



# 1 Glyoxal tropospheric column retrievals from TROPOMI, 2 multi-satellite intercomparison and ground-based validation

3  
4 Christophe Lerot<sup>1</sup>, François Hendrick<sup>1</sup>, Michel Van Roozendael<sup>1</sup>, Leonardo M.A. Alvarado<sup>2,3</sup>,  
5 Andreas Richter<sup>3</sup>, Isabelle De Smedt<sup>1</sup>, Nicolas Theys<sup>1</sup>, Jonas Vlietinck<sup>1</sup>, Huan Yu<sup>1</sup>, Jeroen Van  
6 Gent<sup>1</sup>, Trissevgeni Stavrakou<sup>1</sup>, Jean-François Müller<sup>1</sup>, Pieter Valks<sup>4</sup>, Diego Loyola<sup>4</sup>, Hitoshi  
7 Irie<sup>5</sup>, Vinod Kumar<sup>6</sup>, Thomas Wagner<sup>6</sup>, Stefan F. Schreier<sup>7</sup>, Vinayak Sinha<sup>8</sup>, Ting Wang<sup>9</sup>,  
8 Pucai Wang<sup>9</sup>, Christian Retscher<sup>10</sup>

9  
10 <sup>1</sup> Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

11 <sup>2</sup> Alfred Wegner Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany

12 <sup>3</sup> University of Bremen, Institute of Environmental Physics, Bremen, Germany

13 <sup>4</sup> Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft und Raumfahrt (DLR),  
14 Oberpfaffenhofen, Germany

15 <sup>5</sup> Center for Environmental Remote Sensing, Chiba University, Japan

16 <sup>6</sup> Max Planck Institute for Chemistry (MPIC), Mainz, Germany

17 <sup>7</sup> Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences, Vienna,  
18 Austria

19 <sup>8</sup> Indian Institute of Science Education and Research Mohali, Mohali, India

20 <sup>9</sup> Institute of Atmospheric Physics, Chinese Academy of Sciences (CAS), Beijing, China

21 <sup>10</sup> European Space Agency, ESRI, Frascati, Italy

22  
23 *Correspondence to:* Christophe Lerot, [Christophe.Lerot@aeronomie.be](mailto:Christophe.Lerot@aeronomie.be)

24 **Abstract.** We present the first global glyoxal (CHOCHO) tropospheric column product derived from the  
25 TROPOspheric Monitoring Instrument (TROPOMI) on board of the Sentinel-5 Precursor satellite. Atmospheric  
26 glyoxal results from the oxidation of other non-methane volatile organic compounds (NMVOCs) and from direct  
27 emissions caused by combustion processes. Therefore, this product is a useful indicator of VOC emissions. It is  
28 generated with an improved version of the BIRA-IASB scientific retrieval algorithm relying on the Differential  
29 Optical Absorption Spectroscopy (DOAS) approach. Among the algorithmic updates, the DOAS fit now includes  
30 corrections to mitigate the impact of spectral misfits caused by scene brightness inhomogeneity and strong NO<sub>2</sub>  
31 absorption. The product comes along with a full error characterization, which allows providing random and  
32 systematic error estimates for every observation. Systematic errors are typically in the range of 1-3x10<sup>14</sup>  
33 molec/cm<sup>2</sup> (~30-70% in emission regimes). Random errors are larger (>6x10<sup>14</sup> molec/cm<sup>2</sup>) but can be reduced by  
34 averaging observations in space and/or time. Benefiting from a high signal-to-noise ratio and a large number of  
35 small-size observations, TROPOMI provides glyoxal tropospheric column fields with an unprecedented level of  
36 details.

37 Using the same retrieval algorithmic baseline, glyoxal column data sets are also generated from the Ozone  
38 Monitoring Instrument (OMI) on Aura and from the Global Ozone Monitoring Experiment-2 (GOME-2) on board  
39 of Metop-A and Metop-B. Those four data sets are intercompared over large-scale regions worldwide and show  
40 a high level of consistency. The satellite glyoxal columns are also compared to glyoxal columns retrieved from  
41 ground-based Multi-Axis (MAX-) DOAS instruments at nine stations in Asia and Europe. In general, the satellite  
42 and MAX-DOAS instruments provide consistent glyoxal columns both in terms of absolute values and variability.  
43 Correlation coefficients between TROPOMI and MAX-DOAS glyoxal columns range between 0.61 and 0.87.



44 The correlation is only poorer at one mid-latitude station, where satellite data appears low biased during  
45 wintertime. The mean absolute glyoxal columns from satellite and MAX-DOAS generally agree well for  
46 low/moderate columns with differences less than  $1 \times 10^{14}$  molec/cm<sup>2</sup>. A larger bias is identified at two sites where  
47 the MAX-DOAS columns are very large. Despite this systematic bias, the consistency of the satellite and MAX-  
48 DOAS glyoxal seasonal variability is excellent.

## 49 1. Introduction

50 Exposure to poor air quality kills millions of people annually (e.g. Vohra et al., 2021; World Health Organization,  
51 2016) due to natural and human emissions of a large range of particulate matters and gases, including among  
52 others nitrous oxides (NO<sub>x</sub>), sulphur dioxide, carbon monoxide, methane and volatile organic compounds  
53 (VOCs). The latter, in combination with NO<sub>x</sub>, play a significant role in the secondary production of tropospheric  
54 ozone (Jacob, 2000), which is highly toxic for the respiratory system and also contributes to global warming  
55 because of its absorption in the thermal infrared. Global measurements of atmospheric concentrations of the ozone  
56 precursors is therefore crucial. The number of VOCs that can be found in the atmosphere is manifold, but only a  
57 few of them can be probed using remote sensing techniques. For example, formaldehyde (HCHO) measurements  
58 have been used in many studies as a proxy for probing emissions of non-methane VOCs of biogenic, pyrogenic  
59 and anthropogenic origin (e.g. Abbot et al., 2003; Barkley et al., 2013; Bauwens et al., 2016; Beekmann and  
60 Vautard, 2010; Curci et al., 2010; Jin et al., 2020; Marais et al., 2012; Palmer et al., 2006; Stavrakou et al., 2016;  
61 Wells et al., 2020).

62 With a lifetime of a few hours, glyoxal (CHOCHO) is another short-lived VOC that can be detected remotely,  
63 offering the potential to provide information on Non-Methane VOC (NMVOC) emissions. Over the past few  
64 years, an increasing number of studies (e.g. Cao et al., 2018; Chan Miller et al., 2017; Fu et al., 2008; Li et al.,  
65 2016; Liu et al., 2012; Stavrakou et al., 2009, 2016; Wittrock et al., 2006) have exploited glyoxal measurements  
66 from space, often in combination with formaldehyde. Being produced from similar sources, those two species are  
67 complementary as they have different production yields. For example, the oxidation of aromatics produces glyoxal  
68 with a much higher yield than formaldehyde (Cao et al., 2018). Although being both mostly produced via the  
69 oxidation of other VOCs, direct emissions from anthropogenic and fire activities also occur, and contribute more  
70 to the glyoxal global budget than to the formaldehyde one (Stavrakou et al., 2009b, 2009a). This motivated many  
71 studies to investigate the ratio of glyoxal to formaldehyde concentrations or columns as a possible metric to  
72 discriminate between different types of VOC emissions (e.g. Chan Miller et al., 2014; DiGangi et al., 2012; Hoque  
73 et al., 2018; Kaiser et al., 2015; Vrekoussis et al., 2010). Glyoxal measurements are also essential for establishing  
74 the global budget of secondary organic aerosols (SOAs). Indeed, with a high solubility in water, glyoxal undergoes  
75 heterogeneous uptake on aerosols and cloud droplets where the subsequent aqueous-phase chemistry forms SOA  
76 (Chan et al., 2010; Fu et al., 2008; Hallquist et al., 2009; Knote et al., 2014; Li et al., 2016; Volkamer et al., 2007).

