_2 profile. For specific shapes the ratio can be calculated explicitly. Here, vertical integration is performed to infinity. Below, we derive the ratio $\frac{h_{O_2}}{h_{O_4}}$, which allows for simpler

395   notation avoiding compound fractions. For application in eq. 6, the inverse ratio has to be taken.

A1   Isothermal atmosphere

For the simple assumption of a barometric pressure profile with constant T, the O_2 number density decreases exponentially with altitude:

$$n_{O_2} = n_{O_2,0} \cdot \exp\left(-z'/H\right) \tag{A1}$$

400   with the scale height H. In this case, the integral of eq. 3 directly yields H, i.e. the effective height equals the scale height for exponential profiles. For O_4, the profile is exponentially decreasing as well, with the scale height being half of that for O_2. Thus, for O_2 profiles declining exponentially with z, the ratio of effective heights is just

$$\frac{h_{O_2}}{h_{O_4}} = 2. \tag{A2}$$

A2   Polytropic atmosphere

405   If the temperature is changing linearly with altitude, i.e. the dependence of $T(z) = T_0 + \Gamma \cdot (z - z_0)$ is described by a constant lapse rate Γ, the resulting profile of O_2 follows a power function:

$$n_{O_2} = n_{O_2,0} \cdot \left(1 + \frac{\Gamma}{T_0} z'\right)^{-\alpha}, \tag{A3}$$

with

$$z' = z - z_0 \tag{A4}$$

410   being altitude above surface, and

$$\alpha = 1 + \frac{g \cdot M}{R \cdot \Gamma} \tag{A5}$$

being the constant exponent.

Integration of eq. 3 yields

$$
\begin{aligned}
h_{O_2} &= \int_0^\infty \left(1 + \frac{\Gamma}{T_0} z'\right)^{-\alpha} dz' \\
&= \left[\frac{1}{-\alpha+1}\left(1 + \frac{\Gamma}{T_0} z'\right)^{-\alpha+1} \cdot \frac{T_0}{\Gamma}\right]_0^\infty \\
&= \frac{1}{-\alpha+1} \cdot \frac{T_0}{\Gamma}
\end{aligned}
\tag{A6}
$$

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

415   For O_4, the number density profile is

$$n_{O_4} = n_{O_4,0} \cdot \left(1 + \frac{\Gamma}{T_0} z'\right)^{-2\alpha}, \qquad (A7)$$

and thus

$$h_{O_4} = \frac{1}{-2\alpha + 1} \cdot \frac{T_0}{\Gamma} \qquad (A8)$$

The ratio of effective heights can then be calculated as

$$
\begin{aligned}
\frac{h_{O_2}}{h_{O_4}} &= \frac{2\alpha - 1}{\alpha - 1} \\
&= \frac{2\frac{g \cdot M}{R \cdot \Gamma} + 1}{\frac{g \cdot M}{R \cdot \Gamma}} \\
&= 2 + \frac{R}{g \cdot M} \cdot \Gamma. \qquad (A9)
\end{aligned}
$$

420

For a lapse rate of 0 this equals the result for exponential profile (=2). For a typical lapse rate of e.g. -6.5 K/km, the ratio of effective heights is 1.81.

Note that for solving the integral in eq. A6 analytically, a constant lapse rate has to be assumed throughout the atmosphere, while in reality, the temperature profile is far more complex. For the calculation of the O_4 VCD, however, the troposphere, 425   where the assumption of a constant lapse rate is appropriate, contributes more than 95% of the total column. For the column above the tropopause, the assumption of a constant lapse rate causes an overestimation. In terms of the total O_4 VCD, results based on eq. 9 are biased high by about 0.47% compared to the respective VCDs calculated by the method described in Wagner et al. (2019), assuming constant temperature above 12 km. This effect is quite small and thus neglected in eq. 9. For the empirical correction in 12, however, this effect is corrected implicitly.

