



Reassessment of the glyoxal-to-formaldehyde ratio R_{GF} as a proxy for VOC source identification

Simon Bittner¹, Andreas Richter¹, Bianca Zilker¹, Sebastian Donner², Thomas Wagner², Leonardo M. A. Alvarado³, and Mihalis Vrekoussis^{1,4,5}

¹Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany

²Max Planck Institute for Chemistry, Mainz, Germany

³German Aerospace Center (DLR), Earth Observation Center (EOC), Wessling, Germany

⁴Center of Marine Environmental Sciences (MARUM), University of Bremen, Bremen, Germany

⁵Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, Nicosia, Cyprus

Correspondence: Simon Bittner (simon.bittner@uni-bremen.de)

Abstract. The glyoxal-to-formaldehyde ratio (R_{GF}) has been proposed as a proxy to distinguish sources of volatile organic compounds (VOCs) in the atmosphere. However, the interpretation of its variability remains uncertain because of the diverse processes that affect VOC emissions and chemistry. In this study, we revisit the applicability and limitations of R_{GF} using multi-year ground-based MAX-DOAS measurements at four distinct sites: two biogenic (Orléans, France, and ATTO Tower, Brazil) and two anthropogenic (Athens, Greece, and Incheon, South Korea).

The results show higher R_{GF} in anthropogenic environments and lower at biogenic sites. Seasonal R_{GF} patterns are broadly similar across sites, with reduced values in summer and enhanced values in winter, driven by HCHO variability. Diurnal cycles, caused by CHOCHO variability, are more pronounced at urban sites, with weekend effects of 10%. Correlations between R_{GF} and NO_2 vary, even among anthropogenic stations, indicating the importance of local emission contributions.

Additionally, our analysis shows that increasing temperatures leads to a decrease of R_{GF} by up to 1.9 percentage points across all sites, due to the more rapid increase of HCHO levels with temperature than CHOCHO.

Moreover, we discuss four effects that reduce the comparability between different R_{GF} values: measurement volume, vertical sensitivity, time dependence, and the impact of averaging-ratioing order.

Our findings suggest that ground-based remote sensing R_{GF} contains valuable diagnostic information about VOC source environments. However, its use as a universal proxy remains challenging, as our incomplete understanding of the various effects currently limits the reliable use of R_{GF} for VOC source attribution.

1 Introduction

The European Environmental Agency reported in 2024 that meeting World Health Organization (WHO) air quality standards across EU Member States could prevent 239000 annual deaths from fine particulate matter ($\text{PM}_{2.5}$), 70000 from tropospheric ozone (O_3), and 48000 from nitrogen dioxide (NO_2) exposure (European Environment Agency, 2024). Globally, the situation is comparable, with particularly high numbers of premature deaths occurring in Asia (Lelieveld et al., 2015).



Among the three pollutants, tropospheric O_3 —which has strongly enhanced concentrations in summer smog—is associated with increased cardiovascular and respiratory mortality (Bell et al., 2004; Turner et al., 2016). Its formation requires two precursors in the presence of sunlight: nitrogen oxides ($NO_x = NO_2 + NO$) and volatile organic compounds (VOCs) (Haagen-Smit, 1952). Understanding the role of these individual components is essential for effective ozone mitigation strategies. NO_x emissions originate primarily from fossil fuel combustion, followed by natural sources such as biomass burning, soil emissions, and lightning (IPCC, 2001; Seinfeld and Pandis, 2006).

In contrast, investigating the origin of VOCs, focussing on non-methane VOCs (NMVOCs), is more challenging, as they encompass a large and diverse group of compounds. In addition to their role in tropospheric ozone formation, they contribute to the formation of secondary organic aerosols (SOA) (Hallquist et al., 2009; Derwent et al., 2010) and cloud condensation nuclei (Zheng et al., 2020; Liu and Matsui, 2022). Their sources are generally categorised as biogenic, pyrogenic, or anthropogenic (Vrekoussis et al., 2010).

Among these categories, biogenic VOC emissions represent the largest share of total VOC emissions (Guenther et al., 1995; Stavrakou et al., 2009b). Vegetation emits up to 10000 different VOCs (Goldstein and Galbally, 2007), which are involved in a wide range of processes, including growth, development, communication, and defence against herbivores. Isoprene (C_5H_8) is the most commonly emitted VOC species, followed by monoterpenes ($C_{10}H_{16}$). Emission rates are influenced by many factors and vary across plant species, plant parts, and even leaf age (Laothawornkitkul et al., 2009; Zhang et al., 2023).

Another significant share of VOC emissions originates from pyrogenic sources, such as biomass burning. The combustion of biogenic material releases a complex mixture of species into the atmosphere, including a wide variety of VOCs. The composition of these emissions strongly depends on the material being burned (Gilman et al., 2015) and on moisture content (Paris et al., 2022).

Anthropogenic VOCs are emitted by a variety of sources. The Community Emissions Data System (CEDS) inventory indicates that energy production, road transportation, residential activities, and solvent usage are the dominant processes/sectors on a global scale (McDuffie et al., 2020).

Among VOC species, glyoxal ($CHOCHO$) and formaldehyde ($HCHO$) are key intermediate products of VOC oxidation in the atmosphere (Fu et al., 2008; Chan Miller et al., 2016). $HCHO$ is the most abundant atmospheric aldehyde, with primary emissions from vehicle exhausts (Nelson et al., 2008) and biomass burning (Lee et al., 1998; Andreae and Merlet, 2001). Its main source, however, is formation through secondary production from VOC oxidation (Fortems-Cheiney et al., 2012) and methane (CH_4) oxidation, which determines its background levels (Franco et al., 2016). $HCHO$ is removed from the atmosphere by photolysis, reaction with hydroxyl radicals (OH), and deposition (Stavrakou et al., 2009b). Its typical tropospheric lifetime around midday is about 3 hours (Dienhart et al., 2021).

Glyoxal, the smallest dicarbonyl compound, shares similar sources with $HCHO$: primary emissions from biomass burning (Zarzana et al., 2017, 2018) and biofuel use (Fu et al., 2008), as well as secondary formation via VOC oxidation. Primary glyoxal emissions are generally small compared to its secondary production (Stavrakou et al., 2009a; Silva et al., 2018). Its tropospheric lifetime is short, on the order of a few hours (Volkamer et al., 2007; Myriokefalitakis et al., 2008; Fu et al., 2008).



Glyoxal is removed through photolysis, reactions with OH, and both dry and wet deposition (Myriokefalitakis et al., 2008), with an additional important sink via secondary organic aerosol (SOA) formation (Stavrakou et al., 2009a).

The ratio of glyoxal to formaldehyde, R_{GF} , was proposed by Wittrock et al. (2006) and Vrekoussis et al. (2010) as a potential proxy for differentiating VOC source types. Because CHOCHO and HCHO have similar sources and loss processes, subtle differences in VOC mixtures or source-specific yields are expected to be reflected in R_{GF} .

In their 2010 study, Vrekoussis et al. analysed two years of GOME-2 satellite data (ground pixel size: 80 km×40 km) and found a strong spatial correlation between R_{GF} and VOC source categories (anthropogenic, biogenic, and pyrogenic). They concluded that R_{GF} values below 4% indicate predominantly anthropogenic sources, whereas values above 4% suggest biogenic or pyrogenic origins. Consistently, they observed a decrease of R_{GF} with increasing NO₂ levels and an increase with higher vegetation density, quantified by the Enhanced Vegetation Index (EVI).

Since the initial study, further investigations led to mixed and sometimes contradictory findings, making it difficult to form a consistent interpretation of R_{GF} . Several studies supported or challenged the earlier conclusions (Irie et al., 2011; DiGangi et al., 2012; MacDonald et al., 2012; Li et al., 2014; Chan Miller et al., 2014). For example, DiGangi et al. (2012), based on in-situ data from two campaigns, proposed an interpretation opposite to that of Vrekoussis et al. (2010): lower R_{GF} for biogenic sources and higher values for anthropogenic or pyrogenic origins.

In 2015, Kaiser et al. reassessed R_{GF} and shifted the focus toward VOC precursors. Using airborne in-situ data, they found that R_{GF} primarily depends on the speciation of hydrocarbon precursors, with monoterpenes yielding high R_{GF} values and isoprene low values.

At one rural (Thailand) and one semi-urban site (India), Hoque et al. (2018a, b) and Rawat et al. (2024) used MAX-DOAS observations to investigate R_{GF} . They reported higher values during the wet season, a diurnal cycle peaking at noon, and seasonal variability between sites, periods, and different platforms. Similarly, Xing et al. (2020) used two weeks of MAX-DOAS vertical profile data from Chongqing, China, and found altitude-dependent changes in the diurnal cycle of R_{GF} .

A recent satellite-based study by Chen et al. (2023) using TROPOMI data reported a positive correlation of R_{GF} with both EVI and NO₂. Their results support the latitude dependence noted earlier by Myriokefalitakis et al. (2008) and they reported reduced R_{GF} values during the COVID-19 lockdown. Chen et al. (2023) suggested that regional identification of anthropogenic VOC emissions is feasible under the following conditions: $R_{GF} > 4\%$ and $EVI < 0.2$, and HCHO columns $> 0.5 \times 10^{16}$ molec cm⁻².

Another study, by Hong et al. (2024), using MAX-DOAS data from four megacities in China, proposed excluding primary HCHO emissions from the analysis. They suggested that the different primary emissions between CHOCHO and HCHO reduce R_{GF} accuracy to diagnose VOC sources. By excluding the primary HCHO with a multiple linear regression model, they propose the ratio of CHOCHO to secondary HCHO as a more reliable metric at surface level.

In summary, although considerable progress has been made over the past 15 years, the interpretation of R_{GF} remains challenging and its behaviour differs between platforms and datasets.

This study systematically investigates the drivers and limitations of R_{GF} using MAX-DOAS data from four stations located in different environments. We analyse its magnitude, temporal patterns, and correlations with meteorological variables. More-



over, we discuss four effects, that potentially hinder comparisons of different R_{GF} to, in the end, enhance our understanding and reassess the suitability of R_{GF} as a proxy for VOC origin.

2 Methods and Datasets

2.1 MAX-DOAS

95 Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) is a remote sensing technique that uses scattered sunlight in the ultraviolet (UV) and visible (VIS) spectral ranges to determine trace gas concentrations, integrated along the average atmospheric light path. By computing optical depth from the measured spectrum and a reference spectrum, and comparing it to the known absorption cross-sections of specific trace gases, their atmospheric abundance can be quantified. The spectral fitting process focuses on the differential absorption structures within absorber-specific wavelength intervals, known
 100 as fit windows (Hönninger et al., 2004; Platt and Stutz, 2008).

The term Multi-Axis refers to the instrument’s ability to scan in multiple viewing directions. By measuring at various elevations (vertical) and azimuths (horizontal), different atmospheric layers can be probed. Observations at high elevation angles (around 90° , known as zenith-sky direction) are used for stratospheric absorbers, while low-elevation, off-axis measurements in various azimuth directions are more sensitive to boundary layer trace gas concentrations (Hönninger et al., 2004; Wittrock
 105 et al., 2004; Platt and Stutz, 2008).

The DOAS retrieval yields the measured slant column density (SCD_{meas}), relative to a reference spectrum with its own SCD (SCD_{ref}). Because DOAS captures only the differential absorption between the measured and reference spectra, it provides the differential slant column density (dSCD), see Eq. (1). Mathematically, the SCD is defined as the integral of the absorber number density (n_{meas} or n_{ref}) along the effective light path (ds_{meas} or ds_{ref}) from the top of the atmosphere (TOA) to the
 110 ground, see Eq. (2).

$$dSCD = SCD_{meas} - SCD_{ref} \quad (1)$$

$$= \int_0^{TOA} n_{meas} ds_{meas} - \int_0^{TOA} n_{ref} ds_{ref}. \quad (2)$$

In this study, we primarily use off-axis measurements at low elevation angles (1° – 3°). The atmospheric abundances are retrieved with sequential fits, where the reference spectrum is the corresponding zenith-sky measurement closest in time or
 115 interpolated to the measurement time. This setup has the advantage that most stratospheric influences and diurnal changes in viewing geometry cancel out, so that changes in the dSCD reflect enhancements of the trace gas in the boundary layer near the ground.

The multi-year dataset used here, comprises four stations in different environments: ATTO Tower (Brazil), Orléans (France), Athens (Greece), and Incheon (South Korea). The stations are introduced again in Sect. 2.5; measurement periods and instru-
 120 ment alignments are listed in table A6. Three of the four instruments (Athens, Orléans, and Incheon) were developed and deployed by the University of Bremen and therefore use identical fit settings for NO_2 , $CHOCHO$, and $HCHO$. Measurements



from ATTO were obtained using a different instrument developed and evaluated by the Max Planck Institute for Chemistry (MPIC) (Donner, 2024), and thus, different fit settings were applied. All fit settings of the Bremen instruments are listed in the supplementary material (Tables A1, A2, A3, and A4). The fit settings for the instrument evaluated by MPIC are given in
 125 Donner (2024) in Tables 9.1-9.5.

We apply several quality filters based on the root mean square ($RMS < 0.001$) of the fit residual, intensity (with separate thresholds for UV and vis per station), solar zenith angle ($SZA < 80^\circ$), and the relative slant column density error ($< 50\%$). The relative error filter for the dSCDs constrains the propagated uncertainty of R_{GF} to below 71%. All applied thresholds are summarized in the appendix (Table A5). To facilitate data merging with the meteorological data, the measurements are binned
 130 to 30-minute resolution.

2.2 Computation of R_{GF}

We calculate R_{GF} for each 30-minute bin using the corresponding quality-filtered CHOCHO and HCHO dSCDs. Since CHOCHO and HCHO are retrieved in different spectral ranges, atmospheric scattering processes, such as Rayleigh scattering, vary, resulting in different effective light path lengths. This discrepancy can introduce systematic differences between the two
 135 dSCDs, as each trace gas effectively samples a slightly different part of the boundary layer.

To estimate and correct for differences in light path lengths, we use the collision-induced absorption of O_2 – O_2 , typically approximated as O_4 , which must be included as a pseudo-absorber in DOAS retrievals (Finkenzeller and Volkamer, 2022). The vertical profile of O_4 is well characterised; its vertical column density (VCD) can be accurately calculated because O_2 concentration decreases approximately exponentially with altitude, producing a known vertical distribution of O_4 . To reduce
 140 systematic differences between fits in the UV and the visible part of the spectrum, we apply a first-order correction by multiplying R_{GF} with the inverse of the O_4 dSCDs from the corresponding wavelength regions, see Eq. (3). This correction approach is effective because the O_4 VCD cancels out in the process, leaving only the ratio of the respective air mass factors (AMFs), which accounts for differences in physical processes, see Eq. (4). In the equations, the AMF of the measured spectrum is shown as is (AMF^{O_4}) , while the AMF from the reference spectrum is denoted as $AMF_{ref}^{O_4}$.

The O_4 correction is valid under the assumption that the vertical profiles of CHOCHO and HCHO closely resemble the profile of O_4 , since the O_4 air mass factor (AMF) is used to correct differences in effective light path length. This assumption is reasonable for our dataset, since we focus on the lowest elevation angles, where the slant columns are dominated by absorption in the layer close to the ground. Enhancements at elevated layers, e.g. from fire plumes, would weaken this approach.

$$R_{GF}^* = \frac{dSCD_{vis}^{CHOCHO}}{dSCD_{UV}^{HCHO}} \cdot \frac{dSCD_{UV}^{O_4}}{dSCD_{vis}^{O_4}} \quad (3)$$

$$= R_{GF} \cdot \frac{VCD^{O_4} (AMF_{UV}^{O_4} - AMF_{UV,ref}^{O_4})}{VCD^{O_4} (AMF_{vis}^{O_4} - AMF_{vis,ref}^{O_4})} \quad (4)$$

The O_4 dSCDs used for correction are obtained from the respective NO_2 fits in the visible and UV ranges for the Bremen and MPIC instruments. As the MPIC instrument covers a shorter wavelength range, the NO_2 vis fit window extends only until 460 nm and therefore does not cover the O_4 absorption band at 477 nm, which reduces the quality of O_4 dSCDs in the visible.

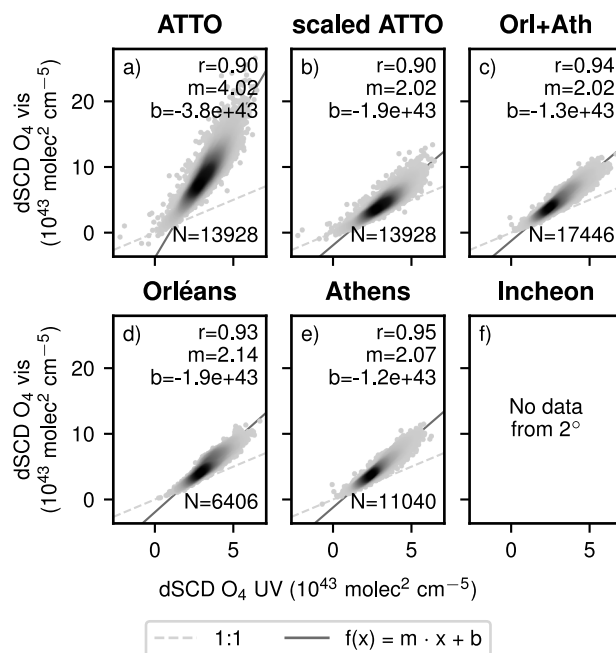


Figure 1. Scatter plots showing the correlation of O_4 dSCDs from the UV wavelength range with the vis wavelength range for different datasets at 2° viewing elevation. The light grey dashed line indicates the 1-to-1 line and the gray solid line indicates a orthogonal linear fit with the specified parameters. The density of the data points is indicated by the hue, denser regions are shown in dark grey.

The ratio of O_4 dSCDs from UV and vis is shown in Fig. 1 for all sites. The original data from ATTO (Fig. 1a) deviates from the other stations, which we primarily attribute to the lower quality of the O_4 retrieval in the visible.

To remove this bias in our study, we scale the O_4 vis dSCDs at ATTO by using the data from all other sites as the reference. A slope is computed for the original dataset (m_{ATTO}), see Fig. 1a, and the reference dataset (m_{ref}), see Fig. 1c. Both slopes are then used to scale the O_4 vis dSCDs, see Eq. (5), which results in Fig. 1b. This scaling is calculated and applied for all elevation angles individually, see Table A7. The effect of the scaling is, that the magnitude of the O_4 ratio is comparable with the other stations, see Fig. 4+5. A comparison of the original and scaled O_4 ratio is found in the appendix in Fig. A6.

$$\text{dSCD}_{\text{vis,scaled}}^{O_4} = \text{dSCD}_{\text{vis}}^{O_4} \cdot \frac{m_{\text{ref}}}{m_{\text{ATTO}}} \quad (5)$$

Overall, the O_4 correction of R_{GF} is important for interpreting the results because the physical processes, influencing the effective light path, are systematically different in the spectral ranges of CHOCHO and HCHO. Since dSCDs are used to compute R_{GF} , if left uncorrected, these light path effects alter the values of R_{GF} , hiding the influence of the actual drivers. Figure 5 (bottom row) illustrates the impact of the correction. The O_4 ratio is relatively constant over the day and primarily reduces the overall magnitude of the corrected R_{GF} . In the following sections, we denote the uncorrected glyoxal-to-formaldehyde ratio as R_{GF} and the ratio with the O_4 correction as R_{GF}^* .