77 Glyoxal has three absorption bands in the visible spectral range that have been exploited to remotely retrieve  
78 information on its atmospheric abundance using the Differential Optical Absorption Spectroscopy method  
79 (DOAS, Platt and Stutz, 2008) applied to ground-based (e.g. Benavent et al., 2019; Hoque et al., 2018; Javed et  
80 al., 2019; Schreier et al., 2020), air-borne (e.g. Kluge et al., 2020; Volkamer et al., 2015), ship-borne (e.g. Behrens  
81 et al., 2019; Sinreich et al., 2010) and space-based instruments. The first global glyoxal tropospheric column



82 observations from space have been realized by Wittrock et al. (2006) using nadir measurements from the  
83 SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CartographY) instrument. Based  
84 on this pioneering work, different glyoxal data products were derived from the Global Ozone Monitoring  
85 Experiment-2 (GOME-2) (Lerot et al., 2010; Vrekoussis et al., 2009) and from the Ozone Monitoring Instrument  
86 (OMI) (Alvarado et al., 2014; Chan Miller et al., 2014). All those different products rely on a similar DOAS  
87 approach, but generally differ from each other by the choice of the fit settings and of the auxiliary input data.

88 In general, the glyoxal optical depth is very low ( $< 5e-4$ ), typically one order of magnitude smaller than the  $\text{NO}_2$   
89 optical depth in the same spectral range. This results in retrievals prone-to-noise, requiring to average many of  
90 them to extract meaningful glyoxal signals. With an enhanced spatial resolution resulting in a number of  
91 observations more than ten times larger than provided by its predecessor OMI, the TROPospheric Monitoring  
92 Instrument (TROPOMI), operating since 2017, allows observing weak atmospheric absorbers with an  
93 unprecedented level of spatio-temporal details. This has been illustrated by Alvarado et al. (2020a) who  
94 investigated the large amounts of formaldehyde and glyoxal emitted by the intense North-American wildfires in  
95 August 2018 as observed by TROPOMI for several days and over long distances. Theys et al. (2020) have  
96 evaluated the respective contributions to the hydroxyl radical production in fresh fire plumes from nitrous acid,  
97 VOCs and other sources with the support of different TROPOMI data sets, including the glyoxal data product  
98 described here.

99 This work presents the latest version of the BIRA-IASB scientific glyoxal tropospheric column retrieval algorithm  
100 that has been applied to three years of TROPOMI measurements, and also to data from the predecessor nadir  
101 instruments OMI and GOME-2A/B. The quality of the TROPOMI glyoxal retrievals is investigated with (1) a  
102 global intercomparison of the satellite glyoxal data products generated with a common algorithm and (2)  
103 comparisons with independent glyoxal measurements from a series of Multi-AXis DOAS (MAX-DOAS)  
104 instruments located at nine stations in Asia and Europe.

105 After a brief introduction of the satellite instruments used in this study in Section 2, the retrieval algorithm and its  
106 different steps are described in Section 3, with emphasis on the updated and innovative aspects compared to  
107 heritage studies. This section also presents the typical random and systematic errors associated to the retrievals  
108 and how they are estimated for each individual measurement. Section 4 presents the evaluation of the inter-satellite  
109 consistency by comparing both seasonal global spatial patterns as seen from different instruments as well as  
110 monthly mean time series and seasonal cycles in a series of selected large-scale regions. Finally, Section 5 presents  
111 validation results based on MAX-DOAS data.

## 112 2. TROPOMI and other nadir-viewing satellite sensors

113 TROPOMI was launched on 13 October 2017 on board of the Sentinel-5 precursor platform. It flies on a sun-  
114 synchronous Low Earth Orbit (LEO) with an ascending node crossing the equator at the local time of 13:30. In  
115 the series of Sentinel missions from the European Union Copernicus programme, it is the first one dedicated to  
116 atmospheric composition. The instrument operates in a nadir viewing mode and measures Earthshine radiances  
117 and solar irradiances in the ultraviolet (UV), visible, near infrared and short infrared spectral bands. It aims at  
118 providing column amounts of a number of key pollutants, such as ozone ( $\text{O}_3$ ),  $\text{NO}_2$ ,  $\text{SO}_2$ , HCHO, CO,  $\text{CH}_4$  as well  
119 as cloud and aerosol parameters. TROPOMI offers a quasi-daily global coverage at the unprecedented spatial



120 resolution of  $3.5 \times 5.5 \text{ km}^2$  ( $3.5 \times 7 \text{ km}^2$  before August 2019) in the UV-visible spectral range. It is an imager-type  
121 instrument using a two-dimensional Charge Coupled Device (CCD) for the light measurements, the detector  
122 columns being used for the spectral resolution while the rows are binned to resolve spatially the 2600 km across-  
123 track swath into 450 individual ground pixels. The spectral resolution of the instrument is about 0.5 nm and offers  
124 a remarkably high signal-to-noise ratio of about 1500 in band 4 (405-500 nm) used in this study. More details on  
125 the instrument and its performance can be found in (Kleipool et al., 2018; Ludewig et al., 2020; Schenkeveld et  
126 al., 2017; Veeffkind et al., 2012). The TROPOMI measurements allow to derive the vertical columns of multiple  
127 species, some of them not included among the operational products listed above. Glyoxal is one of them and the  
128 details on how its column quantities are retrieved will be described in the next section.

129 The TROPOMI design strongly inherits from past nadir-viewing sensors, and in particular from the Ozone  
130 Monitoring Instrument (OMI) that we use to evaluate the TROPOMI glyoxal product presented in this work. OMI  
131 (Levelt et al., 2006) is also an imager instrument and flies on an early afternoon orbit since October 2004. The  
132 OMI swath, divided into 60 across-track pixels with a size varying from  $13 \times 24 \text{ km}^2$  (at nadir) to  $13 \times 150 \text{ km}^2$  (at  
133 the edges), allowed a daily global coverage before being limited in 2008 by the so-called row anomaly. The latter  
134 consists in a modification of the signal recorded by OMI at specific rows, due to a mechanical obstruction of the  
135 field of view, and leads to lower quality spectral measurements (Torres et al., 2018). We also exploit spectral  
136 measurements from the Global Ozone Monitoring Experiment-2 (GOME-2) instruments aboard the Metop-A and  
137 Metop-B platforms. In contrast to OMI and TROPOMI, the GOME-2 instruments (Munro et al., 2016) fly on  
138 early morning LEOs with local equator crossing times around 09:30 and are scanning spectrometers, meaning that  
139 across-track pixels are successively sounded. The scan is divided into 24 pixels for a total swath of 1920 km,  
140 providing global coverage in 1.5 day. Each pixel has a size of  $80 \times 40 \text{ km}^2$ . After the launch of Metop-B, the GOME-  
141 2A swath was reduced to 960 km in July 2013, leading to ground pixel two times smaller.