430   **A3   Real atmosphere**

For real atmospheric conditions, the lapse rate can generally not be considered to be constant. However, the ratio of effective heights can still be described by eq. A9 if an *effective* lapse rate is considered:

$$\Gamma_{\text{eff}} = \left(\frac{h_{O_2}}{h_{O_4}} - 2\right) \cdot \frac{g \cdot M}{R} \qquad (A10)$$
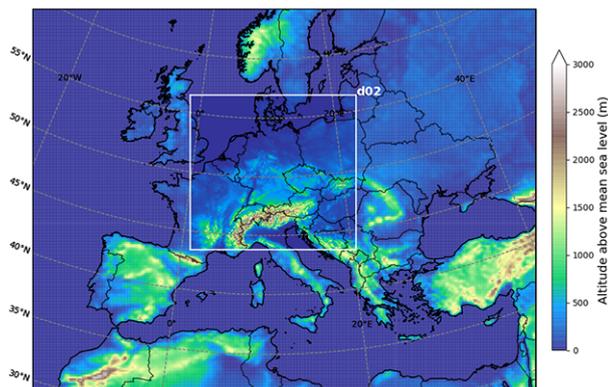
Appendix B: WRF model domains

Figure B1. Nested model domains used for the WRF simulations.
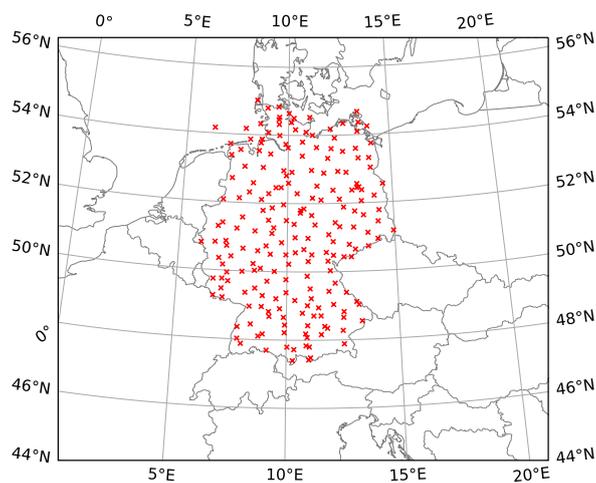
435 **Appendix C: DWD stations**

Figure C1. Location of the 206 DWD ground stations providing simultaneous measurements of surface values of T, p and RH during 1-9 May 2018.

Appendix D: Validation of surface values from WRF

We use the DWD network of surface stations for investigating the accuracy and precision of the WRF simulations. Fig. D1 displays correlations between surface values from the DWD station network and the respective WRF simulations. For this purpose, each station is associated with the nearest neighbor from the WRF simulation. We do not interpolate the WRF data as

440 we still want to compare the parameterized O_4 VCD with the true VCD derived from vertical integration of the WRF profiles.

Surface altitude (a) is lower in the gridded elevation map used as input in the WRF simulations by $20\,\mathrm{m}$ on average, and by almost $1\,\mathrm{km}$ for the station on Germany's highest mountain Zugspitze. This is a consequence of the spatial resolution of the WRF simulations of $1\,\mathrm{km}$ not resolving single mountains. The systematic negative bias of WRF surface altitude indicates that the DWD stations tend to be located on hill and mountain tops.

445 This difference in altitude would directly affect the comparisons of T and particularly p. Thus, we apply a simple correction of station values and extrapolate them to the respective WRF surface altitude assuming a lapse rate of $-6.5\,\mathrm{K\,km^{-1}}$. For RH, no correction is applied.

The reason for keeping the WRF values and adjusting the station data is that for WRF we have the full vertical profile and can calculate the true O_4 VCD according to eq. 13.
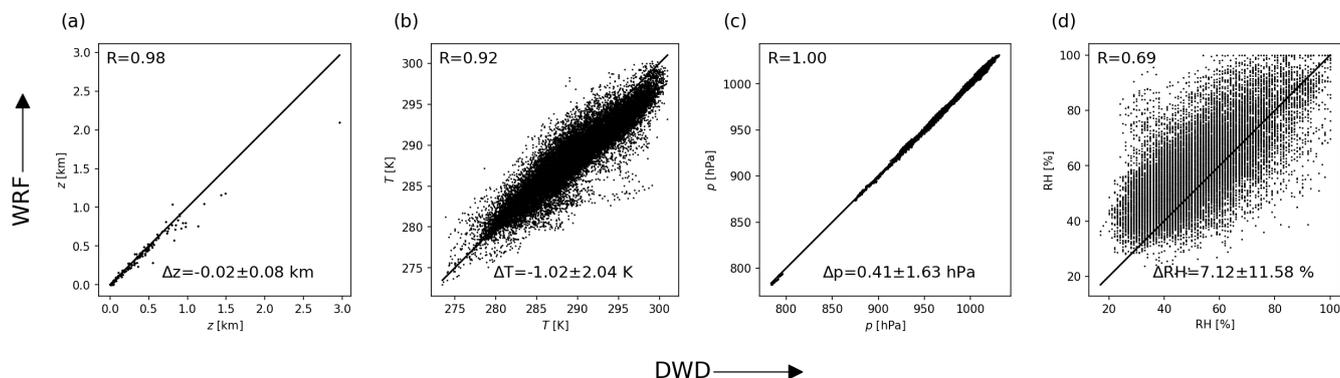