2.3 Meteorological data

We use meteorological data to associate changes in R_{GF}^* with specific meteorological conditions, thereby extending our understanding of its driving factors. To ensure consistency across all stations throughout the measurement periods, we selected data from the ECMWF Reanalysis v5 (ERA5) dataset. ERA5 provides hourly gridded data ($0.25^\circ \times 0.25^\circ$ grid spacing). The meteorological variables included in this analysis are temperature at 2 m, short-wave radiation, and wind speed and direction at 100 m (Hersbach et al., 2023). To merge the ERA5 data with the MAX-DOAS datasets, the ERA5 datasets are interpolated in time from hourly values to every 30 minutes.

2.4 VOC emission data

To investigate differences in anthropogenic VOC sources between the sites Athens and Incheon, we use the CAMS-GLOB-ANT version 6.2 emission dataset created by Copernicus Atmosphere Monitoring Service (CAMS) (Soulie et al., 2024) and provided by ECCAD (Granier et al., 2019). The annually gridded product ($0.1^\circ \times 0.1^\circ$ grid spacing) for NMVOCs (in Tg) are summed up over a region enclosing the two cities, see Fig. A1. Afterwards the annual emissions are averaged over the years of our ground-based coverage for both cities: Athens (2021-2023), Incheon (2021, 2022).

2.5 Measurement sites

Four measurement sites were selected according to their predominant environmental characteristics. Each site was classified based on its surroundings and the chosen viewing direction (Fig. 2). Athens and Incheon represent anthropogenic environments due to enhanced NO_2 levels (Mavroidis and Ilia, 2012; Nguyen et al., 2015; Gratsea et al., 2016; Lange et al., 2024) and high population density within their metropolitan areas (Kim et al., 2021; Hellenic Statistical Authority, 2024).

The third station, Orléans, is classified primarily as a biogenic environment. This classification is supported by relatively low observed NO_2 levels, and a viewing direction aimed directly over forest canopies. The fourth station, ATTO, is similarly considered biogenic, given its remote location within the Amazon rainforest.

Potential pyrogenic influences at ATTO (wildfires during the dry season) and Athens (occasional wildfires) are neglected, as we expect such events to occur only infrequently within our measurement periods.

2.5.1 Athens

The instrument in Athens is located at the National Observatory of Athens in Penteli, Greece. The Athens metropolitan area, with approximately 3 million inhabitants in the Attica region (Hellenic Statistical Authority, 2024), is strongly influenced by anthropogenic activity.

The MAX-DOAS instrument is installed on a hill approximately 500 m above ground level to the east of the city (see Fig. 2). Measurements are routinely conducted in multiple directions. For this analysis, we use data collected between January 2021 and December 2023 from the viewing direction oriented toward the city centre (232.5° relative to north, indicated by the purple line in Fig. 2). Additional details regarding the instrument hardware and setup are given by Gratsea et al. (2016).

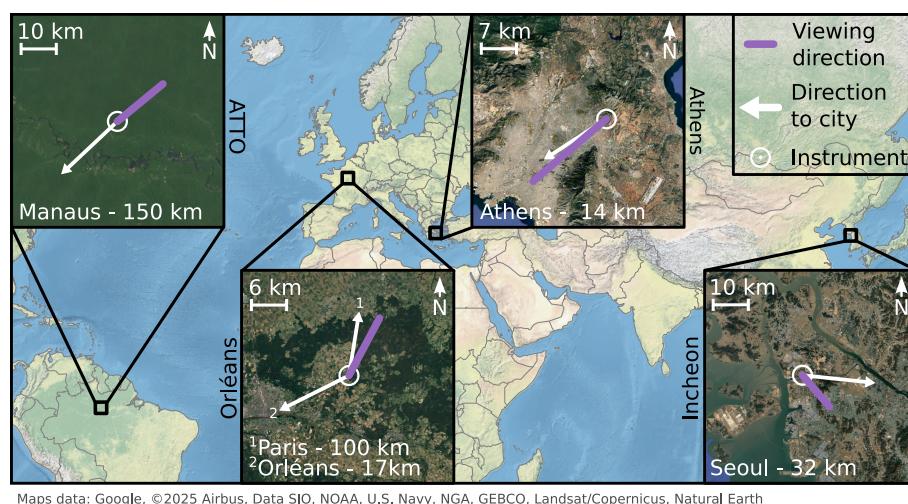


Figure 2. World map showing the location of all stations, their surroundings and distances to neighbouring cities. The white circles indicate the instrument positions, the white arrows show the direction to the cities, whereas the purple lines correspond to the relevant viewing direction of the instruments.

Under certain meteorological conditions, local topography causes pollutants to accumulate within the urban area (Kas-
 200 somenos et al., 1995). Additionally, due to its hot and dry climate, Athens occasionally experiences wildfires, as observed, for
 example, in 2018 and 2024 (Lagouvardos et al., 2019; Castro-Melgar et al., 2025). Mountains with Mediterranean vegetation
 are located to the north. To the east, the landscape features mountainous vegetation interspersed with smaller residential areas,
 while to the south lie the airport and lower-density residential and industrial zones. The city centre of Athens and the port of
 Piraeus are situated to the west.

205 Meteorologically, the region experiences low precipitation, pronounced diurnal and seasonal cycles in short-wave radi-
 ation, and relatively high temperatures exhibiting clear seasonal and daily variations (Fig. 3). The prevailing winds during the
 measurement period come from northern directions, frequently reaching speeds above 9 ms^{-1} (Fig. A2).

2.5.2 Orléans

The second station is located near Orléans, France, on the premises of a radio station in Traînou, which is regularly used
 210 for scientific measurements, including the ICOS project (Ramonet et al., 2025). Traînou (approx. 3500 inhabitants; Institut
 national de la statistique et des études économiques, 2025) is situated about 100 km south of Paris and 17 km northeast
 of Orléans (116000 inhabitants; Institut national de la statistique et des études économiques, 2024). The exact location and
 surroundings of the instrument are shown in Fig. 2.

Crucially for this study, the site is adjacent to a large forested region. The Orléans State Forest covers roughly 350 km^2 , and
 215 comprises a mixture of broadleaf and evergreen tree species (Bello et al., 2019). Thus, this measurement site is primarily influ-

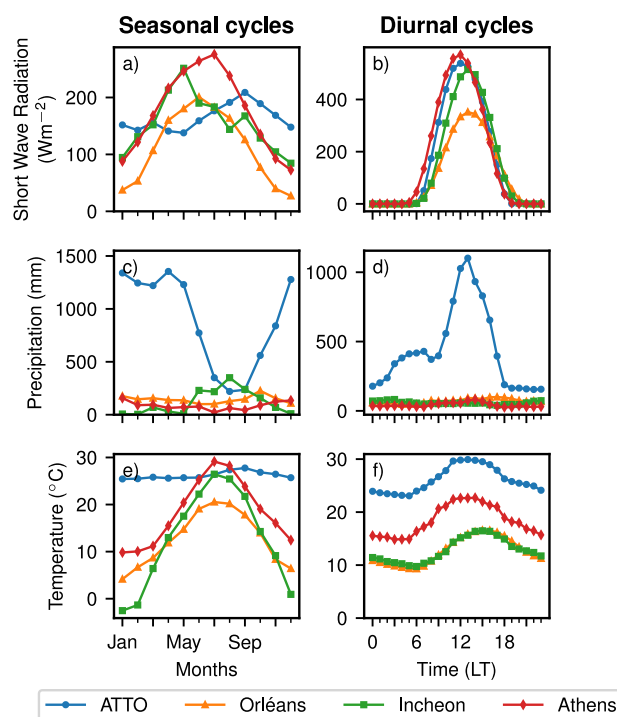


Figure 3. Meteorological overview showing the seasonal cycles (left column) and the diurnal cycles (right column) of mean short wave radiation (first row), sum of precipitation (second row) and mean temperature (third row) for the analysed stations based on ERA5 data for the individual measurement periods.

enced by biogenic activity, with minimal local anthropogenic emissions, although pollutant plumes from Paris can occasionally be detected under northerly winds.

The MAX-DOAS instrument is mounted on an elevated position (approx. 10 m above ground level), enabling low-elevation scans directly above the forest canopy. Data analysed in this study cover the period from July 2023 to July 2024, focusing on measurements taken towards the forest (28° relative to north, Fig. 2).

Meteorologically, Orléans experiences strong seasonal and diurnal variations in short-wave radiation, though its maximum values are comparatively lower due to its higher latitude (Fig. 3). Precipitation is moderate without a clear seasonality. Temperatures are among the lowest of the investigated sites, with a less pronounced seasonal cycle than at Incheon and Athens. The prevailing wind direction is from the southwest, frequently exhibiting high wind speeds exceeding 9 ms^{-1} (Fig. A2).

2.5.3 Incheon

The third instrument was installed on the roof of the Environmental Satellite Center in Incheon, part of the Seoul Metropolitan Area (SMA) in South Korea. With approximately 3 million inhabitants, Incheon is South Korea's third-largest city. The SMA is the most densely populated region in the country (Kim et al., 2021).



The instrument's exact location and surroundings are shown in Fig. 2. It is situated in an anthropogenically dominated environment, with Seoul city centre approximately 32 km to the east, Incheon city centre to the south, and the harbour area to the west. The northern edge of the metropolitan area borders North Korea (about 20 km north), where some forested mountains are located.

As part of the GEMS Map of Air Pollution (GMAP) 2021 campaign and the Satellite Integrated Joint Monitoring of Air Quality (SIJAQ) 2022 campaign, MAX-DOAS measurements were conducted for about one year. For this study, we analyse data from October 2021 to November 2022, focusing on the urban azimuth viewing direction at 137.5° relative to north.

Meteorologically, Incheon is strongly influenced by the East Asian monsoon. Heavy rainfall occurs between June and September, while the rest of the year is comparatively dry. No pronounced diurnal precipitation cycle is observed. The seasonal precipitation pattern affects short-wave radiation, which declines during the wet months but otherwise shows strong seasonal and diurnal cycles with high peak values. Temperatures also exhibit strong seasonal and diurnal variability, with the lowest temperatures across all sites recorded in December and January. The prevailing wind direction is from the northwest (especially during Winter) or west (Fig. A2).

2.5.4 ATTO

The fourth station is located on the tall ATTO Tower in Brazil, deep within the Amazon rainforest. Situated approximately 150 km northeast of Manaus (population 2 million, (Instituto Brasileiro de Geografia e Estatística, 2022)), the ATTO site serves as a remote research site in the heart of the rainforest (Fig. 2). The surrounding area is sparsely populated, resulting in the site being predominantly influenced by biogenic activity. During the dry season wildfires are more frequent and affect local atmospheric conditions (Andreae et al., 2015; Donner, 2024).

The instrument was installed at a height of 80 m in October 2017 and measurements are still ongoing. However, not all data are yet analysed in scientific quality, so the used dataset ends in August 2022. Some data gaps occurred due to the challenging hot and humid climate affecting hardware and electronics. The instrument is oriented northeast (50° relative to north). The dataset analysed in this study was originally obtained by Donner (2024), who also provided a detailed description of the site and instrumentation. Some figures from that publication are reproduced here using our own processing methodology based on their dataset. In such cases, the figure captions indicate which panels are affected.

Meteorologically, the ATTO Tower is characterised by a tropical climate, see Fig. 3. Precipitation predominantly occurs during the wet season (December–May), with much drier conditions during the rest of the year. A pronounced diurnal precipitation cycle is observed. Temperatures are consistently high, showing daily but minimal annual variation. Short-wave radiation exhibits a strong diurnal pattern but remains relatively stable on seasonal timescales, with only a slight reduction during the wet season. Prevailing winds are from the northeast, but compared to the other sites, wind speeds are predominantly low, typically below 3 ms^{-1} (Fig. A2).



260 3 Results

In this chapter, we investigate multiple factors influencing R_{GF}^* to assess its potential as an indicator for VOC source identification. As outlined in the introduction, previous studies have identified a wide range of influencing factors. This work specifically examines diurnal, weekly, and seasonal cycles, as well as the dependence of R_{GF}^* on elevation angle, temperature, and NO_2 levels across all four stations.

265 3.1 Diurnal cycle

A diurnal cycle describes the variation over a day. It allows to compare with other variables that change regularly over the day, e.g. incoming solar radiation or car traffic. For the case of R_{GF} , multiple diurnal cycles were reported. At two sites, one in India (semi-urban) and one in Thailand (rural), Hoque et al. (2018a, b); Rawat et al. (2024) observed a diurnal cycle with a noon maximum for R_{GF} based on VCDs retrieved from MAX-DOAS measurements. The values ranged from 2–4%. Hoque et al. (2018a) also found the diurnal cycle of R_{GF} to be less pronounced in the dry season compared to the wet season, which we will revisit for ATTO in Sect. 3.2.

DiGangi et al. (2012) investigated the diurnal cycle at two predominantly biogenic sites at higher altitudes (Sierra Nevada Mountains, 1315 m; Rocky Mountains 2286 m) with field campaigns in July 2009 and August 2010 utilizing in-situ instruments. The average R_{GF} values for both campaigns were about 2% and 1.7%. At the Sierra Nevada site, R_{GF} increased to about 3% around midday to to afternoon. Whereas the Rocky Mountains campaign showed only minor diurnal variability. DiGangi et al. (2012) attributed the observed enhancements primarily to anthropogenic VOCs and biomass burning plumes encountered during the campaigns.

The diurnal cycles of R_{GF}^* for our dataset (all four stations) are shown in Fig. 4, plotted against local solar time (LST). The diurnal cycles differ strongly across the stations. Anthropogenic sites show pronounced diurnal variability, while the biogenic sites show relatively little variation over the day. At ATTO and Orléans, the diurnal cycles are relatively flat, with average values of 3.1 ± 0.4 %pt. and 2.2 ± 0.2 %pt., respectively. In contrast, Athens and Incheon exhibit higher average values (3.5 ± 0.4 %pt. and 3.7 ± 0.7 %pt.) and strong diurnal patterns, with peaks around 10 am in Athens and noon in Incheon. In Athens, the cycle follows morning rush hour, whereas Incheon has a noon maximum, indicating different drivers of R_{GF}^* over the day for both cities.

Except at ATTO, data availability decreases in the early morning and late afternoon due to the applied SZA filter, which excludes measurements when the sun is low. Data from these times of day are mainly collected during summer months, introducing a seasonal bias in the early and late portion of the diurnal cycle. Despite this, the diurnal patterns are representative across the year, although seasonal offsets are present. It is important to note that the number of valid data points decreases substantially during the winter months at Orléans, Incheon, and Athens, primarily because we filter by relative error of CHOCHO dSCDs, which increases for low atmospheric concentrations. Filtering by relative error is needed to limit the scatter of R_{GF}^* , but it means that R_{GF}^* is more representative for high CHOCHO and HCHO columns.

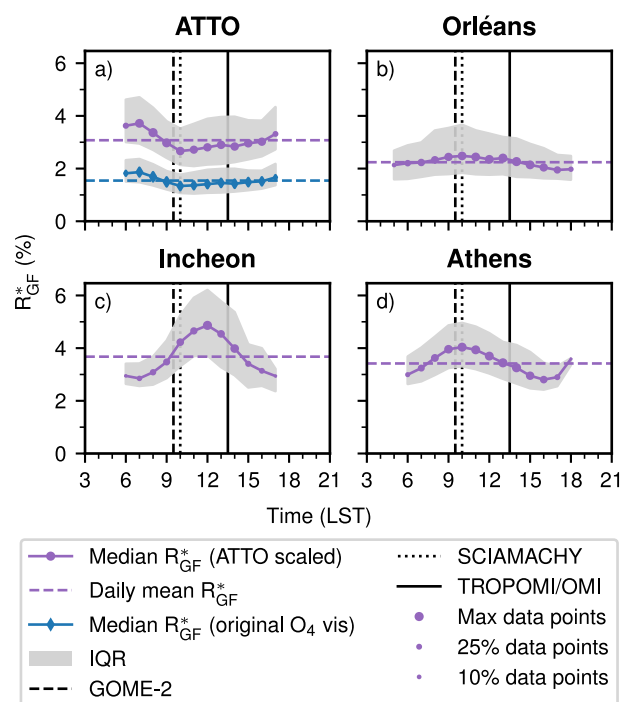


Figure 4. Diurnal cycle of R_{GF}^* for ATTO a), Orléans b), Incheon c) and Athens d) relative to local solar time (LST). The stations are sorted in a clockwise order from the most biogenic environment (ATTO, top left) to the most anthropogenic environment (Incheon, bottom left). To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The solid line corresponds to the median value, the dashed line to the daily mean, and the gray shaded area to the interquartile range (IQR). The original R_{GF}^* without scaling O_4 vis dSCDs is shown for ATTO in blue with diamond markers. In addition, the overpass times of GOME-2, SCIAMACHY and TROPOMI/OMI are highlighted with black vertical bars. Only data from the lowest viewing elevation is included.

The effect of scaling the vis O_4 dSCDs for ATTO is highlighted again by showing the original data with the blue lines in Fig. 4+5. The overall high O_4 dSCDs in the visible lead to a really low O_4 ratio (Fig. 5i), which is mirrored in the overall low level of R_{GF}^* (Fig. 4a).

295 Examining the three components separately (CHOCHO dSCDs in the top row, HCHO dSCDs in the center row, and the O_4 ratio in the bottom row of Fig. 5) reveals that CHOCHO behaves differently across the four stations. The O_4 ratio, after scaling ATTO, is of similar magnitude across sites and does not contribute to a pronounced diurnal cycle. HCHO dSCDs follow a U-shaped diurnal cycle at all stations, with a maximum in the morning and evening and a minimum around noon. This pattern has previously been attributed to enhanced sinks (photolysis and OH oxidation) dominating around midday (Nussbaumer et al.,
 300 2021; Donner, 2024). However, the underlying processes are more complex as they can also promote secondary formation of CHOCHO and HCHO by breaking down VOC precursors. In addition, potential dilution effects associated with boundary layer growth may contribute to the observed shape.



The diurnal cycle of the CHOCHO dSCDs varies in magnitude and shape across stations. ATTO also shows a pronounced U-shape. Orléans exhibits a relatively flat diurnal cycle. Contrasting to that, the anthropogenic stations show a different behaviour. Here, we find higher daily averages plus a maximum in Athens around 10 am and in Incheon over noon. The diurnal cycle shapes of CHOCHO at the anthropogenic stations suggest a link to anthropogenic activity, and their difference from HCHO indicates a weaker anthropogenic influence on HCHO. Since direct CHOCHO emissions are suspected to be low (Stavrakou et al., 2009a; Silva et al., 2018), anthropogenically emitted precursors with a high CHOCHO yield might be a possible explanation, like aromatics (Chan Miller et al., 2016) or acetylene/ethylene (Fu et al., 2008).

Considering these curves, the diurnal cycle of R_{GF}^* appears to be driven by CHOCHO. The enhanced daily mean R_{GF}^* in Athens and Incheon can be explained by the overall higher CHOCHO levels. The shape of R_{GF}^* diurnal cycle can be attributed to the behaviour of CHOCHO dSCDs. Similar shapes between CHOCHO dSCDs and HCHO dSCDs lead to flat cycles at ATTO and Orléans, whereas the different shapes of CHOCHO dSCDs and HCHO dSCDs lead to a pronounced diurnal cycle of R_{GF}^* at the anthropogenic stations.

Compared with previous studies, these observations are consistent with the reported midday peaks in R_{GF} at sites in Thailand (described as rural) and India (described as semi-urban) (Hoque et al., 2018a, b; Rawat et al., 2024), and moderate to flat diurnal cycles with lower absolute RGF (<2%) at predominantly biogenic sites (DiGangi et al., 2012).