### 142 3. Description of the Algorithm

143 The algorithm for retrieving tropospheric vertical columns of glyoxal relies on a classical DOAS approach (Platt  
144 and Stutz, 2008). This approach consists first in fitting measured optical depths in an optimized spectral window  
145 to derive the so-called slant column densities *SCDs* (atmospheric concentration integrated along the effective light  
146 path) of the absorbers. The latter are thereafter converted into vertical column densities *VCDs* (concentration  
147 vertically integrated from the satellite ground pixel up to the top of the atmosphere) with air mass factors obtained  
148 by modelling the radiative transfer through the atmosphere. An additional background correction procedure is  
149 often applied for weak absorbers such as glyoxal in order to reduce as much as possible the presence of systematic  
150 biases caused by spectral interferences.

151 The glyoxal algorithm presented here largely inherits from past developments for predecessor nadir-viewing  
152 satellite sensors (Alvarado et al., 2014, 2020; Chan Miller et al., 2014; Lerot et al., 2010; Vrekoussis et al., 2009;  
153 Wittrock et al., 2006). In the following subsections, we further describe each algorithmic component, with  
154 emphasis on its specificities. The retrievals are provided with estimates for the random and systematic errors,  
155 which are discussed in subsection 3.4.



156        **3.1. DOAS fit**

157        To exploit the glyoxal absorption bands, we use a fitting window from 435 to 460 nm encompassing the two most  
158        intense bands, which has shown in the past to provide reliable results (Barkley et al., 2017; Lerot et al., 2010).  
159        This has been confirmed by sensitivity tests carried out by Alvarado et al. (2014) and Chan Miller et al. (2014).  
160        Owing to its low optical depth ( $<5 \times 10^{-4}$ ), any poorly fitted feature in the radiance measurements may affect the  
161        retrieved glyoxal SCD. It is therefore crucial to account for any physical or instrumental effect in order to optimise  
162        the fit quality as much as possible. Different aspects of the algorithm contribute to achieve this.

163        The wavelength grids of the measured spectra are recalibrated before the actual DOAS fits with a cross-correlation  
164        procedure (Danckaert et al., 2017; De Smedt et al., 2018) during which the position of the lines in the measured  
165        irradiance spectrum is fitted to an external solar atlas (Chance and Kurucz, 2010), convolved to the satellite  
166        spectral resolution. This recalibration procedure is done once per orbit and separately for every detector row of  
167        the instrument.

168        Although the DOAS fit generally uses an irradiance as the reference spectrum, it is common practice, in the case  
169        of weak tropospheric absorbers, to replace it by a mean radiance spectrum recorded in a remote region where the  
170        concentration of the gas of interest is low (e.g. De Smedt et al., 2018). This allows reducing the presence of  
171        systematic biases caused by spectral interferences and/or instrumental limitations. In particular, the use of one  
172        separate mean radiance spectrum per detector row minimizes the presence of so-called stripes in the product  
173        typical of imager-type instruments such as OMI or TROPOMI. Here we compute those mean radiance spectra on  
174        a daily basis by averaging for each row all spectra located within the equatorial Pacific Ocean (15°S-15°N;  
175        120°W-180°W).

176        The selected settings for the DOAS fits rely on the aforementioned past studies and are summarized in Table 1.  
177        The latest available cross-sections for species absorbing in the selected fitting window are included in the fit, i.e.  
178        O<sub>3</sub>, NO<sub>2</sub>, O<sub>2</sub>-O<sub>2</sub>, water vapour and liquid water in addition to glyoxal. Note that the water vapour cross-section is  
179        based on the HITRAN2012 database (Rothman et al., 2013) as we found that the latest HITRAN2016 version  
180        (Gordon et al., 2017) led to poorer fit quality. The temperature dependence of the NO<sub>2</sub> absorption is taken into  
181        account by including a second cross-section, taken as the difference between NO<sub>2</sub> cross-sections reported at 2  
182        temperatures (220 and 294K) as proposed by Alvarado et al. (2014) and Chan Miller et al. (2014) for their  
183        respective OMI glyoxal products. Consistently with Alvarado et al. (2014), we found that fitting the liquid water  
184        optical depth in the glyoxal fitting window performs as well as fixing it to a value previously determined in a  
185        larger spectral interval as proposed in the past (Lerot et al., 2010). A number of additional cross-sections are  
186        included in the fit to consider (1) Inelastic scattering (Ring effect) introduces high-frequency structures that are  
187        treated as a pseudo-absorber (Chance and Spurr, 1997); (2) Intensity offsets in the spectra, caused for example by  
188        residual straylight, are corrected for by fitting the inverse of the reference spectrum (Danckaert et al., 2017); (3)  
189        heterogeneity of the scene brightness may also introduce high frequency structures, which are considered with  
190        pseudo-cross-sections (more details hereafter). All those cross-sections are generated at the instrumental spectral  
191        resolution by using the key data Instrumental Spectral Response Functions provided for all individual detector  
192        rows. During the DOAS procedure, the earthshine radiance spectrum is further aligned with the reference, by  
193        allowing it to be shifted and stretched in wavelength. In addition, the DOAS fit procedure includes a spike removal



194 scheme as described in Richter et al. (2011) enabling to filter out from the fit individual corrupted radiance  
 195 measurements, and hence to reduce the noise in the product.

196 **Table 1 : Absorption cross-sections and settings used for the retrieval of glyoxal slant columns**

Fitting interval	435-460 nm
<b>Absorption cross-sections</b>	
Glyoxal	Volkamer et al. (2005)
Ozone	Serdyuchenko et al. (2014), 223K
NO <sub>2</sub>	Vandaele et al. (1998), 220K and 294K, I <sub>0</sub> effect-corrected (Aliwell et al., 2002)
O <sub>4</sub> (O <sub>2</sub> -O <sub>2</sub> )	Thalman and Volkamer (2013), 293K
H <sub>2</sub> O (vapour)	Rothman et al. (2013), 293K
H <sub>2</sub> O (liquid)	Mason et al. (2016)
Scene Heterogeneity	2 pseudo-absorbers (Richter, 2018) – Internally generated
Ring effect	Pseudo-absorber (Chance and Spurr, 1997; Wagner et al., 2009)
<b>Other parameters</b>	
Polynomial	3rd order
Intensity offset correction	1 <sup>st</sup> -order offset (additional cross-section taken as the inverse of the reference spectrum)
Earthshine wavelength shift	1 <sup>st</sup> -order shift
<b>Reference spectrum (E<sub>0</sub>)</b>	Daily average of radiances, per detector row, selected in equatorial Pacific (Lat: [-15° 15°], Long: [180°-240°])