Figure D1. Comparison of WRF surface values to DWD ground stations. For T and p, station values are adjusted to the mean altitude of the respective gridded elevation map used as input for WRF siumulations (see text for details).

Appendix E: GRUAN stations

Fig. E1 displays the location of the available GRUAN stations. Table E1 lists the stations, including their full name, and provides information on latitude, longitude, altitude, and the number of available profiles with SZA$< 85°$.
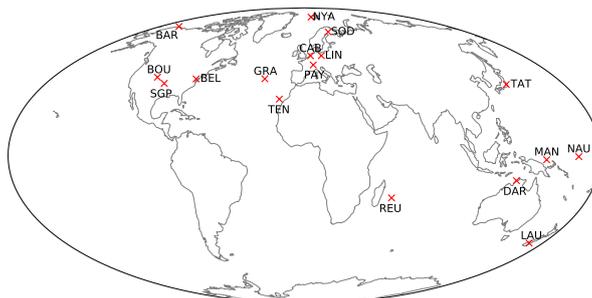

Figure E1. Location of GRUAN stations considered in this study. For station names and further details see table E1.

Table E1. List of GRUAN stations and number of available sonde flights (only considering SZA$<85°$) used in this study.

Label	Name	Lat [°N]	Lon [°E]	z_0 [m]	Profiles
BAR	Barrow	71.32	-156.62	8	1855
BEL	Beltsville	39.05	-76.88	53	93
BOU	Boulder	39.95	-105.20	1743	128
CAB	Cabauw	52.10	5.18	1	381
DAR	Darwin	-12.42	130.89	35	4
GRA	Graciosa	39.09	-28.03	30	417
LAU	Lauder	-45.05	169.68	371	203
LIN	Lindenberg	52.21	14.12	103	4997
MAN	Manus	-2.06	147.43	4	67
NAU	Nauru	-0.52	166.92	7	29
NYA	NyAlesund	78.92	11.92	15	1915
PAY	Payerne	46.81	6.95	491	59
REU	LaReunion	-21.08	55.38	2156	8
SGP	Lamont	36.61	-97.49	315	3368
SOD	Sodankyla	67.37	26.63	179	1262
TAT	Tateno	36.06	140.13	30	589
TEN	Tenerife	28.32	-16.38	121	935

Author contributions. CB initiated this study by proposing to express the O_4 VCD by surface number density and column density of O_2. VK performed the WRF simulations. SD, CB and TW provided input on O_4 VCD calculation and meteorology. SB developed the full formalism,

455 performed the intercomparisons to external datasets, and wrote the manuscript, with input and feedback from all co-authors.

Competing interests. None.

References

Acarreta, J. R., de Haan, J. F., and Stammes, P.: Cloud pressure retrieval using the O2-O2 absorption band at 477 nm, J. Geophys. Res., 109, D05204, doi:10.1029/2003JD003915, 2004.

Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y., and Wagner, T.: The Mainz profile algorithm (MAPA), Atmos. Meas. Tech., 12, 1785–1806, https://doi.org/10.5194/amt-12-1785-2019, 2019.

Bodeker, G. E., Bojinski, S., Cimini, D., Dirksen, R. J., Haeffelin, M., Hannigan, J. W., Hurst, D. F., Leblanc, T., Madonna, F., Maturilli, M., Mikalsen, A. C., Philipona, R., Reale, T., Seidel, D. J., Tan, D. G. H., Thorne, P. W., Vömel, H. and Wang, J.: Reference Upper-Air Observations for Climate: From Concept to Reality, Bulletin of the American Meteorological Society, 97(1), 123–135, doi:10.1175/BAMS-D-14-00072.1, 2016.

Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, Atmos. Meas., Tech., 7, 4463–4490, https://doi.org/10.5194/amt-7-4463-2014, 2014.