In summary, R_{GF}^* shows enhanced average values over the day for anthropogenic stations, due to enhanced CHOCHO levels. This indicates that R_{GF}^* contains information about the different environments, which supports its usage as a proxy for VOC origin. The pronounced diurnal cycles for anthropogenic stations, however, complicate the interpretation as the timing of the measurement becomes important. The implications for comparing R_{GF} values of different studies are discussed in Sect. 3.7.3.

3.2 Seasonal cycle

The variation over the year, the seasonal cycle, enables to investigate how the variable is connected to changes of other variables based on seasons. Multiple studies have reported seasonal cycles for R_{GF} so far: Hoque et al. (2018a, b); Rawat et al. (2024) found a relatively flat seasonal pattern at Pantnagar (India, described as semi-urban) based on MAX-DOAS VCDs. At a second site, Phimai (Thailand, described as rural), the seasonal cycle showed an increase from January to September. Similarly, Xing et al. (2025), analysing one year of MAX-DOAS VCDs from Guangzhou (China), found enhanced R_{GF} values from November to April and lower values during the rest of the year.

The seasonal cycles of R_{GF}^* for our dataset of all four stations are shown in Fig. 6. The annual mean R_{GF}^* of the stations ranges from 2.8 ± 0.7 %pt. in Orléans to 4.2 ± 0.8 %pt. at Incheon.

The overall shape of the seasonal cycle is similar across stations, with one minimum and one maximum per year. At Orléans, Athens, and Incheon, the lowest values occur in July and August (late summer), while the highest values are observed between October and March (winter). At ATTO, the seasonal cycle is shifted by several months, with a minimum in October (dry season) and a maximum extending into June (wet season). Notably, the minimum phase at the biogenic sites tends to be more

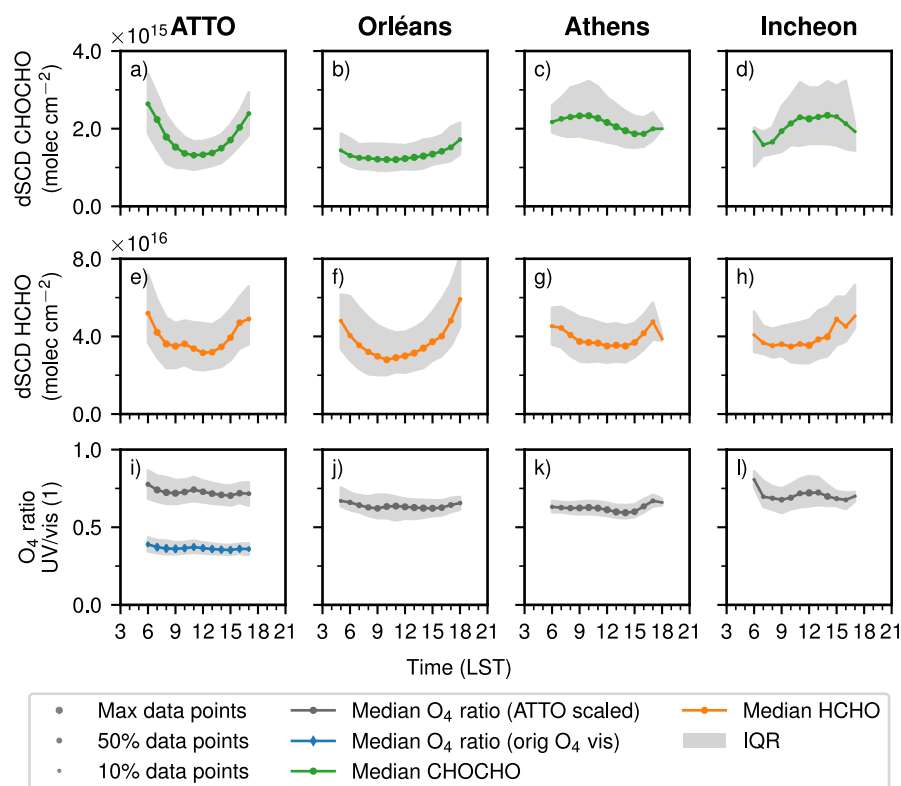


Figure 5. Diurnal cycles of CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) for ATTO, Orléans, Athens, and Incheon relative to local solar time (LST). To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The line corresponds to the median value and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included. The original O_4 ratio without scaling O_4 vis dSCDs is shown for ATTO in blue with diamond markers. Panel a) + e) are self-created based on Donner (2024).

prolonged compared to the anthropogenic sites. Fewer data points are available in winter due to filtering based on relative error (see Sect. 3.1).

Examining the three components separately (CHOCHO dSCDs in the top row, HCHO dSCDs in the center row, and the O_4 ratio in the bottom row of Fig. 7) reveals that both trace gases behave differently for all four stations, whereas the O_4 ratio is similar. Looking at the seasonal cycles of HCHO dSCDs we see one enhanced period during the middle of the year, a narrow peak during June and July for Athens, and an extended peak over four months spanning from June to October for the other stations. The annual means and the amplitude are comparable between the stations.

The seasonal cycle of CHOCHO dSCDs is relatively flat with one peak in different months from June (Athens) to October (Incheon). One can see a shift to higher annual mean values from ATTO to Incheon. The anthropogenic stations show the highest CHOCHO dSCDs and more variability over the year.

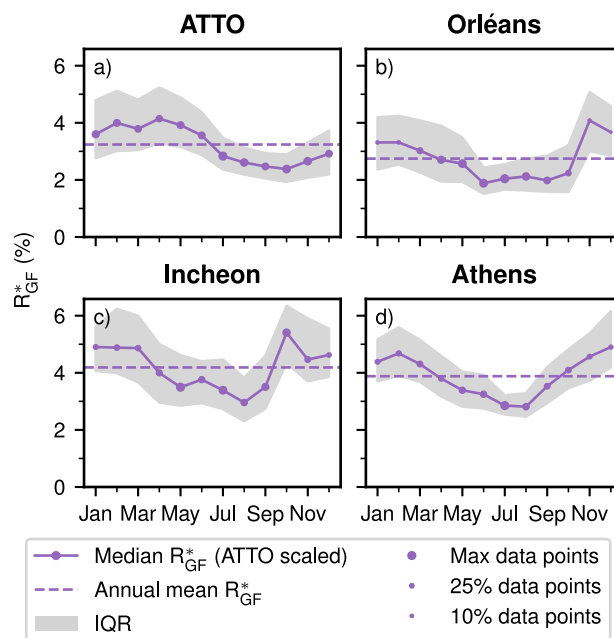


Figure 6. Seasonal cycle of R_{GF}^* for ATTO a), Orléans b), Incheon c) and Athens d). The stations are sorted in a clockwise order from the most biogenic environment (ATTO, top left) to the most anthropogenic environment (Incheon, bottom left). To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The solid line corresponds to the median value, the dashed line to the annual mean, and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included. Panel a) is self-created based on Donner (2024).

The second axis in Fig. 7 shows the temperature at 2 m from ERA5, matched to each measurement and then binned in conjunction with the dSCDs. For all stations except ATTO, the seasonal cycle of HCHO dSCDs closely resembles the seasonal cycle of temperature. The seasonal cycle of CHOCHO does not show a clear influence by temperature. This comparison highlights ATTOs unique tropical conditions: The temperature does not change significantly over the year and therefore both trace gases are driven by different processes.

Donner (2024) suggested, that both trace gases undergo different processing in the dry and wet season and that both seasons probably have a different precursor composition. Figure 8 shows the diurnal cycle of R_{GF}^* during the wet and dry season. As reported by Donner (2024), enhanced R_{GF}^* values in the wet season and reduced values in the dry season are found. The daily mean is reduced by 0.7%pt. in the dry season. The shape of the diurnal cycle is relatively flat. Since forest fires predominantly occur in the dry period, previously excluded pyrogenic activity may contribute to the observed changes. This would be supported by enhanced NO_2 levels and aerosols during the dry season as shown by Donner (2024). However as biomass burning has been reported to lead to higher R_{GF} levels (DiGangi et al., 2012; Zarzana et al., 2017; Chan Miller et al., 2014; Alvarado et al., 2020), the lower R_{GF} levels are unlikely to originate from pyrogenic activity in the vicinity.

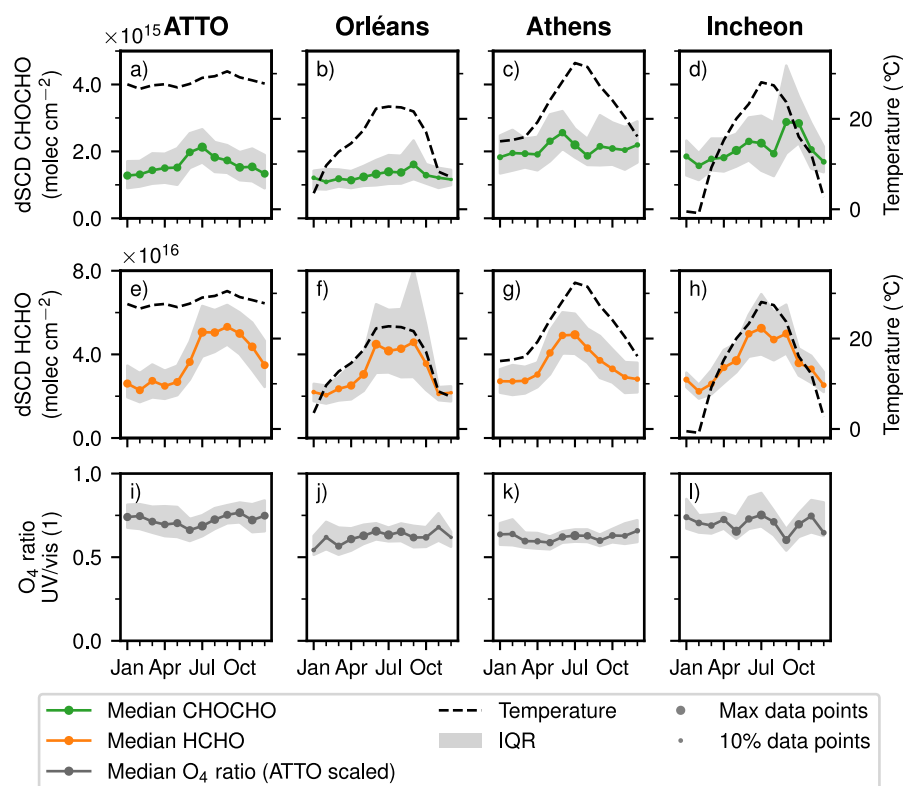


Figure 7. Seasonal cycle of CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) for ATTO, Orléans, Athens, and Incheon. To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The line corresponds to the median value and the gray shaded area to the IQR. The seasonal cycle of temperature is shown on a secondary axes with a solid black line. Only data from the lowest viewing elevation is included. Panels a) + e) are self-created based on Donner (2024).

The discrepancy between wet and dry season is in agreement with the findings of Hoque et al. (2018a), where they found higher R_{GF} during the wet season and lower R_{GF} during the dry season in Phimai (Thailand). Furthermore, both seasons share the same diurnal cycle for Hoque et al. (2018a). However, their diurnal cycle had a pronounced noon maximum, which is not present in this dataset, which might, even though the Phimai site is described as rural, hint at a stronger anthropogenic influence than at ATTO, see Sect. 3.1.

Having these points in mind, the seasonal cycle of R_{GF}^* seems to be driven, contrary to the diurnal cycles, by the variability of HCHO. The variability of HCHO dSCDs is strongly connected to the variability of temperature for non-tropical stations and seems to be connected to the dry/wet season for ATTO. The enhanced annual mean R_{GF}^* at anthropogenic stations can be explained by higher CHOCHO levels.

Comparing with the previous studies, these observations of anthropogenic stations match best with the enhancements during the colder months reported by Xing et al. (2025). Comparing the absolute values, lower absolute R_{GF}^* values are observed, but

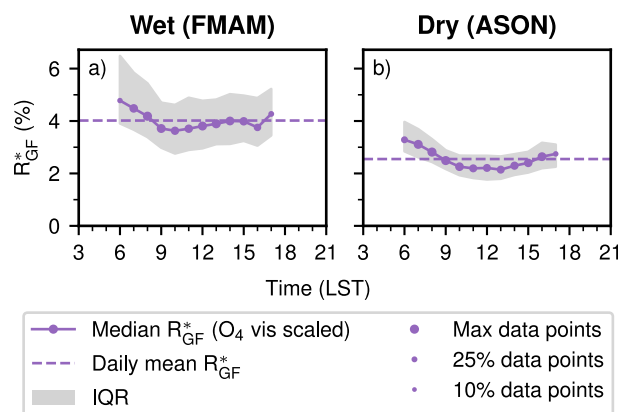


Figure 8. Diurnal cycles in the wet a) and dry b) season of R_{GF}^* at ATTO. To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The solid line corresponds to the median value, the dashed line to the daily mean, and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included.

caution is required, as the R_{GF}^* is based on corrected dSCDs for the lowest viewing elevation instead of VCDs. Contrary to Hoque et al. (2018b), no increase over the year is visible in our dataset, but a similar magnitude for remote stations is found in this study.

To conclude, we see a similar pattern for seasonal cycles as for diurnal cycles: R_{GF}^* exhibits a cycle and its average and its amplitude are more pronounced for anthropogenic stations, but this time originating from variations in HCHO. This complicates the interpretation of R_{GF}^* as a proxy for VOC origin, because many other seasonal effects can contribute to its variation, e.g. temperature, which are difficult to disentangle from changes in VOC origin over the year. Moreover, longer time series are needed for measurement campaigns to avoid sampling biases.

3.3 Weekly cycle

As anthropogenic emissions are typically lower on the weekend, the weekly cycles can be used as an indicator for the contribution of anthropogenic emissions (Beirle et al., 2003). Gratsea et al. (2016) reported a weekly cycle in Athens for glyoxal and to a lesser extent for formaldehyde, but only for measurements dominated by urban air.

To our knowledge, no previous study has investigated weekly cycles specifically for R_{GF} , but from the findings of Gratsea et al. (2016), we expect a weekend effect may occur. The weekly cycle of R_{GF}^* for our four sites is shown in Fig. 9. ATTO and Orléans display flat weekly cycles, while the anthropogenic stations show an offset between weekday and weekend. Moreover, as seen and discussed before for diurnal and seasonal cycles, the average value throughout the week is higher for the anthropogenic stations.

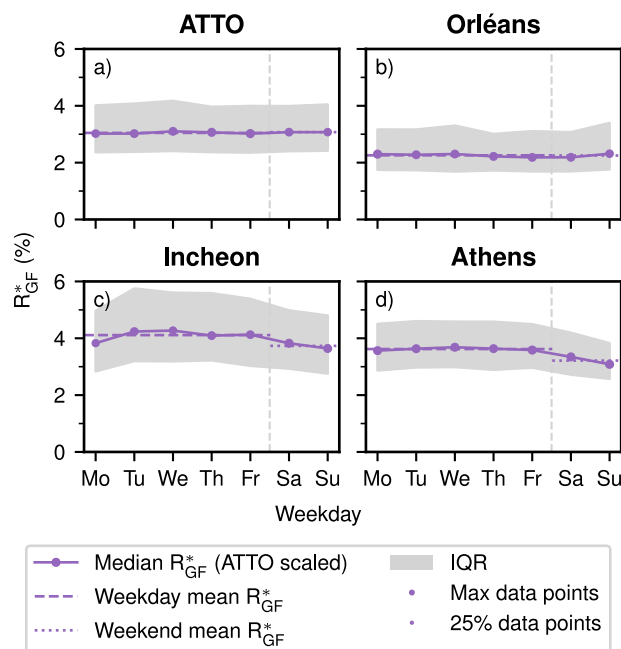


Figure 9. Weekly cycle of R_{GF}^* for ATTO a), Orléans b), Incheon c) and Athens d). The stations are sorted in a clockwise order from the most biogenic environment (ATTO, top left) to the most anthropogenic environment (Incheon, bottom left). To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The solid line corresponds to the median value, the dashed line to the mean weekday value, the dotted line to the mean weekend value, and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included.

Based on the four sites, we observe a weekend effect in R_{GF}^* at the anthropogenic sites, Incheon and Athens. Specifically, R_{GF}^* decreases on weekends in both Incheon and Athens, with the strongest reduction seen in Athens by up to 0.4 %pt., i.e. about 10% lower than during weekdays.

390 In Figure 10 weekly cycles of both oxygenated VOCs (OVOCs) and the O_4 ratio are shown. For ATTO and Orléans the weekly cycles of both OVOCs are relatively flat and show no weekend effect. Comparing both OVOCs over the week for the anthropogenic stations, CHOCHO dSCDs show a strong weekend effect for Athens and Incheon. HCHO dSCDs, however, do not show a strong decrease on the weekend, therefore the weekend effect observed for R_{GF}^* is driven by the weekend effect from CHOCHO dSCDs.

395 To summarize, R_{GF}^* exhibits a weekend effect for anthropogenic stations, driven by the weekend effect of CHOCHO dSCDs. Showing a weekend effect supports R_{GF}^* usage as a proxy for different VOC origin, as changes in anthropogenic emissions are mirrored in R_{GF}^* .

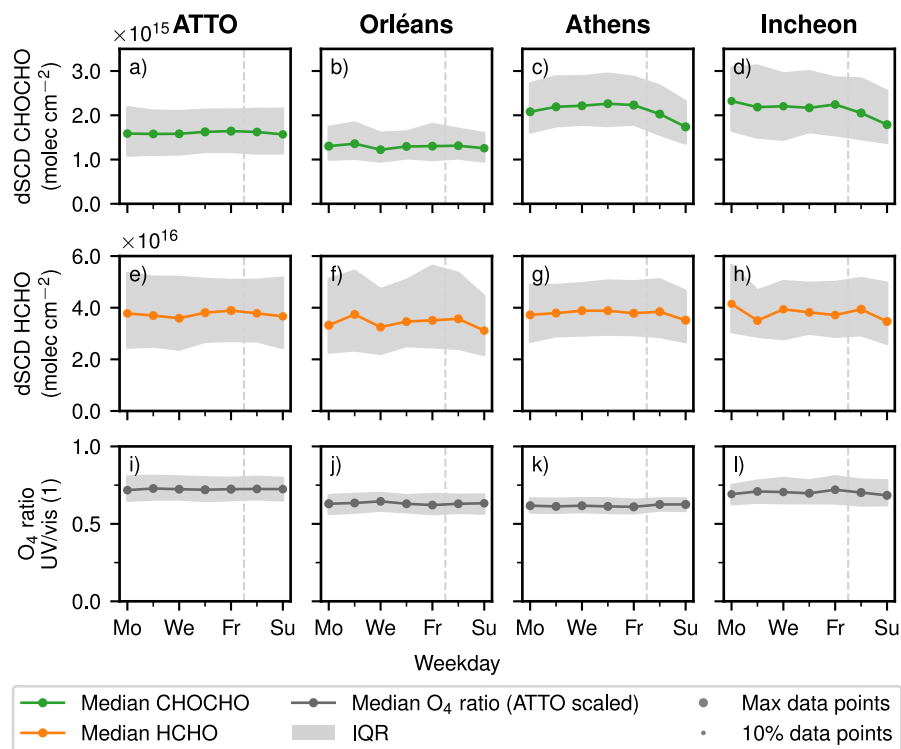


Figure 10. Weekly cycle of CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) for ATTO, Orléans, Athens, and Incheon. To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The line corresponds to the median value and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included.