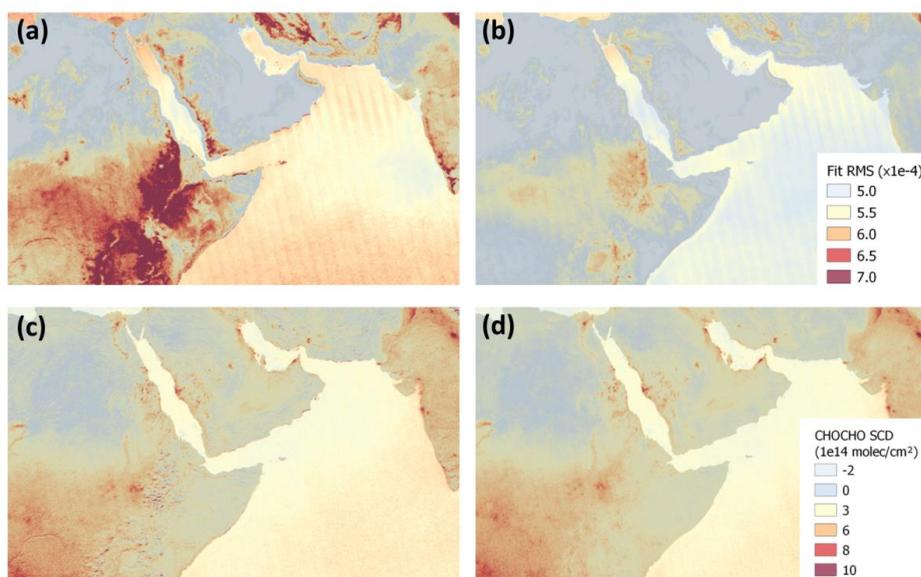
197

198 **3.1.1. Scene heterogeneity**

199 Any intensity variation within the probed scene taking place perpendicularly to the instrumental slit (i.e. along  
 200 track) leads to perturbations of the instrumental spectral slit function (ISRF) (Noël et al., 2012; Voors et al., 2006).  
 201 Richter et al. (2018) have shown that those perturbations lead to a degradation of the NO<sub>2</sub> DOAS spectral fit quality  
 202 and to systematic biases on the retrieved slant columns. Such abrupt intensity changes occur for example along  
 203 the coasts, mountains or cloud edges. Glyoxal retrievals are also affected by such scene heterogeneity as illustrated  
 204 in Figure 1 over the Horn of Africa and Middle East. This figure shows in the panel (a) that the root mean square  
 205 (RMS) of the DOAS fit residuals is systematically higher along the coasts but also over land where contamination  
 206 by broken clouds or abrupt elevation changes cause discontinuities in brightness fields. The stripes visible in this  
 207 figure are due to the smaller pixel size (and hence lower signal-to-noise ratio) on the edges of the across-track  
 208 field of view. The panel (c) shows that there are some collocated artificial patterns (positive/negative biases) in  
 209 the mean retrieved glyoxal slant column field. The latter result from spectral interferences with the signature  
 210 introduced by the ISRF distortion. Richter et al. (2018) showed that those spectral interferences can be  
 211 significantly reduced with additional cross-sections in the DOAS fit scaling the possible scene heterogeneity  
 212 signature. Those cross-sections are generated with a statistical analysis of the fit residuals for many observations



213 in a remote region as a function of the level of scene heterogeneity. The latter can be computed using radiance  
214 measurements at higher spatial resolution available in the TROPOMI level-1 data at a limited number of  
215 wavelengths. Following this approach, two additional cross-sections have been added to the DOAS baseline and  
216 both the fit residuals and the identified glyoxal biases have been reduced as illustrated in the right panels (b) and  
217 (d) of Figure 1. This effect is particularly visible along coasts and mountains but also over lands where some  
218 pseudo-noise caused by persistent broken clouds is also largely reduced. Note that a third cross-section derived  
219 from the mean residuals of homogeneous scenes is also added, which explains why the fit RMS are also reduced  
220 (but less drastically) in homogeneous scenes. This cross-section has no impact on the retrieved glyoxal SCDs and  
221 allowed mostly isolating systematic residuals due to scene heterogeneity only for the pseudo cross-sections  
222 creation.



223

224 **Figure 1 : Impact of scene brightness heterogeneity on glyoxal retrievals in the fitting window 435-460 nm over the**  
225 **Horn of Africa and Middle-East. The panels (a) and (b) show mean fit residuals RMS for the year 2019 without and**  
226 **with (left and right) pseudo-cross sections to correct for spectral signatures introduced by scene heterogeneity. The**  
227 **panels (c) and (d) show the corresponding mean glyoxal slant column densities. Only observations with cloud**  
228 **fractions less than 20% are considered.**

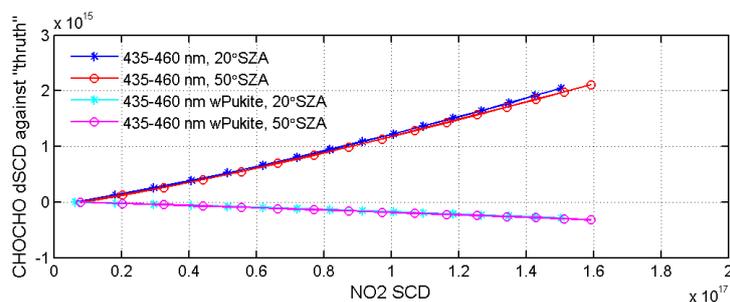
### 229 3.1.2. Empirical correction for strong NO<sub>2</sub> absorption

230 The DOAS approach assumes that the wavelength dependence of the effective light path within the fit interval  
231 can be neglected. Although this assumption is generally reasonable, it may fail in case of strong absorption by one  
232 (or more) species, of which the slant column density becomes dependent on the wavelength (Puķīte et al., 2010).  
233 In that case, fitting the optical depth of that species by a simple scaling of its cross-section is inaccurate and the  
234 fit quality is degraded. Puķīte et al. (2010) have shown that fitting additional cross-sections resulting from a Taylor



235 expansion of the wavelength-dependent slant column corrects for its variability within the fit window. As  
236 mentioned before, the high sensitivity of glyoxal retrievals to potential sources of misfit was a motivation to  
237 further investigate its sensitivity to extreme  $\text{NO}_2$  concentration levels.

238 For this purpose, synthetic spectra were generated at a spectral resolution of 0.5 nm with the radiative transfer  
239 model SCIATRAN (Rozanov et al., 2005) for a satellite nadir-viewing geometry and two different solar zenith  
240 angles. In those simulations, inelastic scattering was neglected and a large range of tropospheric  $\text{NO}_2$  columns  
241 was covered by scaling the  $\text{NO}_2$  a priori profile. The TROPOMI DOAS baseline described above was then applied  
242 to those simulated spectra in order to retrieve CHOCHO SCDs and evaluate the error as a function of the  $\text{NO}_2$   
243 SCD as illustrated in Figure 2. Results clearly point to a CHOCHO SCD error increasing with the  $\text{NO}_2$  SCD. Note  
244 that the exact error magnitude may change slightly depending on the  $\text{NO}_2$  vertical distribution and on the actual  
245 atmospheric content. On the other hand, adding the so-called Pukite cross-sections (Puķīte et al., 2010) to account  
246 for the wavelength-dependence of the  $\text{NO}_2$  SCD significantly reduces the errors.