Greenblatt, G. D., Orlando, J., Burkholder, J. B., and Ravishankara, A. R.: Absorption measurements of oxygen between 330 and 1140 nm, J. Geophys. Res., 95, 18577–18582, 1990.

Heckel, A., Richter, A., Tarsu, T., Wittrock, F., Hak, C., Pundt, I., Junkermann, W., and Burrows, J. P.: MAX-DOAS measurements of formaldehyde in the Po-Valley, Atmos. Chem. Phys., 5, 909–918, https://doi.org/10.5194/acp-5-909-2005, 2005.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q J R Meteorol Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., and Deutschländer, T.: Monitoring of climate change in Germany – data, products and services of Germany's National Climate Data Centre, in: Advances in Science and Research, 12th EMS Annual Meeting and 9th European Conference on Applied Climatology (ECAC) 2012, 99–106, https://doi.org/10.5194/asr-10-99-2013, 2013.

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., van Gent, J., Volkamer, R., Vrekoussis, M., Wang, S., Wang, Z., Wenig, M., Wittrock, F., Xie, P. H., Xu, J., Yela, M., Zhang, C., and Zhao, X.: Intercomparison of NO2, O4, O3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2, Atmos. Meas. Tech., 13, 2169–2208, https://doi.org/10.5194/amt-13-2169-2020, 2020.

Lawrence, M. G.: The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications, BAMS 86, 225–234, https://doi.org/10.1175/BAMS-86-2-225, 2005.

Atmospheric
Measurement
Techniques
Discussions

May, R. M., Arms, S. C., Marsh, P., Bruning, E., Leeman, J. R., Goebbert, K., Thielen, J. E., and Bruick, Z., 2021: MetPy: A Python Package for Meteorological Data. Unidata, https://github.com/Unidata/MetPy, doi:10.5065/D6WW7G29.

500 Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy, Springer-Verlag Berlin Heidelberg, 2008.

Romps, D. M.: Exact Expression for the Lifting Condensation Level, J. Atmos. Sci. 74, 3891–3900, https://doi.org/10.1175/JAS-D-17-0102.1, 2017.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, UCAR/NCAR, https://doi.org/10.5065/1DFH-6P97, 2019.

505 Sommer, M., Dirksen, R., and Immler, F.: RS92 GRUAN Data Product Version 2 (RS92-GDP.2), GRUAN Lead Centre, DOI:10.5676/GRUAN/RS92-GDP.2, 2012.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O_2–O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15(37), 15371–15381, doi:10.1039/C3CP50968K, 2013.

Tohjima, Y., Machida, T., Watai, T., Akama, I., Amari, T., and Moriwaki, Y.: Preparation of gravimetric standards for measurements of
510 atmospheric oxygen and reevaluation of atmospheric oxygen concentration, JGR 110 (D11), https://doi.org/10.1029/2004JD005595, 2005.

Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements to the OMI O_2–O_2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9, 6035–6049, https://doi.org/10.5194/amt-9-6035-2016, 2016.

Wagner, T., Erle, F., Marquard, L., Otten, C., Pfeilsticker, K., Senne, T., Stutz, J., and Platt, U.: Cloudy sky optical paths as derived from differential optical absorption spectroscopy observations, JGR 103, 25307–25321, https://doi.org/10.1029/98JD01021, 1998.

515 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.: Is a scaling factor required to obtain closure between measured and modelled atmospheric O4 absorptions? An assessment of uncertainties of measurements and radiative transfer simulations for 2 selected days during the MAD-CAT campaign, Atmos. Meas. Tech.,
520 12, 2745–2817, https://doi.org/10.5194/amt-12-2745-2019, 2019.

Wagner, T., Dörner, S., Beirle, S., Donner, S., and Kinne, S.: Quantitative comparison of measured and simulated O4 absorptions for one day with extremely low aerosol load over the tropical Atlantic, Atmos. Meas. Tech., 14, 3871–3893, https://doi.org/10.5194/amt-14-3871-2021, 2021.

World Meteorological Organization (WMO): Guide to Instruments and Methods of Observation, Volume I: Measurement of Meteorologi-
525 cal Variables, 2018 Edition, WMO, Geneva, ISBN 978-92-63-10008-5, online available at https://library.wmo.int/index.php?lvl=notice_display&id=12407#.YO1QPEyxXYQ, last access 13 July 2021, 2018.