3.4 Elevation dependence

Investigating the elevation dependence of R_{GF} provides insight into how R_{GF} is associated with different altitudes. Previous studies that examined the altitude dependence of R_{GF} reported varying results. Kaiser et al. (2015) observed a slight increase in R_{GF} in the free troposphere compared to the boundary layer. Similarly, Xing et al. (2020) reported lower values within the lowest 100 m and higher values up to 1 km altitude. In contrast, Rawat et al. (2024) found relatively stable R_{GF} values up to approximately 3.7 km, followed by a sharp decrease above that level. Hong et al. (2024), studying Chinese cities, reported a strong increase in R_{GF} with altitude, up to 2 km.

As we use dSCDs for R_{GF}^* in our analysis, we can only investigate elevation dependence, rather than true altitude dependence. However, the elevation dependence still hints at the altitude distribution in the way that the signal from different elevation angles is dominated by different altitudes.

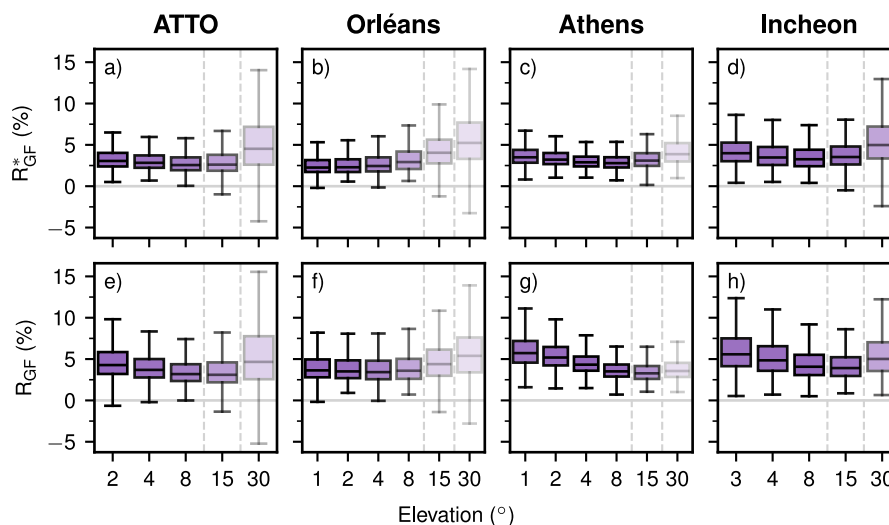


Figure 11. R_{GF}^* (with O_4 correction) and R_{GF} (without O_4 correction) at different viewing elevations for ATTO, Orléans, Athens, and Incheon. The stations are sorted from left to right from the most biogenic environment (ATTO) to the most anthropogenic environment (Incheon). The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. The vertical dashed light gray line indicates the transition to higher viewing elevations. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Note, that the lowest viewing elevation differs across the stations.

Since the O_4 correction is motivated by the lowest elevations close to the surface, Fig. 11 shows R_{GF}^* and the uncorrected R_{GF} . Interpretation of R_{GF}^* at higher elevation angles requires caution, which is why both quantities are shown. R_{GF}^* increases
 410 towards higher elevation angles at all four sites. At lower elevation angles R_{GF}^* is relatively constant.

Looking at the three components, see Fig. 12, we see that CHOCHO and HCHO dSCDs behave consistently across the stations, but differently when directly compared. CHOCHO dSCDs drop consistently towards higher elevation angles, whereas HCHO dSCDs stay constant for the lowest 2-3 elevation angles and drop afterwards. The enhanced levels at higher elevation indicate a different profile and thus a higher lifetime for HCHO compared to CHOCHO. Moreover, the O_4 ratio shows an
 415 increase with elevation and approaches 1 for 30° elevation, implying that no correction is applied at this angle.

The elevation dependency of R_{GF}^* is influenced by two effects. (1) the difference elevation behaviour of CHOCHO versus HCHO described above, (2) the elevation dependence of the O_4 correction. The light path differences in the UV and vis are strongest for the lowest elevations and disappear for 30° direction. This behaviour amplifies the strong decrease of HCHO at higher elevations, leading to higher R_{GF}^* values at higher elevations.

420 To summarize, it is unclear how R_{GF}^* changes with altitude. Our analysis, limited to different elevation angles, suggests that R_{GF}^* increases with elevation.

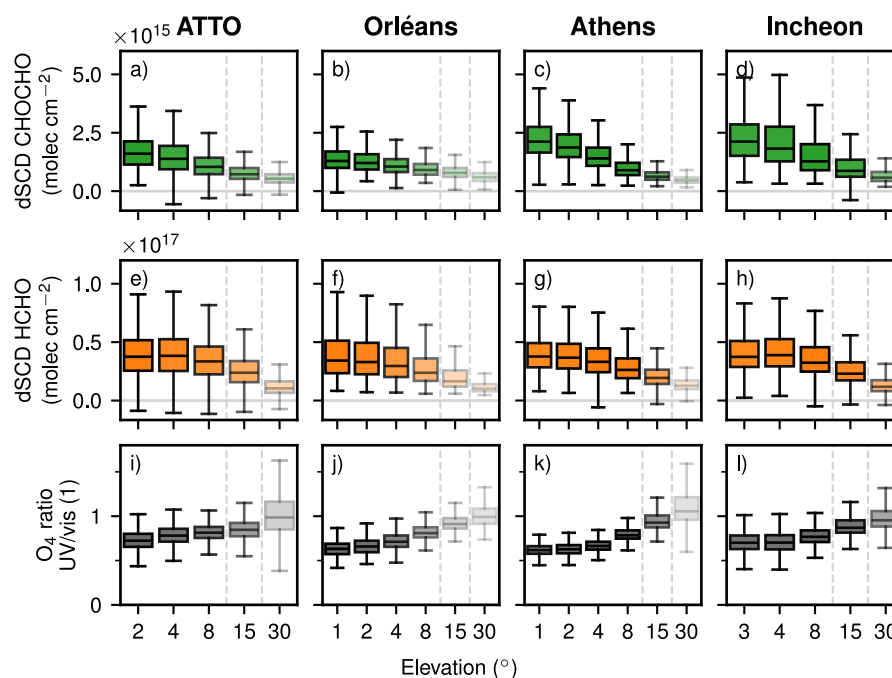


Figure 12. CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) at different viewing elevations for ATTO, Orléans, Athens, and Incheon. The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. The vertical dashed light gray line indicates the transition to higher viewing elevations. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Note, that the lowest viewing elevation differs among the stations.

3.5 Temperature dependence

Atmospheric levels of VOCs are known to be influenced by temperature (Pusede et al., 2014; Bourtsoukidis et al., 2024; Li et al., 2024), which could also impact R_{GF} . Multiple effects can contribute to the temperature dependence of the trace gases, e.g. biogenic emissions and secondary formation rates. A previous study from Guo et al. (2021) used long-path DOAS in Shanghai (China) during summer. They touched the temperature dependence of R_{GF} for their campaign duration and concluded that R_{GF} increases with temperature.

The temperature dependence for our stations is shown in Fig. 13. Across all stations, R_{GF}^* exhibits a similar pattern: at lower temperatures, values remain relatively stable with some fluctuations. However, starting from about 15°C, R_{GF}^* decreases hyperbolically, with a maximum reduction of up to 1.9 %pt. observed at Athens.

The temperature dependence of the three components is shown in Fig. 14. The O_4 ratio does not vary with temperature. Looking at HCHO dSCDs (center row) it is visible, that the HCHO levels grow exponentially with increasing temperatures across all stations. Comparing with CHOCHO dSCDs, we see that the CHOCHO dSCDs also rise with temperature for

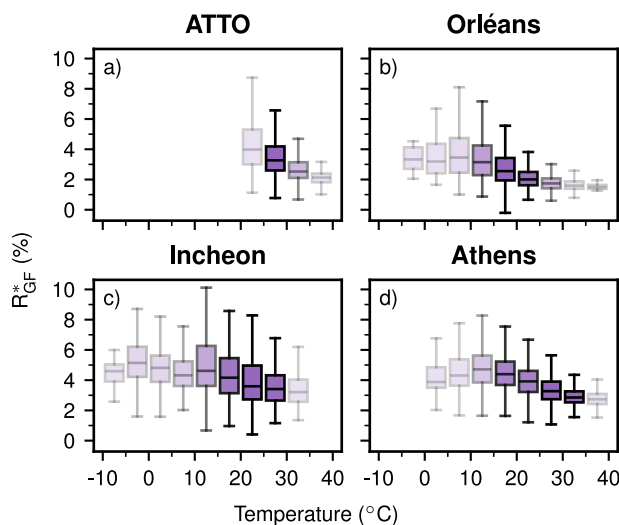


Figure 13. R_{GF}^* at different temperatures intervals for ATTO a), Orléans b), Incheon c) and Athens d). The stations are sorted in a clockwise order from the most biogenic environment (ATTO, top left) to the most anthropogenic environment (Incheon, bottom left). The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Missing box plots indicate that no data points fall within that interval. Only data from the lowest viewing elevation is included.

ATTO and Orléans, but way less pronounced. For Athens and Incheon the bulk of the CHOCHO dSCDs does not increase
 435 with temperature. However, the maximum values occur at the highest temperatures for both sites.

The exponential behaviour is expected, especially for the biogenic stations, as biogenic emissions of precursors are known to increase exponentially with temperature (Guenther et al., 1993, 2012; Bourtsoukidis et al., 2024); higher temperatures enhance biogenic activity, which in turn leads to greater VOC emissions. In addition, the secondary formation via OH oxidation should increase with temperature as reaction rates rise (Berg et al., 2024).

440 For the anthropogenic stations, it is a different situation. Here, we expect the anthropogenic emissions to be temperature independent and attribute the exponential increase with temperature primarily to the increased secondary formation at high temperatures. Adding to that, recent studies suggest that local biogenic VOC emissions in urban environments may play a more important role in local atmospheric chemistry than previously assumed (Liaskoni et al., 2024; Wang et al., 2025).

It is noteworthy that both trace gases do not behave identically at the anthropogenic stations. All above named arguments,
 445 increased secondary formation or potential local biogenic sources, hold for HCHO and CHOCHO, therefore, an important piece of information is still missing.

Overall, the temperature dependency of R_{GF}^* originates from the strong exponential increase of HCHO dSCDs across all stations, which starts around 15°C. As HCHO dSCDs are used in the denominator, R_{GF}^* decreases hyperbolically.

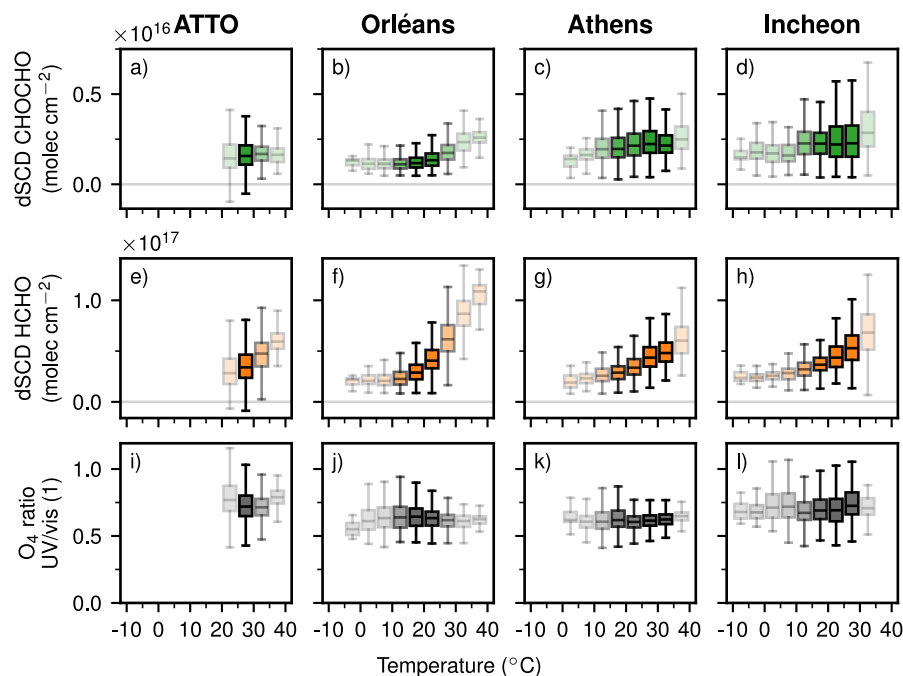


Figure 14. CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) at different temperature intervals for ATTO, Orléans, Athens, and Incheon. The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Missing box plots indicate that no data points fall within that interval. Only data from the lowest viewing elevation is included. Panels a) + e) are self-created based on Donner (2024).

To summarize, R_{GF}^* decreases with higher temperatures in our dataset, which is driven by the strong exponential increase with temperature from HCHO. Similar to the other chapters, this complicates the interpretation of R_{GF} as a proxy for VOC origin, as the R_{GF}^* values depend on temperature regardless of the environment of the sites. This has to be considered in the interpretation of R_{GF}^* when using simple thresholds.

3.6 NO_2 dependence

For R_{GF} usage to discriminate between different VOC origin, it is important to examine how R_{GF} reacts to changing NO_2 levels as NO_2 is a good indicator for anthropogenic activity.

Various studies investigated the NO_2 dependency in the past. Vrekoussis et al. (2010) using GOME-2 satellite data reported a clear link between R_{GF} and NO_2 levels, with smaller R_{GF} values found in polluted environments. Other studies agreed on this finding, e.g. Hoque et al. (2018a) using VCDs from MAX-DOAS in Phimai (Thailand), Xing et al. (2020) using VCDs from MAX-DOAS in Chongqing (China), Hong et al. (2024) using with VCDs from MAX-DOAS in four megacities (China). Chan Miller et al. (2017), however, observed no clear dependence on NO_2 levels using in-situ data from the flight-days with

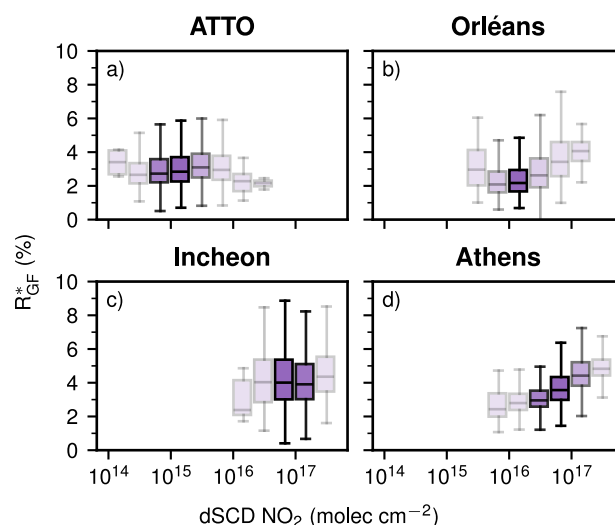


Figure 15. R_{GF}^* at different NO_2 dSCD intervals for ATTO a), Orléans b), Incheon c) and Athens d). The stations are sorted in a clockwise order from the most biogenic environment (ATTO, top left) to the most anthropogenic environment (Incheon, bottom left). The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Missing box plots indicate that no data points fall within that interval. Only data from the lowest viewing elevation is included.

the SENEX aircraft. Another study, by Chen et al. (2023) using TROPOMI satellite data found the opposite trend, where R_{GF} increased with NO_2 levels.

Figure 15 shows the R_{GF}^* dependency on NO_2 levels for all of our sites. The four stations cover a wide range of NO_2 dSCD values, from 10^{14} to 10^{18} molec cm^{-2} . R_{GF}^* does not show a common behaviour for all stations. They can be grouped in two
 465 categories: stations where no clear correlation is observed (ATTO and Incheon) and stations where R_{GF}^* increases with higher NO_2 dSCD (Orléans and Athens).

Figure 16 presents how both OVOCs and the O_4 ratio depend on NO_2 levels. For the first group of stations where R_{GF}^* shows no clear correlation to NO_2 , we see that both trace gases increase with NO_2 and this is cancelled in the ratio. For the other group, we see a different behaviour in each station. In Orléans, HCHO dSCDs decrease with higher NO_2 levels and
 470 therefore R_{GF}^* increases. In Athens, CHOCHO dSCDs increases more rapidly with higher NO_2 levels compared to HCHO dSCDs. This rapid increase is mirrored in the increased R_{GF}^* at high NO_2 levels.

The different behaviour of both anthropogenic stations is interesting. R_{GF}^* exhibits a different behaviour even though, following the interpretation of R_{GF}^* as a proxy for VOC origin, we would expect more similarity. The major difference is that CHOCHO dSCDs increase more rapidly with NO_2 in Athens compared to Incheon. The discrepancy hints at different ratios
 475 between the emissions of NO_x and VOCs at these locations. This can be supported by the relative contributions of the top 5 sectors to the total NMVOC emissions from CAMS-GLOB-ANT in both cities, see Fig. 17. Emissions are similar from the

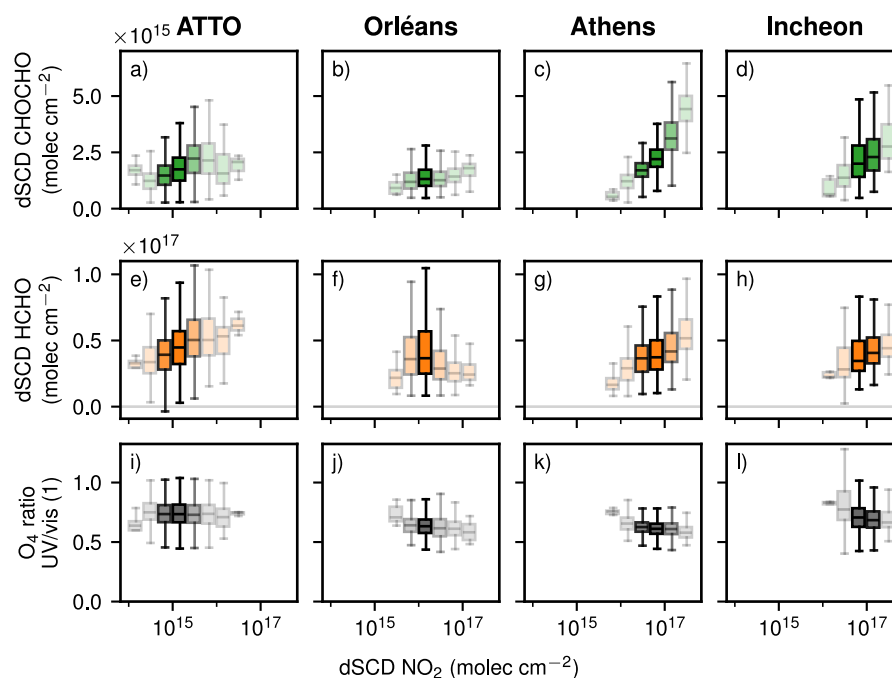


Figure 16. CHOCHO dSCD (top row), HCHO dSCD (center row), and O_4 ratio (bottom row) over NO_2 bins for ATTO, Orléans, Athens, and Incheon. The line within the coloured box indicates the median value, the box itself the IQR. The whiskers extend to 1.5 IQR. To indicate the number of measurements that contribute to each box, the transparency of the box is varied (lower transparency meaning lower number of measurements). Missing box plots indicate that no data points fall within that interval. Only data from the lowest viewing elevation is included. Panels a) + e) are self-created based on Donner (2024).

solvents sector and the road transport. The biggest difference between both sites is the large contribution by industrial processes in Incheon and the strong contributions from refineries and fugitives in Athens.

To summarize, R_{GF}^* shows an inconsistent behaviour with changing NO_2 levels and differs between anthropogenic sites.
 480 This implies that (1) systematic R_{GF}^* differences cannot be reduced only to differing NO_2 levels; R_{GF}^* carries additional environmental information. And (2) local factors strongly influence R_{GF}^* , so using it as a proxy for VOC sources likely requires site-specific considerations.