247

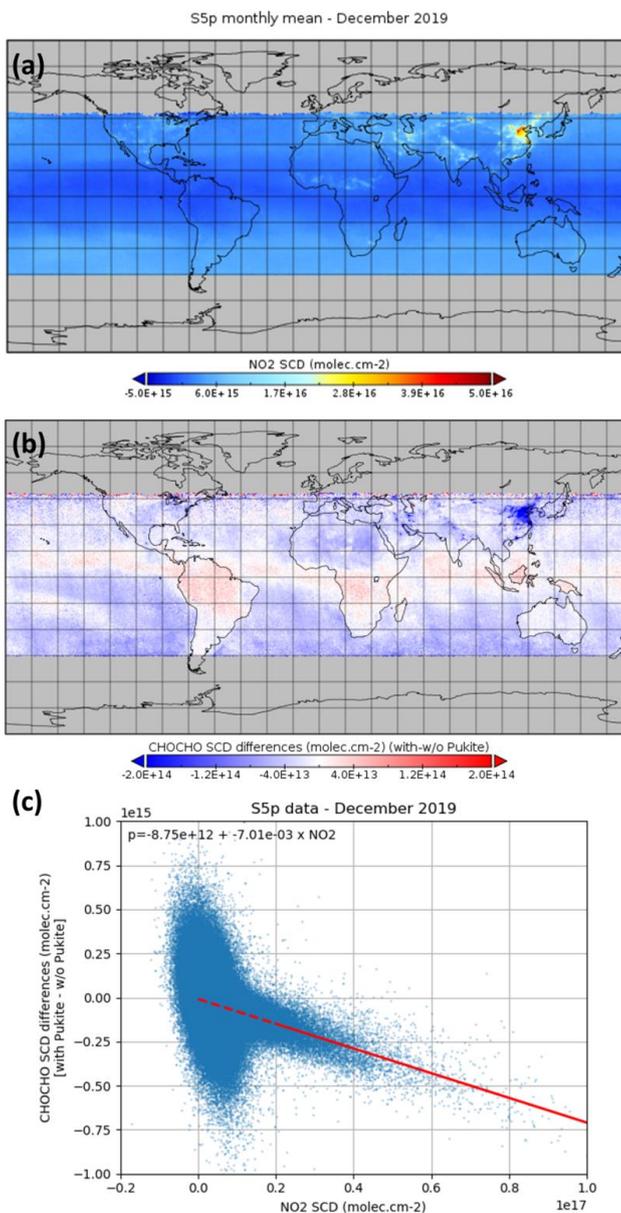
248 **Figure 2 : Absolute error (in molec/cm<sup>2</sup>) on the retrieved CHOCHO SCD as a function of the  $\text{NO}_2$  SCD for simulated**  
249 **spectra in a nadir-viewing satellite geometry and for two solar zenith angles. The reference “true” CHOCHO SCD is**  
250 **taken as the value retrieved for the lowest  $\text{NO}_2$  SCD scenario. The error increases with the  $\text{NO}_2$  SCD when Pukite**  
251 **cross-sections are not included in the fit, but remains small otherwise.**

252 On this basis, the impact of adding the Pukite cross-sections to the DOAS baseline has been investigated using  
253 one month of TROPOMI data. A wintertime period was chosen (December 2019) to favour the number of  
254 observations with large  $\text{NO}_2$  concentrations, in particular in China but also in other megacities in the Northern  
255 Hemisphere. Figure 3 (upper panel) displays the monthly mean  $\text{NO}_2$  SCDs in December 2019, and (middle panel)  
256 the mean impact on the retrieved CHOCHO SCDs of introducing the Pukite terms in the DOAS spectral fit  
257 baseline. The CHOCHO SCD differences caused by the Pukite terms are also plotted as a function of the  $\text{NO}_2$   
258 SCDs to better visualize the correlation (lower panel). For regions with enhanced  $\text{NO}_2$  concentrations ( $>2 \times 10^{16}$   
259 molec/cm<sup>2</sup>) (e.g. China, India, Teheran), the Pukite cross-sections lead to a systematic reduction of the CHOCHO  
260 SCDs, consistent with the closed-loop tests described above. A small improvement of the fit quality is found (not  
261 shown). Unexpectedly, the impact of those additional cross-sections on the CHOCHO SCDs can also be non-  
262 negligible in regions with low  $\text{NO}_2$  columns: positive differences are for example observed over equatorial oceans,  
263 but also over South America and Africa. The correlation plot of Figure 3 clearly shows these two regimes. While  
264 the impact of the Pukite cross-sections on the glyoxal retrievals is understood and reliable for large  $\text{NO}_2$  SCDs,



265 their influence at low  $\text{NO}_2$  SCD is more questionable and likely results from spectral interferences occurring  
266 between the different fitted spectra (e.g. with the Ring signature), which introduces additional noise in the product.

267 To avoid this, rather than fitting additional cross-sections, we introduce an empirical correction applied to the  
268 glyoxal SCDs. This correction consists in subtracting from the glyoxal SCD an  $\text{NO}_2$ -SCD dependent value,  
269 directly prescribed from the linear regression fit through the sensitivity test results for all observations worldwide  
270 from December 2019, with  $\text{NO}_2$  SCDs larger than  $2 \times 10^{16}$  molec/cm<sup>2</sup> as illustrated in Figure 3 (c). It is worth  
271 noting that the regression fit results agree well with the glyoxal SCD errors estimated from the simulations  
272 presented above (Figure 2). For extreme pollution conditions such as what can be found in China during  
273 Wintertime, this correction may lead to glyoxal column reduction up to 30%.



274

275 **Figure 3 : (a) Monthly mean NO<sub>2</sub> SCDs retrieved from TROPOMI data in December 2019. Panel (b) illustrates the**  
276 **CHOCHO SCD absolute differences (molec/cm<sup>2</sup>) due to the incorporation of the Pukite et al. (2010) cross-sections in**  
277 **the DOAS spectral fit and panel (c) shows the correlation between those differences and the NO<sub>2</sub> SCDs. The red line**  
278 **corresponds to a linear regression fit through all points with NO<sub>2</sub> SCD larger than 2x10<sup>16</sup> molec/cm<sup>2</sup>.**



279 **3.2. Air Mass Factor computation**

280 The computation of the air mass factor (AMF) used to convert the retrieved glyoxal slant columns (SCD) to  
281 vertical columns (VCD) relies on the formulation of Palmer et al. (2001), which decouples the radiative transfer  
282 through the atmosphere from the vertical distribution of the gas of interest. Radiative transfer simulations are  
283 performed with the vector model VLIDORT at the middle of the fitting window (448 nm) to compute so-called  
284 altitude-dependent air mass factors or box-AMFs representing the sensitivity of the slant column to a small  
285 concentration change at any altitude. The AMF is obtained as the weighted mean of those box-AMFs using as  
286 weights the vertical distribution of the glyoxal concentration.

287 Typically, the sensitivity of nadir-viewing UV-Visible instruments is reduced in the lowermost atmospheric layers  
288 because of Rayleigh scattering. However, this sensitivity depends strongly on the observation geometry, on the  
289 surface reflectivity and altitude and on the presence of clouds. For example, the sensitivity is generally further  
290 reduced for low sun elevation. For this reason, retrievals with solar zenith angles larger than 70° are filtered out.  
291 We use a pre-computed five-dimensional look-up table of Box-AMFs spanning all observation conditions (see  
292 Table 2) and from which appropriate values are linearly interpolated for every TROPOMI observation. This  
293 interpolation uses as input the observation angles provided in the level-1 data, surface elevation taken from the  
294 GMTED2010 topography (Danielson and Gesch, 2011) and surface albedo extracted from the OMI minimum  
295 Lambertian Equivalent Reflectivity climatology (Kleipool et al., 2008). The spatial resolution of the latter  
296 database (0.5°x0.5°) is coarse compared to the TROPOMI footprint and neglects anisotropy, which may introduce  
297 significant errors (Lorente et al., 2018). However, at the time of writing, it is the only database available at the  
298 S5p overpass time although new Lambertian Equivalent Reflectivity climatologies relying on past works (e.g.  
299 Loyola et al., 2020; Tilstra et al., 2021, 2017) are currently being prepared. On the other hand, the level of noise  
300 in glyoxal retrievals generally requires averaging in space and/or time which in turn will reduce part of those error  
301 sources. We also neglect the impact of clouds and aerosols on the radiative transfer. Instead we apply a stringent  
302 cloud filtering approach: only observations with an effective cloud fraction (as retrieved in the same spectral range  
303 and provided in the TROPOMI operational NO<sub>2</sub> product (van Geffen et al., 2019) lower than 20% are conserved.  
304 This approach is motivated by the fact that glyoxal slant columns tend to be biased high over bright scenes because  
305 of poorly understood residual spectral interferences (e.g. with the Ring signature). Similarly, scenes covered by  
306 snow and ice are also discarded.