3.7 Comparability between different R_{GF}

Measurements of R_{GF} reported in the literature lead to varying and partially contradicting results. This section examines effects
 485 that may hinder comparisons between R_{GF} from different measurement techniques. We focus on four key effects and discuss how each is expected to influence the R_{GF} values.

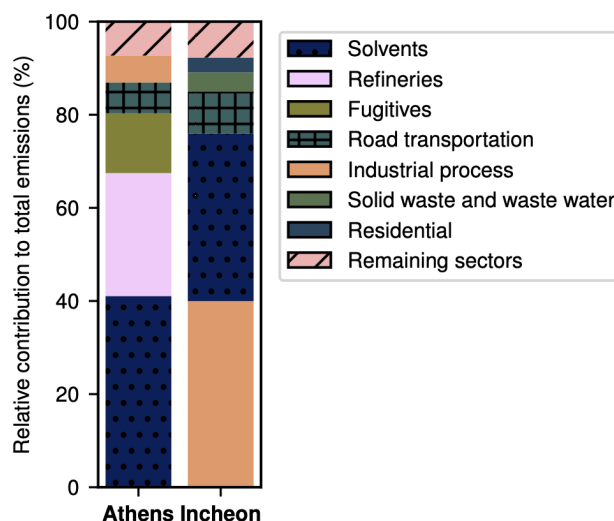


Figure 17. Relative contributions to the CAMS NMVOC emissions of the top 5 sectors in Athens and Incheon. All remaining sectors are summarized in one element. Colormap from Crameri et al. (2020).

3.7.1 Measurement volume

R_{GF} has been computed from data gathered by various different platforms and with different measurement techniques since its first usage. Table 1 shows an overview of the different computations found in literature. Volume mixing ratios (VMRs), dSCDs with correction terms, and mean VCDs have been used to compute R_{GF} . All of these quantities represent R_{GF} in a different measurement volume. For the particular case of VMR R_{GF} and satellite column-averaged R_{GF} , DiGangi et al. (2012) discusses possible causes for disagreements and also briefly mentions the topic of different measurement volumes. We want to further generalize and expand on this inherent difference between the measurement techniques.

Firstly, VMRs obtained by in-situ measurements determine the R_{GF} at the position of the instrument at the sampling time. Here R_{GF} represents the smallest measurement volume, a point measurement.

For R_{GF} values computed via dSCDs from a low elevation angle with O_4 correction (this work), the situation is similar to R_{GF} via VMRs. However, a different volume is probed. Looking towards the horizon, the retrieved dSCDs are dominated by absorption in the lowest layer. Therefore, the resulting R_{GF} is dominated by the volume along the average light path close to the surface until the scattering point.

Lastly, there is column-averaged R_{GF} from either ground-based instruments or satellite-based instruments. Both platforms allow to probe the whole atmospheric column, however with different vertical sensitivities, see Sect. 3.7.2. The column-averaged R_{GF} represents the whole column including the vertical information about the trace gases. However, satellite columns are obtained for the whole ground pixel area, which is larger than the inherent spatial averaging for ground-based columns due to the field of view (FOV).



Table 1. List of different ways to compute R_{GF}

Reference	Method	R_{GF} computation
Vrekoussis et al. (2010)	GOME-2	Annual mean VCDs
DiGangi et al. (2012)	In-situ	surface VMRs ¹
Kluge et al. (2020)	Aircraft	NEMRs ²
Lerot et al. (2023)	TROPOMI	dSCDs multiplied with NO ₂ ratio
Hong et al. (2024)	MAX-DOAS	VMRs removing primary HCHO
This study	MAX-DOAS	dSCDs multiplied with O ₄ ratio

¹ volume mixing ratio

² normalised excess mixing ratio

505 So even though, all ratios of CHOCHO and HCHO are called R_{GF} , they do not necessarily represent the same measurement volume. Different measurement volumes go along with different kinds of averaging or no averaging at all in the case of in-situ R_{GF} . For the comparability of R_{GF} , it should be noted that processes of different scales (spatial or temporal) contribute differently to the R_{GF} from different measurement volumes.

3.7.2 Vertical sensitivity

510 As discussed in the validation study of the TROPOMI HCHO product using ground-based MAX-DOAS observations by De Smedt et al. (2021), satellites and ground-based MAX-DOAS instruments have opposite vertical sensitivity profiles. Satellite-based instruments have minimal sensitivity near the surface, whereas MAX-DOAS instruments are most sensitive at the surface, with sensitivity decreasing to near zero above approximately 3 km altitude. De Smedt et al. (2021) found that accounting for these sensitivity differences can reduce the bias between the two platforms by up to 20%.

515 For R_{GF} , this implies that satellite-derived values are biased toward higher atmospheric layers compared to ground-based measurements, even when vertical profiles or vertical column densities (VCDs) are used. Notably, previous studies have shown that R_{GF} can vary with altitude. For example, Xing et al. (2020) demonstrated that the diurnal behaviour of R_{GF} changes significantly within the lowest 1 km, which may help explain some discrepancies between satellite and ground-based observations.

520 3.7.3 Time dependence

Pronounced diurnal and seasonal cycles in R_{GF}^* are visible at anthropogenic sites in our dataset. In the presence of such cycles, the time/period of measurement becomes critical. For short duration campaigns, the seasonal cycle has to be considered to avoid a sampling bias.

525 The diurnal cycle plays an important role when intercomparing satellites or comparing satellites with ground-based instruments. Sun-synchronous low-Earth orbit satellites, such as the one hosting the TROPOMI instrument, pass at a fixed local solar time over the equator and thus only capture a snapshot of the diurnal variability. Given that our observed diurnal cycles



are relatively symmetric around noon and the overpass times surround noon (see Fig. 4), only minor differences are expected between commonly used satellite instruments such as GOME-2, SCIAMACHY, TROPOMI, and OMI. Only for Athens, the diurnal cycle is shifted to earlier hours, so a notable effect is observed: the measurements during morning overpass are biased higher by approximately 0.5 %pt. than the afternoon.

When comparing satellite measurements to ground-based instruments, however, systematic differences can emerge in daily averages. In the most extreme case, for Incheon, this could result in an overestimation by TROPOMI of about 0.5 %pt. relative to the daily average. Importantly, since diurnal variability is most pronounced at anthropogenic sites, the magnitude of this effect differs across environments. Consequently, a spatially variable bias is expected between studies relying solely on satellite data and those based on ground-based observations. When directly comparing both platforms, it is important to use only data close to the overpass time to eliminate this bias. It is worth noting that new and upcoming geostationary satellites (e.g., GEMS, TEMPO, Sentinel-4) provide diurnal coverage, which should help eliminate such biases when comparing R_{GF} from different platforms.

3.7.4 Impact of averaging-ratioing order

Next to R_{GF} being computed for different measurement volumes, there are two different methodologies found in the literature for the computation of R_{GF} values. Firstly, R_{GF} as the mean of the individual ratios (in the following called instantaneous R_{GF}) and, secondly, R_{GF} as the ratio of the mean of the HCHO and CHOCHO columns (in the following called global R_{GF}).

Both approaches can be applied to any dataset, but in practise the global ratio is often used for R_{GF} based on satellite data. Satellite retrievals are more challenging than ground-based retrievals: the increased distance to surface-level trace gases and the satellite viewing geometry result in lower sensitivity near the surface (Lerot et al., 2021) and the short integration time limits the signal to noise ratio of the individual measurement. To improve the signal-to-noise ratio, satellite measurements are commonly averaged over a defined period and area (Lerot et al., 2021) before calculating R_{GF} from the averaged VCDs. The instantaneous R_{GF} is primarily applied for datasets from ground-based instruments, as in this work, since the ground-based instruments generally provide a higher signal-to-noise ratio due to a longer integration time.

The order of operations matters as the division and the mean do not commute in general. So the ratio of means is not the mean of ratios. Equation (6) shows this for two datasets $A = \{1, 2\}$ and $B = \{3, 4\}$ in a simplified form, i.e. taking the sum instead of the mean.

$$\sum \frac{A}{B} = \frac{1}{3} + \frac{2}{4} = \frac{5}{6} \neq \frac{3}{7} = \frac{1+2}{3+4} = \frac{\sum A}{\sum B} \quad (6)$$

Coming back to R_{GF} : Both ratios are shown in Eq. (7) and Eq. (8). Here N refers to the number of CHOCHO dSCDs and M refers to the number of HCHO dSCDs; N and M are not necessarily identical.

$$R_{GF}^{\text{instantaneous}} = \frac{1}{N} \sum_i^N \frac{\text{dSCD}_i^{\text{CHOCHO}}}{\text{dSCD}_i^{\text{HCHO}}} \quad (7)$$

$$R_{GF}^{\text{global}} = \frac{\frac{1}{N} \sum_i^N \text{dSCD}_i^{\text{CHOCHO}}}{\frac{1}{M} \sum_i^M \text{dSCD}_i^{\text{HCHO}}} \quad (8)$$



Another interesting aspect is the application of quality filters. The instantaneous R_{GF} requires pairs of simultaneous CHOCHO and HCHO measurements, therefore ensuring direct comparability but reducing data coverage. The global R_{GF} is more for-
 560 giving and allows filtering every trace gas individually, leading to potential sampling biases. If valid CHOCHO data occur mainly in summer while valid HCHO measurements are available throughout the whole year, the resulting global R_{GF} would mix a seasonal average with an annual average and thus misrepresent the true annual relationship of CHOCHO and HCHO.

As the usage of the global R_{GF} is required for practical reasons, we investigate how both approaches differ by applying both methodologies to our ground-based dataset. The quality filters are applied in a way, consistent with the previous chapters, that
 565 only valid pairs of simultaneous CHOCHO and HCHO measurements are considered for the analysis.

The instantaneous R_{GF} consistently yields higher values compared to the global R_{GF} across all analyses in this study, but the largest discrepancies occur in hourly binned data, see Fig. 18. It shows the diurnal cycle of R_{GF} (without O4 correction) computed with both methodologies. A clear systematic bias is visible for the two methods, and the magnitude of this bias varies depending on the station and time of day.

At the ATTO Tower (Fig. 18a), a large systematic bias is present throughout the entire day. The strongest difference is
 570 observed in Orléans (Fig. 18b), with the largest discrepancies just below 1 %pt. occurring around 10am. In contrast, at the anthropogenic stations Incheon and Athens (Fig. 18c+d), the two approaches yield more similar results, with smaller differences overall.

Diving into the 10am bin in Orléans and examining the underlying distributions allows highlighting the differences. As
 575 illustrated in Fig. 19, the instantaneous R_{GF} directly incorporates the full distributions of HCHO and CHOCHO dSCDs. By calculating the ratio for each individual measurement, a resulting distribution of R_{GF} values (shown in light purple on the right) is obtained, from which the instantaneous R_{GF} value is computed. In contrast, to compute the global R_{GF} , the HCHO and CHOCHO dSCD distributions are first averaged and then the averaged dSCDs are used to compute the mean R_{GF} value.

Because the underlying distributions and their shapes depend on the data points within each bin, the bias between the two
 580 approaches varies across different bins. However, in all our tests, the ground processing consistently yielded higher R_{GF} values. This can be explained by the occurrence of small HCHO dSCDs in the denominator, which amplify the calculated ratio. To put simply, the difference depends on the asymmetry of the R_{GF} distribution, as greater skewness shifts the instantaneous R_{GF} further away from the global R_{GF} mean. However, this simplified view ignores the interactions, as the skewness of the R_{GF} distribution depends on the underlying HCHO and CHOCHO dSCD distributions and HCHO–CHOCHO correlation in the
 585 bin. For a more technical mathematical argument under which circumstances the ratio of means can be identical to the mean of ratios for two random variables, see Heijmans (1999).

4 Conclusions

Over the past decade, the literature has reported multiple inconclusive findings regarding R_{GF} and its use as a proxy for VOC source identification. In this study, we use a multi-year ground-based MAX-DOAS dataset at four stations to revisit R_{GF} and
 590 reassess its drivers and limitations. Our dataset includes four MAX-DOAS stations located in different environments, allowing

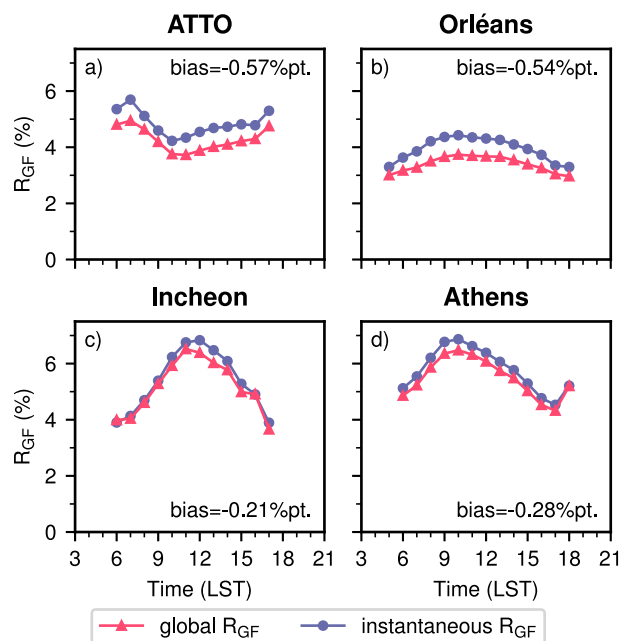


Figure 18. Diurnal cycles of R_{GF} (without O_4 correction) for ATTO a), Orléans b), Incheon c) and Athens d) relative to local solar time (LST). The purple line with round markers corresponds to instantaneous R_{GF} values and the red line with the triangle markers corresponds to global R_{GF} values. The mean bias over the day is shown in the legend. Only data from the lowest viewing elevation for every station is included.

us to systematically investigate patterns in the data. Additionally, we compare the results with various meteorological variables and other trace gases.

We find differences in the absolute magnitudes of R_{GF}^* across environments: lower values at the biogenic sites (ATTO and Orléans) and high values at the strongly anthropogenic sites (Incheon and Athens). While the dSCDs of CHOCHO and HCHO are similarly high across all stations, both trace gases show different behaviours. Glyoxal is notably enhanced at the anthropogenic sites and serves as the primary factor driving the differences in absolute R_{GF}^* magnitudes. This offset is consistently observed in the seasonal, weekly, and diurnal cycles.

In addition, we observe a seasonal cycle characterised by higher R_{GF}^* values during the Northern Hemisphere winter months and lower values during the summer months, primarily driven by the pronounced seasonal cycle of HCHO, which is strongly linked to temperature. This pattern holds across all stations except ATTO, where the seasonal cycle is shifted by several months and exhibits enhanced values during the wet and reduced values during the dry season (Donner, 2024).

The diurnal cycles of R_{GF}^* are relatively flat at the biogenic stations but pronounced at the anthropogenic stations, showing variations of approximately 2 %pt. with peaks occurring before or around noon. Glyoxal, in particular, exhibits a distinct diurnal pattern at the anthropogenic sites compared to the biogenic sites, making it the main driver of the observed diurnal R_{GF}^* variability.

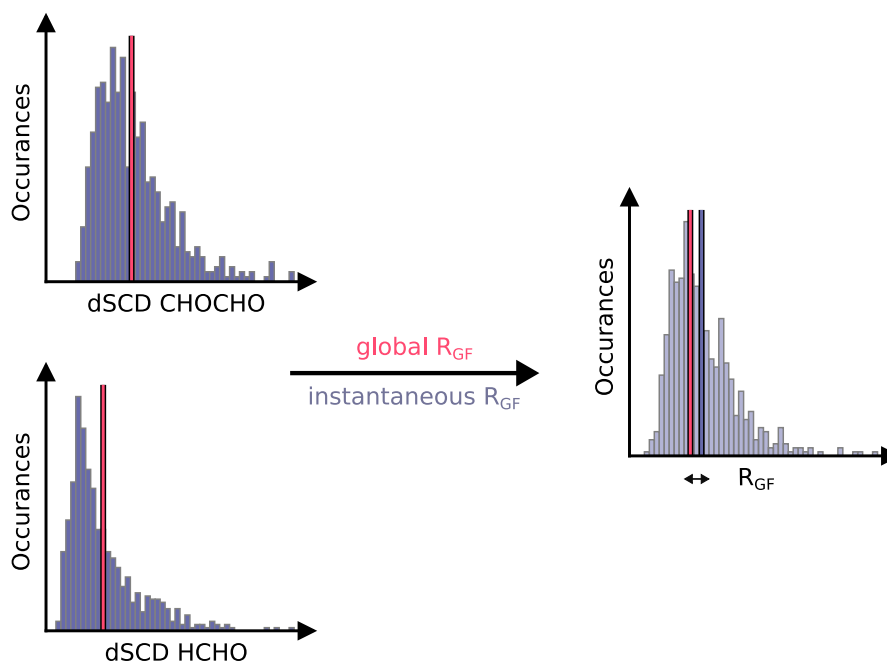


Figure 19. Visualization of the different averaging approaches using the 10am bin from the lowest viewing elevation from Orléans. The distributions of CHOCHO dSCD (top) and of HCHO dSCD (bottom) are shown on the left, with the respective mean indicated with the red vertical line. On the right, the distribution of R_{GF} and the instantaneous R_{GF} (purple vertical line) is shown. The vertical red line corresponds to the global R_{GF} .

We also detect a weak weekend effect at the anthropogenic stations, with R_{GF}^* decreasing by about 10% compared to weekdays, primarily due to a stronger reduction by CHOCHO on the weekend.

Moreover, we report a clear temperature dependence of R_{GF}^* . Above approximately 15°C, HCHO levels increase more strongly with temperature than CHOCHO, leading to a hyperbolic decrease in R_{GF}^* beyond this temperature.

610 Our investigation of the dependence of R_{GF}^* on NO_2 implies that R_{GF}^* cannot be reduced only to different NO_2 levels. Local effects and VOC emission characteristics may play a more significant role than previously assumed. Thus site-specific considerations for the usage of R_{GF}^* as a proxy might be required. Overall, CHOCHO seems to be stronger connected to anthropogenic activity than HCHO.

615 Finally, we examine four factors that can impact comparisons of R_{GF} from different datasets. Firstly, R_{GF} , computed from different measurement techniques, inherently average over different measurement volumes. In addition, the vertical sensitivity of satellites and ground-based instruments is not identical. Moreover, due to the pronounced seasonal and diurnal cycles the time of the measurement becomes critical. No significant difference between different overpass times of GOME-2, SCIAMACHY, and TROPOMI/OMI is observed due to the symmetric diurnal cycles, however the overpass time has to be considered



for ground-based and satellite comparisons. Lastly, we investigate the impact of the order of ratioing and averaging. The global
620 R_{GF} is generally biased low compared to the instantaneous R_{GF} as no extreme values occur in the denominator.

Further work is needed to expand our understanding and strengthen the statistical basis through additional multi-year ground-based observations. Improved coverage of simultaneous VOC precursor measurements would further enhance the interpretation of observational data. An important complementary perspective could come from modeling studies. In particular, sensitivity studies using chemical box models (including gas-phase and heterogeneous chemistry) offer a unique opportunity to challenge
625 current understanding. By systematically turning individual processes on or off, such models can help disentangle complex interconnections that are difficult to isolate in observational data. Additionally, distinguishing between the primary and secondary formation of HCHO and CHOCHO in the context of R_{GF} could help clarify the underlying mechanisms (Hong et al., 2024; Xing et al., 2025).

Given these complexities, using the glyoxal-to-formaldehyde ratio as a proxy for VOC source identification remains chal-
630 lenging. While we observe clear differences in the absolute magnitudes of R_{GF}^* across stations in different environments, suggesting that the ratio carries valuable environmental information, our incomplete understanding of emissions, VOC precursors, and the detailed chemistry of HCHO and CHOCHO currently limits the reliable use of R_{GF}^* as a robust proxy for VOC source attribution.

Data availability. The ERA5 atmospheric reanalysis data were downloaded from the from the Copernicus Climate Change (C3S) climate
635 data store (CDS): <https://doi.org/10.24381/cds.adbb2d47> (Copernicus Climate Change Service, Climate Data Store, 2023). CAMS-GLOB-ANT version 6.2 emission dataset was downloaded from ECCAD: <https://permalink.aeris-data.fr/CAMS-GLOB-ANT> (Granier et al., 2019). MAX-DOAS data are available from the authors on request.

Author contributions. SB, AR, and MV contributed to the conceptualization, methodology, data interpretation and writing of the original draft. SB and BZ set-up and operated the instrument in Orléans. SD and TW conducted the measurements and provided the ATTO data set.
640 LA provided expertise regarding the fit settings for the Bremen instruments. SB processed the data, performed the data analysis. All authors have contributed with scientific discussions to data interpretation.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.

Disclaimer. The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.



645 *Acknowledgements.* AI Usage: Reformatting tables, Improving code for figures, indexing literature, Improved text readability, and spelling correction.

Simulations were performed on the HPC cluster Aether at the University of Bremen, financed by the Deutsche Forschungsgemeinschaft (DFG) within the scope of the Excellence Initiative

650 The ATTO site is operated and managed by INPA and MPG. The authors also acknowledge the whole ATTO team for maintenance and operation of the ATTO site and its infrastructure. Further, Bianca Lauster and Steffen Ziegler are acknowledged for their large contribution in operating and maintaining the ATTO MAX-DOAS instrument. The ATTO research has been supported by the German Federal Ministry of Education and Research (BMBF contract 01LK2101B).

SB, MV, AR acknowledge financial support from the University of Bremen. SB acknowledge financial support from the CLIMADEMY ERASMUS+ (101056066) Grant. MV acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (University Allowance, EXC 2077, University of Bremen).

655



References

- Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Hedegaard, A. B. K., Schneising, O., and Burrows, J. P.: Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus Sentinel-5 Precursor satellite during the 2018 Canadian wildfires, *Atmospheric Chemistry and Physics*, 20, 2057–2072, <https://doi.org/10.5194/acp-20-2057-2020>, 2020.
- 660 Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, *Global Biogeochemical Cycles*, 15, 955–966, <https://doi.org/10.1029/2000gb001382>, 2001.
- Andreae, M. O., Acevedo, O. C., Araújo, A., Artaxo, P., Barbosa, C. G. G., Barbosa, H. M. J., Brito, J., Carbone, S., Chi, X., Cintra, B. B. L., da Silva, N. F., 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., de Oliveira Sá, M., Sales, C. B., dos Santos, R. M. N., Saturno, J., Schöngart, J., Sörgel, M., de Souza, C. M., de Souza, R. A. F., Su, H., Targhetta, N., Tóta, J., Trebs, I., Trumbore, S., van Eijck, A., Walter, D., Wang, Z., Weber, B., Williams, J., Winderlich, J., Wittmann, F., Wolff, S., and Yáñez-Serrano, A. M.: The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols, *Atmospheric Chemistry and Physics*, 15, 10 723–10 776, <https://doi.org/10.5194/acp-15-10723-2015>, 2015.
- 670 Beirle, S., Platt, U., Wenig, M., and Wagner, T.: Weekly cycle of NO₂ by GOME measurements: a signature of anthropogenic sources, *Atmospheric Chemistry and Physics*, 3, 2225–2232, <https://doi.org/10.5194/acp-3-2225-2003>, 2003.
- Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., and Dominici, F.: Ozone and Short-term Mortality in 95 US Urban Communities, 1987–2000, *JAMA*, 292, 2372, <https://doi.org/10.1001/jama.292.19.2372>, 2004.
- 675 Bello, J., Vallet, P., Perot, T., Balandier, P., Seigner, V., Perret, S., Couteau, C., and Korboulewsky, N.: How do mixing tree species and stand density affect seasonal radial growth during drought events?, *Forest Ecology and Management*, 432, 436–445, <https://doi.org/10.1016/j.foreco.2018.09.044>, 2019.
- Berg, F., Novelli, A., Dubus, R., Hofzumahaus, A., Holland, F., Wahner, A., and Fuchs, H.: Temperature-dependent rate coefficients for the reactions of OH radicals with selected alkanes, aromatic compounds, and monoterpenes, *Atmospheric Chemistry and Physics*, 24, 13 715–13 731, <https://doi.org/10.5194/acp-24-13715-2024>, 2024.
- 680 Bourtsoukidis, E., Pozzer, A., Williams, J., Makowski, D., Peñuelas, J., Matthaïos, V. N., Lazoglou, G., Yáñez-Serrano, A. M., Lelieveld, J., Ciais, P., Vrekoussis, M., Daskalakis, N., and Sciare, J.: High temperature sensitivity of monoterpene emissions from global vegetation, *Communications Earth & Environment*, 5, <https://doi.org/10.1038/s43247-023-01175-9>, 2024.
- Castro-Melgar, I., Tsagkou, A., Zacharopoulou, M., Basiou, E., Athinelis, I., Katris, E.-A., Kalavrezou, I.-E., and Parcharidis, I.: Wildfires During Early Summer in Greece (2024): Burn Severity and Land Use Dynamics Through Sentinel-2 Data, *Forests*, 16, 268, <https://doi.org/10.3390/f16020268>, 2025.
- Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J., and Chance, K.: Glyoxal retrieval from the Ozone Monitoring Instrument, *Atmospheric Measurement Techniques*, 7, 3891–3907, <https://doi.org/10.5194/amt-7-3891-2014>, 2014.
- Chan Miller, C., Jacob, D. J., González Abad, G., and Chance, K.: Hotspot of glyoxal over the Pearl River delta seen from the OMI satellite instrument: implications for emissions of aromatic hydrocarbons, *Atmospheric Chemistry and Physics*, 16, 4631–4639, <https://doi.org/10.5194/acp-16-4631-2016>, 2016.
- 690



- 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., Abad, G. G., and Chance, K.: Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data, *Atmospheric Chemistry and Physics*, 17, 8725–8738, <https://doi.org/10.5194/acp-17-8725-2017>, 2017.
- 695 Chance, K. and Kurucz, R.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 1289–1295, <https://doi.org/10.1016/j.jqsrt.2010.01.036>, 2010.
- Chen, Y., Liu, C., Su, W., Hu, Q., Zhang, C., Liu, H., and Yin, H.: Identification of volatile organic compound emissions from anthropogenic and biogenic sources based on satellite observation of formaldehyde and glyoxal, *Science of The Total Environment*, 859, 159 997, <https://doi.org/10.1016/j.scitotenv.2022.159997>, 2023.
- 700 Copernicus Climate Change Service, Climate Data Store: ERA5 hourly data on single levels from 1940 to present, <https://doi.org/10.24381/cds.adbb2d47>, copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023.
- Cramer, F., Shephard, G. E., and Heron, P. J.: The misuse of colour in science communication, *Nature Communications*, 11, <https://doi.org/10.1038/s41467-020-19160-7>, 2020.
- 705 Danckaert, T., Fayt, C., van Roozendaal, M., de Smedt, I., Letocart, V., Merlaud, A., and Pinardi, G.: QDOAS Software user manual - Version 3.2, Royal Belgian Institute for Space Aeronomy, https://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf, 2017.
- De Smedt, I., Pinardi, G., Vigouroux, C., Compernelle, 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., PETERS, 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 Roozendaal, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, *Atmospheric Chemistry and Physics*, 21, 12 561–12 593, <https://doi.org/10.5194/acp-21-12561-2021>, 2021.
- 710 Derwent, R. G., Jenkin, M. E., Utembe, S. R., Shallcross, D. E., Murrells, T. P., and Passant, N. R.: Secondary organic aerosol formation from a large number of reactive man-made organic compounds, *Science of The Total Environment*, 408, 3374–3381, <https://doi.org/10.1016/j.scitotenv.2010.04.013>, 2010.
- 715 Dienhart, D., Crowley, J. N., Bourtsoukidis, E., Edtbauer, A., Eger, P. G., Ernle, L., Harder, H., Hottmann, B., Martinez, M., Parchatka, U., Paris, J.-D., Pfannerstill, E. Y., Rohloff, R., Schuladen, J., Stöner, C., Tadic, I., Tauer, S., Wang, N., Williams, J., Lelieveld, J., and Fischer, H.: Measurement report: Observation-based formaldehyde production rates and their relation to OH reactivity around the Arabian Peninsula, *Atmospheric Chemistry and Physics*, 21, 17 373–17 388, <https://doi.org/10.5194/acp-21-17373-2021>, 2021.
- 720 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.: Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, *Atmospheric Chemistry and Physics*, 12, 9529–9543, <https://doi.org/10.5194/acp-12-9529-2012>, 2012.
- Donner, S.: Retrieving vertical profiles and tropospheric columns of formaldehyde from global long-term MAX-DOAS measurements, Ph.D. thesis, Johannes Gutenberg-Universität Mainz, <https://doi.org/10.25358/OPENSOURCE-11085>, 2024.
- 725 European Environment Agency: Harm to human health from air pollution in Europe, EEA Briefing, Publications Office, <https://doi.org/10.2800/3950756>, 2024.



- Finkenzeller, H. and Volkamer, R.: O₂–O₂ CIA in the gas phase: Cross-section of weak bands, and continuum absorption between 297–500 nm, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 279, 108 063, <https://doi.org/10.1016/j.jqsrt.2021.108063>, 2022.
- Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, *Journal of Photochemistry and Photobiology A: Chemistry*, 168, 117–132, <https://doi.org/10.1016/j.jphotochem.2004.03.026>, 2004.
- Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, *Atmospheric Chemistry and Physics*, 12, 6699–6721, <https://doi.org/10.5194/acp-12-6699-2012>, 2012.
- Franco, B., Marais, E. A., Bovy, B., Bader, W., Lejeune, B., Roland, G., Servais, C., and Mahieu, E.: Diurnal cycle and multi-decadal trend of formaldehyde in the remote atmosphere near 46N, *Atmospheric Chemistry and Physics*, 16, 4171–4189, <https://doi.org/10.5194/acp-16-4171-2016>, 2016.
- Fu, T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M., and Henze, D. K.: Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, *Journal of Geophysical Research*, 113, <https://doi.org/10.1029/2007jd009505>, 2008.
- Gilman, J. B., Lerner, B. M., Kuster, W. C., Goldan, P. D., Warneke, C., Veres, P. R., Roberts, J. M., de Gouw, J. A., Burling, I. R., and Yokelson, R. J.: Biomass burning emissions and potential air quality impacts of volatile organic compounds and other trace gases from fuels common in the US, *Atmospheric Chemistry and Physics*, 15, 13 915–13 938, <https://doi.org/10.5194/acp-15-13915-2015>, 2015.
- Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, *Environmental Science & Technology*, 41, 1514–1521, <https://doi.org/10.1021/es072476p>, 2007.
- Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version), <https://doi.org/10.24380/D0BN-KX16>, 2019.
- Gratsea, M., Vrekoussis, M., Richter, A., Wittrock, F., Schönhardt, A., Burrows, J., Kazadzis, S., Mihalopoulos, N., and Gerasopoulos, E.: Slant column MAX-DOAS measurements of nitrogen dioxide, formaldehyde, glyoxal and oxygen dimer in the urban environment of Athens, *Atmospheric Environment*, 135, 118–131, <https://doi.org/10.1016/j.atmosenv.2016.03.048>, 2016.
- Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., McKay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., and Zimmerman, P.: A global model of natural volatile organic compound emissions, *Journal of Geophysical Research: Atmospheres*, 100, 8873–8892, <https://doi.org/10.1029/94jd02950>, 1995.
- Guenther, A. B., Zimmerman, P. R., Harley, P. C., Monson, R. K., and Fall, R.: Isoprene and monoterpene emission rate variability: Model evaluations and sensitivity analyses, *Journal of Geophysical Research: Atmospheres*, 98, 12 609–12 617, <https://doi.org/10.1029/93jd00527>, 1993.
- Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, *Geoscientific Model Development*, 5, 1471–1492, <https://doi.org/10.5194/gmd-5-1471-2012>, 2012.
- Guo, Y., Wang, S., Zhu, J., Zhang, R., Gao, S., Saiz-Lopez, A., and Zhou, B.: Atmospheric formaldehyde, glyoxal and their relations to ozone pollution under low- and high-NO_x regimes in summertime Shanghai, China, *Atmospheric Research*, 258, 105 635, <https://doi.org/10.1016/j.atmosres.2021.105635>, 2021.



- Haagen-Smit, A. J.: Chemistry and Physiology of Los Angeles Smog, *Industrial & Engineering Chemistry*, 44, 1342–1346, <https://doi.org/10.1021/ie50510a045>, 1952.
- Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, *Atmospheric Chemistry and Physics*, 9, 5155–5236, <https://doi.org/10.5194/acp-9-5155-2009>, 2009.
- Heijmans, R.: When does the expectation of a ratio equal the ratio of expectations?, *Statistical Papers*, 40, 107–115, <https://doi.org/10.1007/bf02927114>, 1999.
- Hellenic Statistical Authority: Attiki census results of population and housing - ELSTAT 2021, https://www.statistics.gr/documents/20181/18409455/population_ATTICA_EN.pdf/f8156495-873d-bb66-bf8e-b87bf6974e14?t=1725436907085, last access: 25 May 2025, 2024.
- 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.: ERA5 hourly data on single levels from 1940 to present, <https://doi.org/10.24381/CDS.ADBB2D47>, copernicus Climate Change Service (C3S) Climate Data Store (CDS), Accessed on DD-MMM-YYYY, 2023.
- Hong, Q., Xing, J., Xing, C., Yang, B., Su, W., Chen, Y., Zhang, C., Zhu, Y., and Liu, C.: Investigating vertical distributions and photochemical indications of formaldehyde, glyoxal, and NO₂ from MAX-DOAS observations in four typical cities of China, *Science of The Total Environment*, 954, 176 447, <https://doi.org/10.1016/j.scitotenv.2024.176447>, 2024.
- Hoque, H. M. S., Irie, H., and Damiani, A.: First MAX-DOAS Observations of Formaldehyde and Glyoxal in Phimai, Thailand, *Journal of Geophysical Research: Atmospheres*, 123, 9957–9975, <https://doi.org/10.1029/2018jd028480>, 2018a.
- Hoque, H. M. S., Irie, H., Damiani, A., Rawat, P., and Naja, M.: First Simultaneous Observations of Formaldehyde and Glyoxal by MAX-DOAS in the Indo-Gangetic Plain Region, *SOLA*, 14, 159–164, <https://doi.org/10.2151/sola.2018-028>, 2018b.
- Hönninger, G., von Friedeburg, C., and Platt, U.: Multi axis differential optical absorption spectroscopy (MAX-DOAS), *Atmospheric Chemistry and Physics*, 4, 231–254, <https://doi.org/10.5194/acp-4-231-2004>, 2004.
- Institut national de la statistique et des études économiques: Populations de référence 2022 - Commune d’Orléans (45234), <https://www.insee.fr/fr/statistiques/8288323?geo=COM-45234>, last access: 25 May 2025, 2024.
- Institut national de la statistique et des études économiques: Dossier complet - Commune de Traînou (45327), <https://www.insee.fr/fr/statistiques/2011101?geo=COM-45327>, last access: 25 May, 2025.
- Instituto Brasileiro de Geografia e Estatística: Panorama Censo 2022 - Manaus, <https://censo2022.ibge.gov.br/panorama/>, last access: 25 May 2025, 2022.
- IPCC: Climate Change 2001: The Scientific Basis, vol. 881., Cambridge University Press, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.)], 2001.
- Irie, H., Takashima, H., Kanaya, Y., Boersma, K. F., Gast, L., Wittrock, F., Brunner, D., Zhou, Y., and Van Roozendael, M.: Eight-component retrievals from ground-based MAX-DOAS observations, *Atmospheric Measurement Techniques*, 4, 1027–1044, <https://doi.org/10.5194/amt-4-1027-2011>, 2011.
- 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.: Reassessing the ratio of glyoxal to formaldehyde



- as an indicator of hydrocarbon precursor speciation, *Atmospheric Chemistry and Physics*, 15, 7571–7583, <https://doi.org/10.5194/acp-15-7571-2015>, 2015.
- Kassomenos, P., Kotroni, V., and Kallos, G.: Analysis of climatological and air quality observations from Greater Athens Area, *Atmospheric Environment*, 29, 3671–3688, [https://doi.org/10.1016/1352-2310\(94\)00358-r](https://doi.org/10.1016/1352-2310(94)00358-r), 1995.
- Kim, M.-J., Baek, K.-M., Heo, J.-B., Cheong, J.-P., and Baek, S.-O.: Concentrations, health risks, and sources of hazardous air pollutants in Seoul-Incheon, a megacity area in Korea, *Air Quality, Atmosphere & Health*, 14, 873–893, <https://doi.org/10.1007/s11869-021-00986-z>, 2021.
- Kluge, F., Hüneke, T., Knecht, M., Lichtenstern, M., Rotermund, M., Schlager, H., Schreiner, B., and Pfeilsticker, K.: Profiling of formaldehyde, glyoxal, methylglyoxal, and CO over the Amazon: normalized excess mixing ratios and related emission factors in biomass burning plumes, *Atmospheric Chemistry and Physics*, 20, 12 363–12 389, <https://doi.org/10.5194/acp-20-12363-2020>, 2020.
- Kurucz, R. L., Furenli, I., Brault, J., and Testerman, L.: Solar flux atlas from 296 to 1300 nm, National Solar Observatory Atlas No. 1, 1984.
- Lagouvardos, K., Kotroni, V., Giannaros, T. M., and Dafis, S.: Meteorological Conditions Conducive to the Rapid Spread of the Deadly Wildfire in Eastern Attica, Greece, *Bulletin of the American Meteorological Society*, 100, 2137–2145, <https://doi.org/10.1175/bams-d-18-0231.1>, 2019.
- Lange, K., Richter, A., Bösch, T., Zilker, B., Latsch, M., Behrens, L. K., Okafor, C. M., Bösch, H., Burrows, J. P., Merlaud, A., Pinardi, G., Fayt, C., Friedrich, M. M., Dimitropoulou, E., Van Roozendaal, M., Ziegler, S., Ripperger-Lukosiunaite, S., Kuhn, L., Lauster, B., Wagner, T., Hong, H., Kim, D., Chang, L.-S., Bae, K., Song, C.-K., Park, J.-U., and Lee, H.: Validation of GEMS tropospheric NO₂ columns and their diurnal variation with ground-based DOAS measurements, *Atmospheric Measurement Techniques*, 17, 6315–6344, <https://doi.org/10.5194/amt-17-6315-2024>, 2024.
- Laothawornkitkul, J., Taylor, J. E., Paul, N. D., and Hewitt, C. N.: Biogenic volatile organic compounds in the Earth system, *New Phytologist*, 183, 27–51, <https://doi.org/10.1111/j.1469-8137.2009.02859.x>, 2009.
- Lee, M., Heikes, B. G., and Jacob, D. J.: Enhancements of hydroperoxides and formaldehyde in biomass burning impacted air and their effect on atmospheric oxidant cycles, *Journal of Geophysical Research: Atmospheres*, 103, 13 201–13 212, <https://doi.org/10.1029/98jd00578>, 1998.
- Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, *Nature*, 525, 367–371, <https://doi.org/10.1038/nature15371>, 2015.
- Lerot, C., Hendrick, F., Roozendaal, M. V., Alvarado, L. M. A., Richter, A., Smedt, I. D., Theys, N., Vlietinck, J., Yu, H., Gent, J. V., 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.: Glyoxal tropospheric column retrievals from TROPOMI - multi-satellite intercomparison and ground-based validation, *Atmospheric Measurement Techniques*, 14, 7775–7807, <https://doi.org/10.5194/amt-14-7775-2021>, 2021.
- Lerot, C., Müller, J.-F., Theys, N., Smedt, I. D., Stavrakou, T., and Roozendaal, M. V.: Satellite Evidence for Glyoxal Depletion in Elevated Fire Plumes, *Geophysical Research Letters*, 50, <https://doi.org/10.1029/2022gl102195>, 2023.
- Li, X., Rohrer, F., Brauers, T., Hofzumahaus, A., Lu, K., Shao, M., Zhang, Y. H., and Wahner, A.: Modeling of HCHO and CHOCHO at a semi-rural site in southern China during the PRIDE-PRD2006 campaign, *Atmospheric Chemistry and Physics*, 14, 12 291–12 305, <https://doi.org/10.5194/acp-14-12291-2014>, 2014.
- Li, X., Zhu, L., De Smedt, I., Sun, W., Chen, Y., Shu, L., Wang, D., Liu, S., Pu, D., Li, J., Zuo, X., Fu, W., Li, Y., Zhang, P., Yan, Z., Fu, T., Shen, H., Wang, C., Ye, J., and Yang, X.: Global Temperature Dependency of Biogenic HCHO Columns Observed