307

308

**Table 2 : Granularity of the Box-AMF look-up table**

Parameter name	Grid of values
Solar zenith angle [deg]	0, 10, 20, 30, 40, 45, 50, 55, 60, 65, 70, 72, 74, 76, 78, 80, 85
Line of sight zenith angle [deg]	0, 10, 20, 30, 40, 50, 60, 65, 70, 75



Relative azimuth angle [deg]	0, 45, 90, 135, 180
Surface albedo	0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.25, 0.3 0.4, 0.6, 0.8, 1.0
Surface pressure [hPa]	1063.10, 1037.90, 1013.30, 989.28, 965.83, 920.58, 876.98, 834.99, 795.01, 701.21, 616.60, 540.48, 411.05, 308.00, 226.99, 165.79, 121.11

309

310 The a priori glyoxal vertical profile shapes necessary to perform the AMF computations are provided by the global  
311 Chemistry Transport Model MAGRITTE developed at BIRA-IASB, which inherits from the IMAGES model  
312 (Bauwens et al., 2016; Müller and Brasseur, 1995; Stavrakou et al., 2009b, 2013). This model runs at  $1^\circ \times 1^\circ$   
313 resolution and calculates the distribution of 182 chemical compounds, of which 141 species undergo transport.  
314 The modelled troposphere is vertically divided in 40 levels between the surface and the lower stratosphere and  
315 meteorological fields are provided by the ECMWF ERA-5 analyses. The chemical mechanism and deposition  
316 scheme have been recently updated (Müller et al., 2018, 2019). Anthropogenic NMVOC emissions of are  
317 provided by the EDGAR 4.3.2 inventory (Huang et al., 2017) for the year 2012. Biomass burning emissions are  
318 obtained from the Global Fire Emission Database version 4 (GFED4s) (Van Der Werf et al., 2017). The emissions  
319 of isoprene and monoterpenes are calculated using the MEGAN-MOHYCAN model (Guenther et al., 2012;  
320 Müller et al., 2008). The model also incorporates biogenic emissions of methanol, methyl-butanol, ethylene,  
321 ethanol, acetaldehyde, formaldehyde and acetone, as well as oceanic emissions of methanol, acetone, acetaldehyde  
322 and alkyl nitrates (Müller et al., 2019). The global source of glyoxal in the model amounts to 47 Tg/yr (in 2013),  
323 of which about 4 Tg/yr are due to direct biomass burning emissions, and 18, 6, 9 and 9 Tg/yr are due to the  
324 atmospheric degradation of isoprene, acetylene, aromatics and monoterpenes, respectively (Müller et al., 2019).

325 To account for the difference in spatial resolution between the model and the observations, a priori profiles are  
326 rescaled to the effective satellite pixel surface elevation using the formulation proposed by Zhou et al. (2009).  
327 Enhanced glyoxal concentrations have been detected over oceans in several studies (Coburn et al., 2014; Lerot et  
328 al., 2010; Sinreich et al., 2010), but current models cannot reproduce this. For this reason, over oceans, we use an  
329 a priori glyoxal concentration profile measured with an air-borne MAX-DOAS instrument over the Pacific Ocean  
330 during the TORERO campaign (Volkamer et al., 2015).

### 331 3.3. Background correction

332 As already mentioned, systematic (row-dependent) biases in the retrieved SCDs often remain due to small residual  
333 interferences with spectral signatures from other absorbers or due to instrumental effects. In the particular case of  
334 pushbroom imaging instruments such as OMI/TROPOMI, across-track row-dependent biases (so-called stripes)  
335 often occur due to the imperfect calibration of the different CCD detector rows. To reduce those biases, a  
336 background correction using observations in a remote reference sector is generally applied as part of the retrieval  
337 algorithm (e.g. Alvarado et al., 2014; Chan Miller et al., 2014; Lerot et al., 2010; Richter and Burrows, 2002; De  
338 Smedt et al., 2018). The principle of this background correction is to add offset values to the retrieved SCDs to



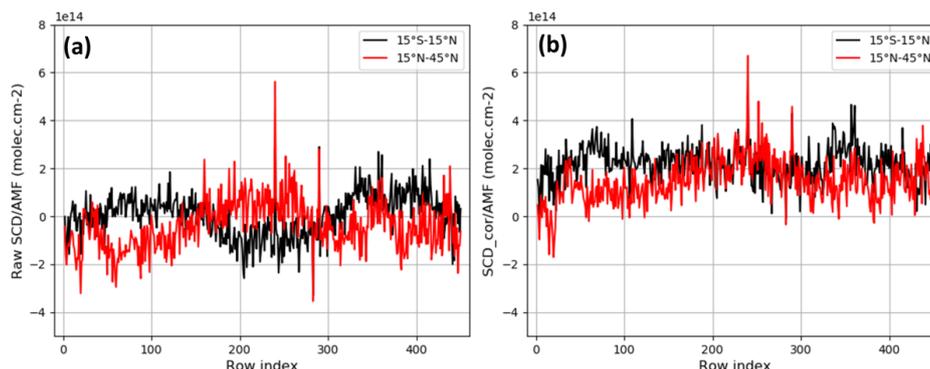
339 ensure that the resulting mean VCD in a clean remote region match an a priori known tropospheric glyoxal  
340 column. Here we use the Pacific Ocean as reference sector with a constant reference VCD of  $1 \times 10^{14}$  molec/cm<sup>2</sup>.  
341 This value was chosen according to independent measurements performed in this region (Sinreich et al., 2010)  
342 since current global models fail to reproduce remote sensing glyoxal levels observed over oceans (Fu et al., 2008;  
343 Myriokefalitakis et al., 2008; Stavrakou et al., 2009b).