- From Space: Interpretation of TROPOMI Results Using GEOS-Chem Model, *Journal of Geophysical Research: Atmospheres*, 129, <https://doi.org/10.1029/2024jd041784>, 2024.
- Liaskoni, M., Huszár, P., Bartók, L., Prieto Perez, A. P., Karlický, J., and Šindelářová, K.: The long-term impact of biogenic volatile organic compound emissions on urban ozone patterns over central Europe: contributions from urban and rural vegetation, *Atmospheric Chemistry and Physics*, 24, 13 541–13 569, <https://doi.org/10.5194/acp-24-13541-2024>, 2024.
- Liu, M. and Matsui, H.: Secondary Organic Aerosol Formation Regulates Cloud Condensation Nuclei in the Global Remote Troposphere, *Geophysical Research Letters*, 49, <https://doi.org/10.1029/2022gl100543>, 2022.
- 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.: DOAS measurements of formaldehyde and glyoxal above a south-east Asian tropical rainforest, *Atmospheric Chemistry and Physics*, 12, 5949–5962, <https://doi.org/10.5194/acp-12-5949-2012>, 2012.
- Mavroidis, I. and Ilia, M.: Trends of NO_x, NO₂ and O₃ concentrations at three different types of air quality monitoring stations in Athens, Greece, *Atmospheric Environment*, 63, 135–147, <https://doi.org/10.1016/j.atmosenv.2012.09.030>, 2012.
- McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), *Earth System Science Data*, 12, 3413–3442, <https://doi.org/10.5194/essd-12-3413-2020>, 2020.
- Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm, *Journal of Geophysical Research: Atmospheres*, 105, 7089–7101, <https://doi.org/10.1029/1999jd901074>, 2000.
- Myriokefalitakis, S., Vrekoussis, M., Tsigaridis, K., Wittrock, F., Richter, A., Brühl, C., Volkamer, R., Burrows, J. P., and Kanakidou, M.: The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, *Atmospheric Chemistry and Physics*, 8, 4965–4981, <https://doi.org/10.5194/acp-8-4965-2008>, 2008.
- Nelson, P. F., Tibbett, A. R., and Day, S. J.: Effects of vehicle type and fuel quality on real world toxic emissions from diesel vehicles, *Atmospheric Environment*, 42, 5291–5303, <https://doi.org/10.1016/j.atmosenv.2008.02.049>, 2008.
- Nguyen, H. T., Kim, K.-H., and Park, C.: Long-term trend of NO₂ in major urban areas of Korea and possible consequences for health, *Atmospheric Environment*, 106, 347–357, <https://doi.org/10.1016/j.atmosenv.2015.02.003>, 2015.
- 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.: Measurement report: Photochemical production and loss rates of formaldehyde and ozone across Europe, *Atmospheric Chemistry and Physics*, 21, 18 413–18 432, <https://doi.org/10.5194/acp-21-18413-2021>, 2021.
- Paris, E., Carnevale, M., Vincenti, B., Palma, A., Guerriero, E., Borello, D., and Gallucci, F.: Evaluation of VOCs Emitted from Biomass Combustion in a Small CHP Plant: Difference between Dry and Wet Poplar Woodchips, *Molecules*, 27, 955, <https://doi.org/10.3390/molecules27030955>, 2022.
- Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy: Principles and Applications*, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-540-21193-8, https://doi.org/10.1007/978-3-540-75776-4_6, 2008.
- Pusede, S. E., Gentner, D. R., Wooldridge, P. J., Browne, E. C., Rollins, A. W., Min, K.-E., Russell, A. R., Thomas, J., Zhang, L., Brune, W. H., Henry, S. B., DiGangi, J. P., Keutsch, F. N., Harrold, S. A., Thornton, J. A., Beaver, M. R., St. Clair, J. M., Wennberg, P. O., Sanders, J., Ren, X., VandenBoer, T. C., Markovic, M. Z., Guha, A., Weber, R., Goldstein, A. H., and Cohen, R. C.: On the temperature



- dependence of organic reactivity, nitrogen oxides, ozone production, and the impact of emission controls in San Joaquin Valley, California, *Atmospheric Chemistry and Physics*, 14, 3373–3395, <https://doi.org/10.5194/acp-14-3373-2014>, 2014.
- 880 Ramonet, M., Lopez, M., and Delmotte, M.: ICOS ATC NRT CO₂ growing time series from Trainou (50.0 m), <https://hdl.handle.net/11676/3aTWbGOmjVhir3TOGMPcLH3V>, 2025.
- Rawat, P., Naja, M., Rajwar, M. C., Irie, H., Lerot, C., Kumar, M., and Lal, S.: Long-term observations of NO₂, SO₂, HCHO, and CHO-CHO over the Himalayan foothills: Insights from MAX-DOAS, TROPOMI, and GOME-2, *Atmospheric Environment*, 336, 120 746, <https://doi.org/10.1016/j.atmosenv.2024.120746>, 2024.
- 885 Rothman, L., Gordon, I., Barber, R., Dothe, H., Gamache, R., Goldman, A., Perevalov, V., Tashkun, S., and Tennyson, J.: HITEMP, the high-temperature molecular spectroscopic database, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 2139–2150, <https://doi.org/10.1016/j.jqsrt.2010.05.001>, 2010.
- Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P., Birk, M., Bizzocchi, L., Boudon, V., Brown, L., Campargue, A., Chance, K., Cohen, E., Coudert, L., Devi, V., Drouin, B., Fayt, A., Flaud, J.-M., Gamache, R., Harrison, J., Hartmann, J.-M., Hill, C., Hodges, J., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R., Li, G., Long, D., Lyulin, O., Mackie, C., Massie, S., Mikhailenko, S., Müller, H., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E., Richard, C., Smith, M., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G., Tyuterev, V., and Wagner, G.: The HITRAN2012 molecular spectroscopic database, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 130, 4–50, <https://doi.org/10.1016/j.jqsrt.2013.07.002>, 2013.
- 890 Seinfeld, J. H. and Pandis, S. N.: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2 edn., 2006.
- Serdyuchenko, A., Gorshchev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections - Part 2: Temperature dependence, *Atmospheric Measurement Techniques*, 7, 625–636, <https://doi.org/10.5194/amt-7-625-2014>, 2014.
- Silva, S. J., Heald, C. L., and Li, M.: Space-Based Constraints on Terrestrial Glyoxal Production, *Journal of Geophysical Research: Atmospheres*, 123, <https://doi.org/10.1029/2018jd029311>, 2018.
- 900 Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Lioussé, C., Crippa, M., Guizzardi, D., Hoesly, R., and Smith, S. J.: Global anthropogenic emissions (CAMSGLOB-ANT) for the Copernicus Atmosphere Monitoring Service simulations of air quality forecasts and reanalyses, *Earth System Science Data*, 16, 2261–2279, <https://doi.org/10.5194/essd-16-2261-2024>, 2024.
- 905 Stavrakou, T., Müller, J.-F., Smedt, I. D., Roozendaal, M. V., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, *Atmospheric Chemistry and Physics*, 9, 8431–8446, <https://doi.org/10.5194/acp-9-8431-2009>, 2009a.
- Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendaal, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, *Atmospheric Chemistry and Physics*, 9, 1037–1060, <https://doi.org/10.5194/acp-9-1037-2009>, 2009b.
- 910 Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O₂-O₂ collision pairs between 340 and 630 nm and at atmospherically relevant pressure, *Physical Chemistry Chemical Physics*, 15, 15 371, <https://doi.org/10.1039/c3cp50968k>, 2013.
- Turner, M. C., Jerrett, M., Pope, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L., and Burnett, R. T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, *American Journal of Respiratory and Critical Care Medicine*, 193, 1134–1142, <https://doi.org/10.1164/rccm.201508-1633oc>, 2016.



- Vandaele, A. C., Hermans, C., Simon, P. C., Roozendael, M. V., Guilmot, J. M., Carleer, M., and Colin, R.: Fourier transform measurement of NO₂ absorption cross-section in the visible range at room temperature, *Journal of Atmospheric Chemistry*, 25, 289–305, <https://doi.org/10.1007/bf00053797>, 1996.
- Volkamer, R., Spietz, P., Burrows, J., and Platt, U.: High-resolution absorption cross-section of glyoxal in the UV–vis and IR spectral ranges, *Journal of Photochemistry and Photobiology A: Chemistry*, 172, 35–46, <https://doi.org/10.1016/j.jphotochem.2004.11.011>, 2005.
- Volkamer, R., Martini, F. S., Molina, L. T., Salcedo, D., Jimenez, J. L., and Molina, M. J.: A missing sink for gas-phase glyoxal in Mexico City: Formation of secondary organic aerosol, *Geophysical Research Letters*, 34, <https://doi.org/10.1029/2007gl030752>, 2007.
- Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmospheric Chemistry and Physics*, 10, 10 145–10 160, <https://doi.org/10.5194/acp-10-10145-2010>, 2010.
- Wang, H., Li, Y., Liu, Y., Lu, X., Zhang, Y., Fan, Q., Shen, C., Lai, S., Zhou, Y., Zhang, T., and Yue, D.: Underappreciated contributions of biogenic volatile organic compounds from urban green spaces to ozone pollution, *Atmospheric Chemistry and Physics*, 25, 5233–5250, <https://doi.org/10.5194/acp-25-5233-2025>, 2025.
- Wittrock, F., Oetjen, H., Richter, A., Fietkau, S., Medeke, T., Rozanov, A., and Burrows, J. P.: MAX-DOAS measurements of atmospheric trace gases in Ny-Ålesund - Radiative transfer studies and their application, *Atmospheric Chemistry and Physics*, 4, 955–966, <https://doi.org/10.5194/acp-4-955-2004>, 2004.
- Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, *Geophysical Research Letters*, 33, <https://doi.org/10.1029/2006gl026310>, 2006.
- Xing, C., Liu, C., Hu, Q., Fu, Q., Lin, H., Wang, S., Su, W., Wang, W., Javed, Z., and Liu, J.: Identifying the wintertime sources of volatile organic compounds (VOCs) from MAX-DOAS measured formaldehyde and glyoxal in Chongqing, southwest China, *Science of The Total Environment*, 715, 136 258, <https://doi.org/10.1016/j.scitotenv.2019.136258>, 2020.
- Xing, J., Hong, Q., Yang, B., Xing, C., Yang, S., Mao, M., Su, W., Chen, Y., and Zhang, C.: Source identification of volatile organic compounds precursors from glyoxal and secondary formaldehyde utilizing MAX-DOAS observations in guangzhou, China, *Atmospheric Research*, 325, 108 256, <https://doi.org/10.1016/j.atmosres.2025.108256>, 2025.
- Zarzana, K. J., Min, K.-E., Washenfelder, R. A., Kaiser, J., Krawiec-Thayer, M., Peischl, J., Neuman, J. A., Nowak, J. B., Wagner, N. L., Dubè, W. P., St. Clair, J. M., Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Ryerson, T. B., and Brown, S. S.: Emissions of Glyoxal and Other Carbonyl Compounds from Agricultural Biomass Burning Plumes Sampled by Aircraft, *Environmental Science & Technology*, 51, 11 761–11 770, <https://doi.org/10.1021/acs.est.7b03517>, 2017.
- Zarzana, K. J., Selimovic, V., Koss, A. R., Sekimoto, K., Coggon, M. M., Yuan, B., Dubé, W. P., Yokelson, R. J., Warneke, C., de Gouw, J. A., Roberts, J. M., and Brown, S. S.: Primary emissions of glyoxal and methylglyoxal from laboratory measurements of open biomass burning, *Atmospheric Chemistry and Physics*, 18, 15 451–15 470, <https://doi.org/10.5194/acp-18-15451-2018>, 2018.
- Zhang, B., Qiao, L., Han, H., Xie, W., and Li, L.: Variations in VOCs Emissions and Their O₃ and SOA Formation Potential among Different Ages of Plant Foliage, *Toxics*, 11, 645, <https://doi.org/10.3390/toxics11080645>, 2023.
- Zheng, G., Kuang, C., Uin, J., Watson, T., and Wang, J.: Large contribution of organics to condensational growth and formation of cloud condensation nuclei (CCN) in the remote marine boundary layer, *Atmospheric Chemistry and Physics*, 20, 12 515–12 525, <https://doi.org/10.5194/acp-20-12515-2020>, 2020.



Appendix

Table A1. Fit settings for NO₂ vis for Orléans, Athens, and Incheon.

Fit parameter		Selection/source
Spectral range		425 nm – 490 nm
Polynomial degree		6
Wavelength calibration		Solar atlas from Chance and Kurucz (2010)
Reference		Closest two zenith-sky spectra in time interpolated to time of measurement
Cross-section	Temperature	Data source
O ₃	223 K	Serdyuchenko et al. (2014) with I_0 correction
NO ₂	294 K, 220 K	Vandaele et al. (1996) with I_0 correction
O ₄	293 K	Thalman and Volkamer (2013)
H ₂ O	296 K	Rothman et al. (2010)
Ring	-	Computed by QDOAS, Danckaert et al. (2017)

Table A2. Fit settings for NO₂ UV for Orléans, Athens, and Incheon.

Fit parameter		Selection/source
Spectral range		338 nm – 370 nm
Polynomial degree		6
Wavelength calibration		Solar atlas from Chance and Kurucz (2010)
Reference		Closest two zenith-sky spectra in time interpolated to time of measurement
Cross-section	Temperature	Data source
O ₃	223 K, 243 K	Serdyuchenko et al. (2014) with I_0 correction
NO ₂	294 K, 220 K	Vandaele et al. (1996) with I_0 correction
BrO	223 K	Fleischmann et al. (2004)
HCHO	298 K	Meller and Moortgat (2000)
O ₄	293 K	Finkenzeller and Volkamer (2022)
Ring	-	Computed by QDOAS with and without Pukite Terms, Danckaert et al. (2017)



Table A3. Fit settings HCHO for Orléans, Athens, and Incheon.

Fit parameter		Selection/source
Spectral range		337.5 nm – 361 nm
Polynomial degree		6
Wavelength calibration		Solar atlas from Kurucz et al. (1984)
Reference		Closest zenith-sky spectrum in time
Cross-section	Temperature	Data source
HCHO	298 K	Meller and Moortgat (2000)
O ₃	223 K	Serdyuchenko et al. (2014)
NO ₂	294 K	Vandaele et al. (1996)
O ₄	293 K	Thalman and Volkamer (2013)
Ring	-	Computed by QDOAS, Danckaert et al. (2017)



Table A4. Fit settings CHOCHO for Orléans, Athens, and Incheon.

Fit parameter		Selection/source
Spectral range		436 nm – 470 nm
Polynomial degree		6
Wavelength calibration		Solar atlas from Kurucz et al. (1984)
Reference		Closest two zenith-sky spectra in time interpolated to time of measurement
Cross-section	Temperature	Data source
CHOCHO	296 K	Volkamer et al. (2005)
O ₃	223 K	Serdyuchenko et al. (2014)
NO ₂	294 K, 220 K	Vandaele et al. (1996)
O ₄	293 K	Thalman and Volkamer (2013)
H ₂ O	293 K	Rothman et al. (2013)
Ring	-	Computed by QDOAS, Danckaert et al. (2017)



Table A5. Station specific thresholds used for data filtering

	ATTO	Orléans
RMS	< 0.001	< 0.001
Intensity (VIS)	> 5e4 counts s ⁻¹	> 5e4 counts s ⁻¹
Intensity (UV)	> 1e4 counts s ⁻¹	> 1e4 counts s ⁻¹
SZA	< 80°	< 80°
rel. Error	< 50%	< 50%
	Incheon	Athens
RMS	< 0.001	< 0.001
Intensity (VIS)	> 5e4 counts s ⁻¹	-
Intensity (UV)	> 3e3 counts s ⁻¹	-
SZA	< 80°	< 80°
rel. Error	< 50%	< 50%



Table A6. Station specific variables. The angle of the azimuth viewing direction follows N=0° and E=90°.

	ATTO	Orléans
period	12.10.2017 –31.07.2022	03.07.2023 –01.07.2024
lowest viewing elevation	2° ¹	1°
azimuth viewing direction	50°	28°
	Incheon	Athens
period	06.10.2021 –15.11.2022	01.01.2021 –31.12.2023
lowest viewing elevation	3° ²	1°
azimuth viewing direction	137.5°	232.5°

¹ The highest O₄ dSCDs occurred for 2° elevation

² Lower elevations are obstructed



Table A7. O₄ vis scaling factors (f) applied to ATTO data per elevation. Scaling factors for ATTO were determined by matching the slope of ATTO to the total data points per elevation without ATTO

Elevation	f	Figure
2°	0.502	Fig. 1
4°	0.533	Fig. A7
8°	0.604	Fig. A8
15°	0.685	Fig. A9
30°	0.623	Fig. A10

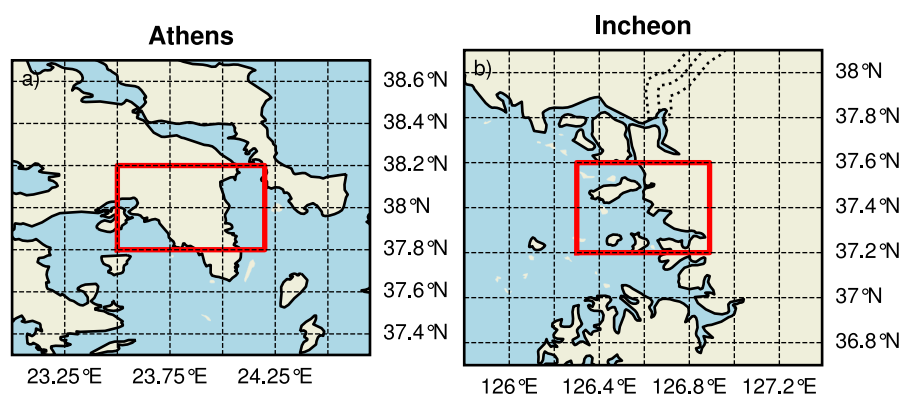


Figure A1. Maps of Athens and Incheon showing the considered area (red square) for computing the CAMS anthropogenic emissions for the relative contributions of the different sectors, see Fig. 17. Maps data: Natural Earth

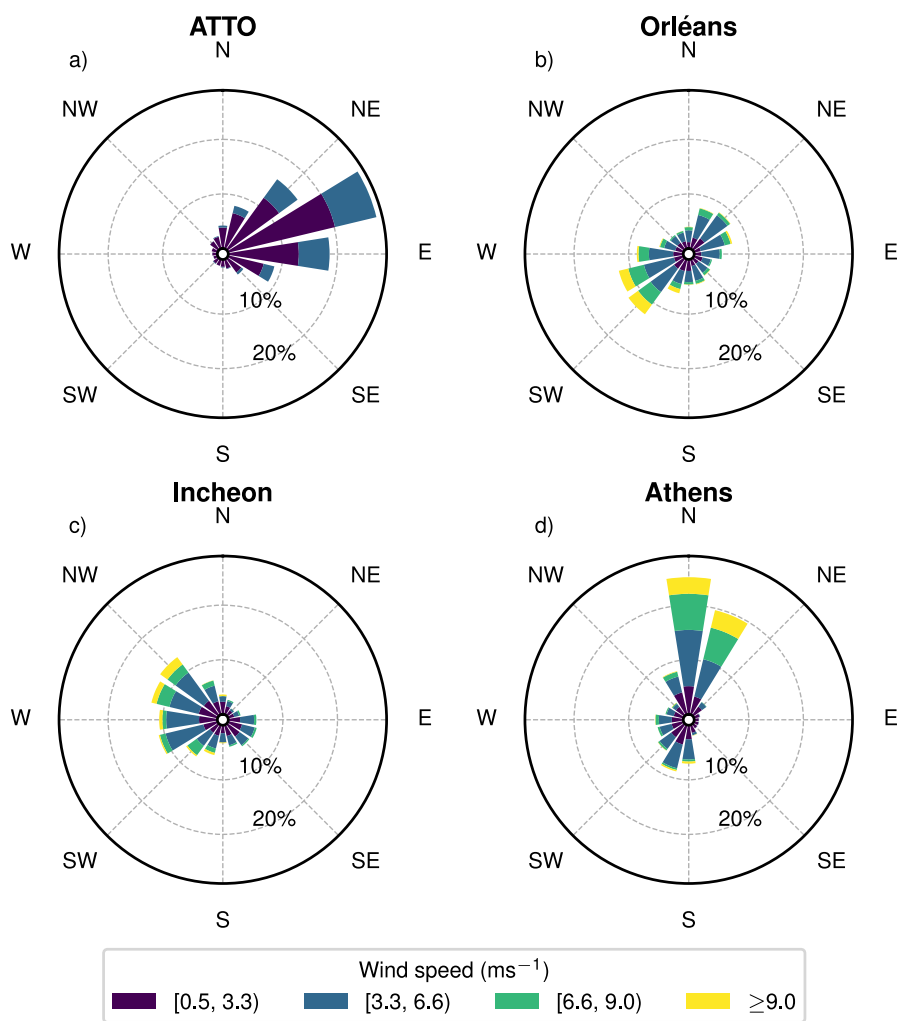


Figure A2. Wind roses showing the prevailing wind directions and wind speed in percentage during the respective measurement period for ATTO a), Orléans b), Incheon c) and Athens d). The bars show the percentage of wind direction from that direction and the colour indicates the wind speed.

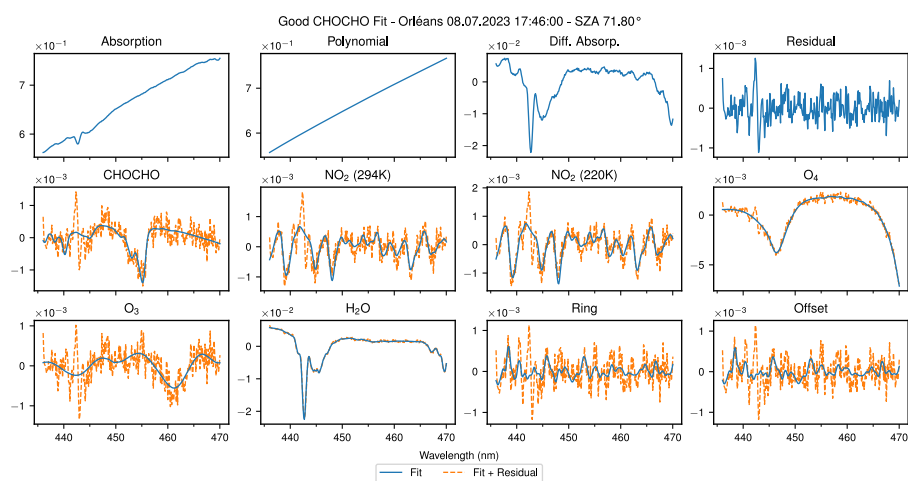


Figure A3. Fit components of a good CHOCHO fit in Orléans. The blue line indicates the fit and the dashed orange line indicates the residual added to the fit.

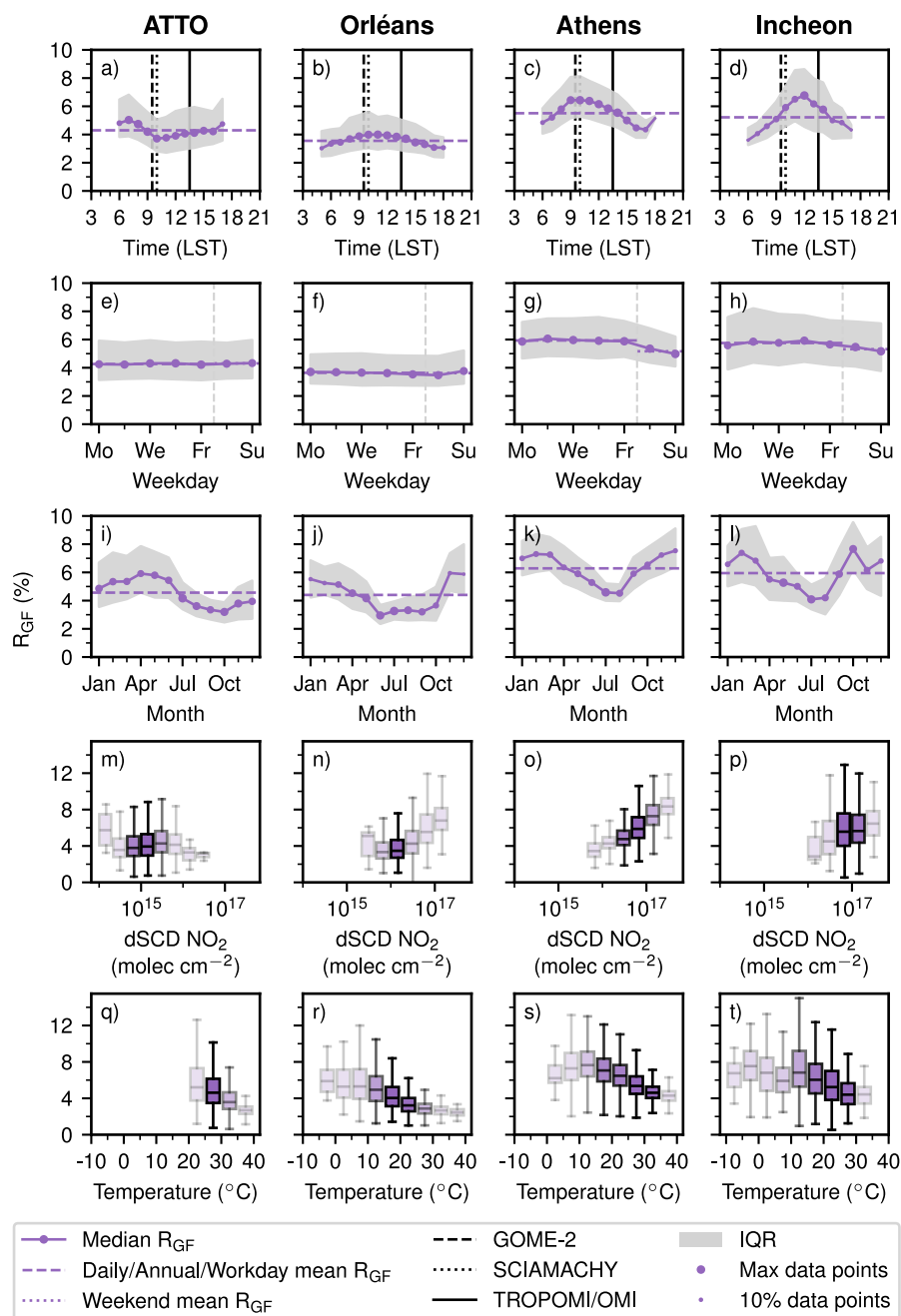


Figure A4. Overview of all figures for R_{GF} (without O_4 correction).

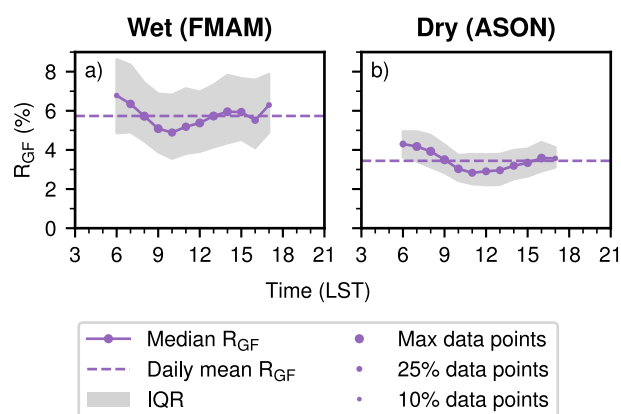


Figure A5. Diurnal cycles in the wet a) and dry b) season of R_{GF} (without O_4 correction) at ATTO. To indicate the number of measurements that contribute to each data point, the size of the marker is varied. The solid line corresponds to the median value, the dashed line to the daily mean, and the gray shaded area to the IQR. Only data from the lowest viewing elevation is included.

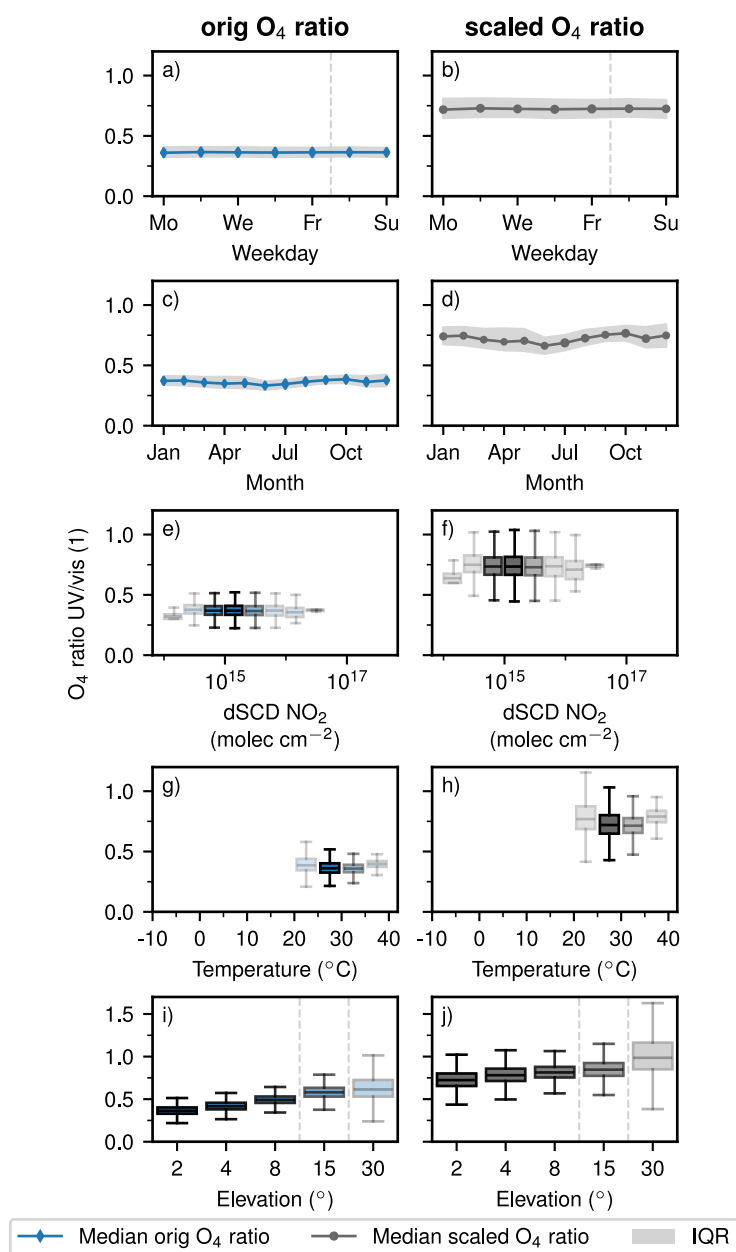


Figure A6. Overview of the behaviour of the O_4 ratio with and without scaling the O_4 vis dSCDs.

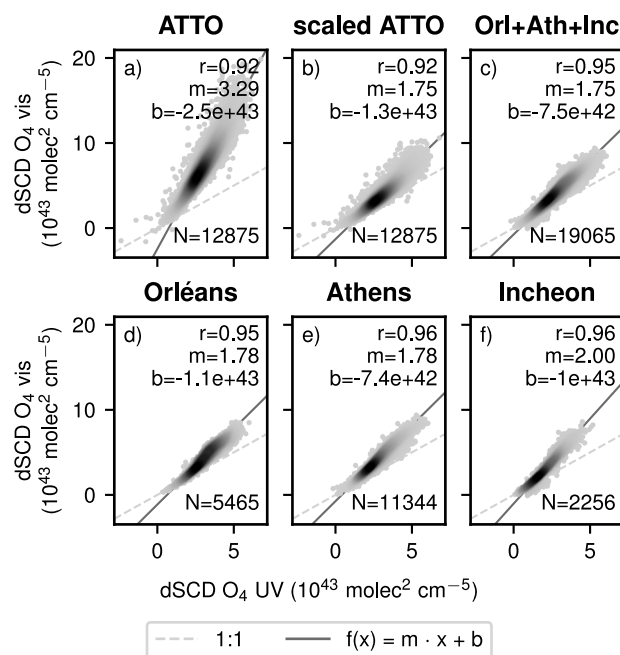


Figure A7. Scatter plots showing the correlation of O₄ dSCDs from the UV wavelength range with the vis wavelength range for different datasets at 4° viewing elevation. The light grey dashed line indicates the 1-to-1 line and the gray solid line indicates a orthogonal linear fit with the specified parameters. The density of the data points is indicated by the hue, denser regions are shown in dark grey.

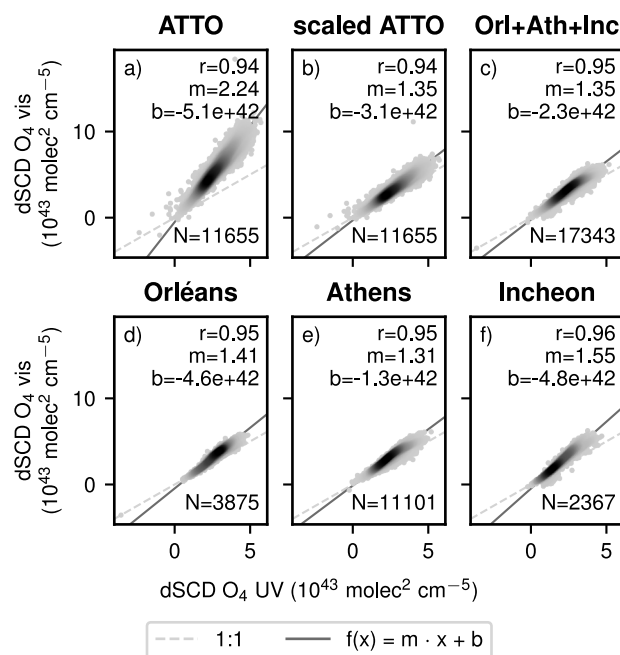


Figure A8. Scatter plots showing the correlation of O₄ dSCDs from the UV wavelength range with the vis wavelength range for different datasets at 8° viewing elevation. The light grey dashed line indicates the 1-to-1 line and the gray solid line indicates a orthogonal linear fit with the specified parameters. The density of the data points is indicated by the hue, denser regions are shown in dark grey.

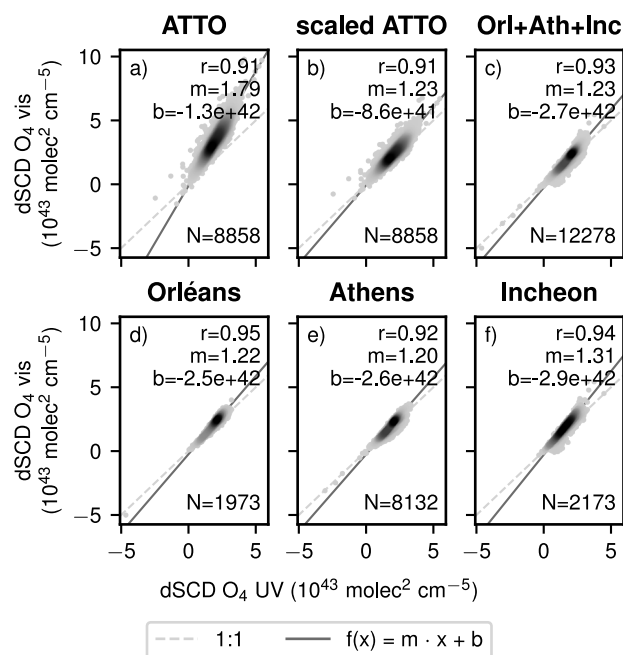


Figure A9. Scatter plots showing the correlation of O_4 dSCDs from the UV wavelength range with the vis wavelength range for different datasets at 15° viewing elevation. The light grey dashed line indicates the 1-to-1 line and the gray solid line indicates a orthogonal linear fit with the specified parameters. The density of the data points is indicated by the hue, denser regions are shown in dark grey.

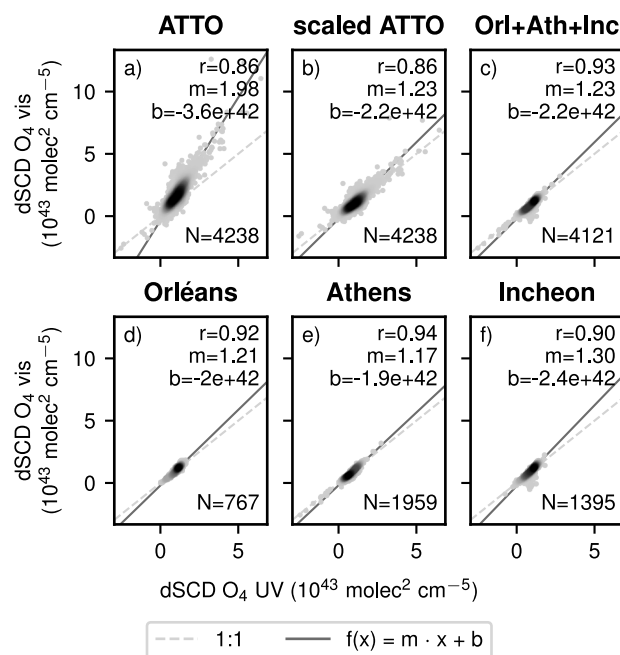


Figure A10. Scatter plots showing the correlation of O₄ dSCDs from the UV wavelength range with the vis wavelength range for different datasets at 30° viewing elevation. The light grey dashed line indicates the 1-to-1 line and the gray solid line indicates a orthogonal linear fit with the specified parameters. The density of the data points is indicated by the hue, denser regions are shown in dark grey.