344 The background correction is applied on a daily basis in different steps:

- 345 1. First, a destriping procedure such as proposed in Boersma et al. (2007) is applied consisting in an offset  
346 correction determined separately for each instrumental row, and relying on clear sky observations from  
347 the Equatorial Pacific Ocean (15°S-15°N, 165°E-220°E). The offset corrections are added to all glyoxal  
348 SCDs worldwide, considering their respective row.
- 349 2. Additionally to the high frequency stripes, a broadband row-dependent structure, of which the shape also  
350 depends on the latitude, was identified as illustrated in Figure 4, panel (a). This figure compares the row-  
351 dependence of mean uncorrected VCDs in the Pacific Ocean at Equatorial and Northern mid-latitudes.  
352 The two curves are somehow anti-correlated, meaning that the destriping correction based on equatorial  
353 latitudes only as applied in step 1 is not sufficient and even reinforces the mid-latitude structure. The  
354 second step of the background correction aims thus at reducing this broadband row-dependent structure  
355 at all latitudes while maintaining the mean latitudinal distribution of the measured background glyoxal  
356 columns. For this, Pacific Ocean measurements (40°S-40°N, 165°E-220°E) are binned per 20° in latitude  
357 and in groups of 15 rows in a 2-dimensional matrix. For this step, we use reference VCDs depending on  
358 the latitude and resulting from the averaging of the binned VCDs along the row dimension. A  
359 corresponding 2-dimensional matrix of SCD offset corrections is then computed in order, once applied  
360 to the binned VCDs, the corrected values match the reference VCDs. Interpolation through this correction  
361 matrix provides offsets to be applied to all SCDs retrieved worldwide.
- 362 3. Finally, the overall level of the product is adjusted with a single offset correction to ensure that the mean  
363 of all clear-sky VCDs within the full reference sector (40°S-40°N, 165°E-220°E) is equal to  $1 \times 10^{14}$   
364 molec/cm<sup>2</sup>. Panel (b) of Figure 4 shows how the identified row dependence in the VCDs at different  
365 latitudes has been reduced. The general level of the columns has also been adjusted.

366

367



368

369 **Figure 4 : Row-dependence of the glyoxal vertical columns of S5p orbit #5877 (December, 1st 2018) averaged in an**  
 370 **equatorial latitude band and in a Northern mid-latitude band. (a) No background correction is applied; (b) a latitude-**  
 371 **dependent background correction is applied.**

### 372 3.4. Uncertainty estimates

373 Glyoxal tropospheric column retrievals are affected by many sources of uncertainties in the different components  
 374 of the algorithm. The low glyoxal optical depth makes its retrieval highly sensitive to measurement noise and to  
 375 spectral interferences with strong absorption signatures of other species or with instrumental features. Although  
 376 the measurement noise can be reduced by averaging column retrievals from individual observations, spectral  
 377 interferences generally lead to residual systematic errors (biases), which cannot be easily eliminated. The  
 378 background correction described above aims at reducing those biases, but it has its own limitations. For example,  
 379 the reference glyoxal tropospheric column within the reference sector is poorly known. In addition to spectral fit  
 380 errors, there are also significant errors associated to the air mass factor calculations, mostly originating from input  
 381 parameters uncertainties. For estimating the total glyoxal column error, we assume that the different error  
 382 components are uncorrelated and can be summed quadratically as in (Boersma et al., 2011; Lerot et al., 2010; De  
 383 Smedt et al., 2008, 2018). If the glyoxal vertical column  $N_v$  is expressed as

$$N_v = \frac{N_s - \overline{(N_{s,0} - N_{v,0,ref} \times M_0)}}{M} \quad (1)$$

384 with  $N_s$  the retrieved slant column,  $M$  the AMF,  $\overline{(N_{s,0} - N_{v,0,ref} \times M_0)}$  the background correction term where  $N_{s,0}$ ,  
 385  $M_0$ ,  $N_{v,0,ref}$  are the slant columns, AMF, and the reference vertical column within the reference sector, the total  
 386 glyoxal vertical column error can be written as

$$\sigma_{N_v}^2 = \frac{1}{M^2} \left( \sigma_{N_s}^2 + N_v^2 \sigma_M^2 + \sigma_{N_{s,0}}^2 + N_{v,0,ref}^2 \sigma_{M_0}^2 + M_0^2 \sigma_{N_{v,0,ref}}^2 \right) \quad (2)$$

387 where  $\sigma_{N_s}$ ,  $\sigma_M$  and  $\sigma_{N_{v,0,ref}}$  are the errors on the slant column, the air mass factor and the reference value used in  
 388 the background correction, respectively. In the following subsections, we discuss the different contributions to  
 389 each of those terms. Errors can affect the retrievals randomly or systematically (biases). While the main random  
 390 error is caused by the propagation of the instrumental photon detector shot noise on the measured radiances, the  
 391 other error components are considered as being systematic.



### 392 3.4.1. Slant column uncertainties

393 As mentioned above, the radiance measurement noise directly propagates into the glyoxal slant column retrieval  
394 and leads to large random errors  $\sigma_{N_s,rand}$  (or precision) due to the low glyoxal optical depth. Those are easily  
395 estimated using the fit residuals RMS and the covariance matrix of the cross-sections included in the fit (Danckaert  
396 et al., 2017). In the visible spectral range, the TROPOMI signal-to-noise ratio is about 1600 over dark scenes.  
397 This leads to a glyoxal VCD precision (i.e.  $\sigma_{N_s,rand}/AMF$ ) in the range of  $6\text{--}10 \times 10^{14}$  molec/cm<sup>2</sup> as illustrated in  
398 Figure 5, panel (d). This range of values is consistent with the scatter observed in the retrieved glyoxal SCDs in  
399 regions without any significant glyoxal source. Over bright scenes, for example covered by clouds or snow, those  
400 errors significantly drop because of the increased signal-to-noise ratio. For individual observations, random errors  
401 dominate and averaging is needed to extract meaningful glyoxal signals.

402 There are also systematic errors associated to the DOAS spectral fit that are mainly dominated by absorption  
403 cross-section uncertainties, by interferences with other species (O<sub>4</sub>, liquid water, Ring ...), or by other effects such  
404 as residual stray light. Those contributions are difficult to assess and can only be estimated from sensitivity tests  
405 (Lerot et al., 2010). In general, this error term can be as high as  $2\text{--}3 \times 10^{14}$  molec/cm<sup>2</sup>. However, the use of a  
406 radiance as reference in the DOAS fit and the application of a background correction removes a large part of the  
407 systematic error in the slant column fit (see section 3). As those corrections are not always sufficient to eliminate  
408 completely the SCD systematic errors due to local conditions (local pollution, residual clouds,...), we set  $\sigma_{N_s,sys}$   
409 to  $1 \times 10^{14}$  molec/cm<sup>2</sup>.

### 410 3.4.2. AMF uncertainties

411 The errors on the air mass factor depend on the input parameter uncertainties and on the sensitivity of the air mass  
412 factor to each of them. This contribution can be broken down into the squared sum (Boersma et al., 2011; Lerot  
413 et al., 2010; De Smedt et al., 2018) as

$$\sigma_{M,sys}^2 = \left( \frac{\partial M}{\partial A_s} \cdot \sigma_{A_s} \right)^2 + \left( \frac{\partial M}{\partial s} \cdot \sigma_s \right)^2 + (0.15M)^2 \quad (3)$$

414 where  $\sigma_{A_s}$  and  $\sigma_s$  are typical uncertainties on the surface albedo and profile shape, respectively.

415 The contribution of each parameter to the total air mass factor error depends on the observation conditions.  
416 Therefore, a small table of air mass factor derivatives spanning all observation conditions was computed using  
417 VLIDORT, considering glyoxal box profile shapes with different effective heights.

418 The AMF error component related to the surface reflectivity (1<sup>st</sup> term of Eq. ((3)) is calculated using an estimated  
419 uncertainty on the albedo  $\sigma_{A_s}$  of 0.02 (Kleipool et al., 2008). Note that this uncertainty can be occasionally larger,  
420 in particular at high latitudes where snow falls may cause abrupt changes in scene albedo. The uncertainty  
421 associated to the a priori profile shapes (the smoothing error) used in the retrieval is more difficult to assess,  
422 especially due to the scarcity of independent glyoxal profile measurements. For every observation, an effective  
423 height corresponding to the a priori glyoxal profile used in the AMF calculation is derived and used to extract the  
424 appropriate AMF derivative and  $\sigma_s$  is taken equal to 50hPa.



425 Formulation ((3) is valid for clear sky pixels and the stringent cloud filtering we use. However, residual clouds  
426 undoubtedly impact the radiative transfer and generally shield the lowermost atmospheric layers. Therefore, we  
427 anticipate that the clear sky assumption generally leads to a low bias on the retrieved glyoxal columns in case of  
428 residual clouds. On the other hand, the spectral interferences over bright (cloudy) scenes as discussed in section  
429 3.2 impact the retrievals the other way round. The third term in equation ((3) accounts for possible errors in the  
430 AMF model itself, including the neglect of aerosols and clouds, wavelength dependence, . . . , and is estimated to  
431 be 15% of the air mass factor (Lorente et al., 2017).

### 432 3.4.3. Background correction uncertainties

433 Although the background correction is designed to overcome systematic features/deficiencies of the slant column  
434 fitting, some errors are also associated to this procedure. In particular, systematic errors on the reference slant  
435 columns and their air mass factors are propagated into the computed correction values. Also, there is an uncertainty  
436 related to the reference glyoxal vertical column value in the reference sector. The three last terms of Eq. (2)  
437 represent the total background correction uncertainty in which  $\sigma_{N_{s,0}}$  is the systematic slant column error fixed to  
438  $1 \times 10^{14}$  molec/cm<sup>2</sup> (see above section 6.5.1), and  $M_0$  and  $\sigma_{M_0}$  are the air mass factors and their associated errors  
439 within the reference sector. In practice, those quantities are treated similarly as the reference slant columns (i.e.  
440 binned in latitude and row bins – see section 3.3).  $\sigma_{N_{v,0,ref}}$  represents the error associated to the reference value  
441  $N_{v,0,ref}$  and is fixed to  $5 \times 10^{13}$  molec/cm<sup>2</sup>.

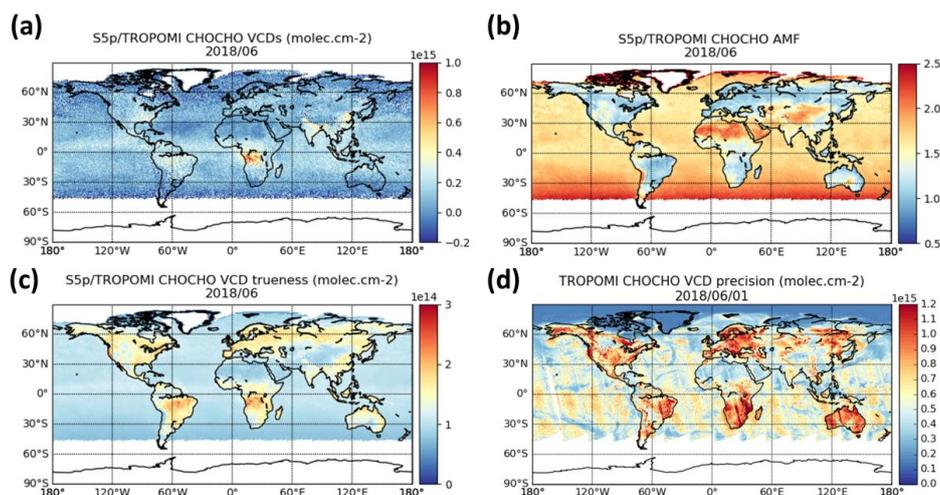
### 442 3.4.4. Total systematic uncertainties

443 Figure 5, panel (c) shows the estimated mean VCD systematic errors for the month of June 2018 when all  
444 systematic error sources are combined together using Eq. (2). Note that the conversion of the AMF error into an  
445 absolute vertical column error (2<sup>nd</sup> term of the equation) requires this error to be multiplied by the corresponding  
446 vertical column. Because of the high level of noise in the product, using the retrieved column for this would lead  
447 to a strong overestimation of the systematic error. To circumvent this, we use instead pre-computed climatological  
448 glyoxal noise-free VCDs.

449 Total glyoxal VCD systematic errors are generally in the range  $1\text{--}3 \times 10^{14}$  molec/cm<sup>2</sup>, corresponding to about 30-  
450 70% for emission regimes (columns larger than  $2 \times 10^{14}$  molec/cm<sup>2</sup>). Note that pixels strongly contaminated by  
451 clouds (cloud fraction > 20%) or covered by snow/ice are discarded. Systematic errors are expected to be large  
452 for those pixels mainly due to spectral interference effects (see section 3.2) and also because the information  
453 content on glyoxal is reduced in case of cloud shielding. Figure 5, panel (b) shows monthly mean AMFs for the  
454 same month. Small AMFs are generally caused by a priori profiles peaking near the surface, which makes the  
455 retrieval more sensitive to albedo uncertainties and to a lesser extent to the a priori profile shape uncertainties.  
456 This explains the anti-correlation between the AMFs and the systematic errors. In contrast, large AMFs are caused  
457 either by bright surface or by background a priori profiles. For such cases, systematic errors are smaller. Note that  
458 satellite column averaging kernels, defined as the Box-AMF divided by the total AMF (Eskes and Boersma, 2003),  
459 are provided for every observation. They can be used to remove the smoothing error component when comparing  
460 the satellite data to any other external data.



461



462

463 **Figure 5: (a) TROPOMI June 2018 monthly means of glyoxal tropospheric columns, glyoxal air mass factors (panel**  
464 **(b)) and glyoxal tropospheric column systematic errors (panel (c)). Scenes contaminated by clouds or Ice/snow have**  
465 **been filtered out. Panel (d) shows glyoxal tropospheric column random errors for one single day, in which all**  
466 **observations have been kept to illustrate the impact of the scene brightness.**

#### 467 4. Comparison with other satellite instruments

##### 468 4.1. Algorithmic differences for GOME-2A/B and OMI glyoxal retrievals

469 Glyoxal tropospheric columns have also been retrieved from other satellite instruments, namely GOME-2 on  
470 board the platforms Metop-A and -B and OMI on board AURA. Retrieval settings very similar to those described  
471 in the previous section were applied. For GOME-2A and B, we use data records recently produced within the  
472 operational environment of the EUMETSAT AC SAF (Valks et al., 2020). We list here the remaining differences  
473 with respect to the TROPOMI algorithmic baseline and the specificities for each instrument.

474 All data sets essentially share the same DOAS fit settings (reference cross-sections, fit window, polynomial  
475 degree...). The heterogeneity cross-sections are omitted for the GOME-2 and OMI retrievals. While the  
476 instrumental design of GOME-2 makes it weakly sensitive to scene heterogeneity, it would be beneficial for OMI  
477 to include similar cross-sections but that would imply a reprocessing of the complete slant column data set data  
478 with limited added-value for the large-scale comparison with TROPOMI that we present in the next subsection.  
479 For the GOME-2 instruments, we also fit two additional cross-sections representative of the instrumental  
480 sensitivity to light polarization as provided from the level-1 key data (EUMETSAT, 2011) as well as one pseudo  
481 cross-section to account for an along-track spectral resolution change occurring due to instrumental temperature  
482 change (Azam and Richter, 2015). Note that for GOME-2 the cross-sections are convolved with an instrumental  
483 slit function optimized as part of the wavelength calibration for every measured irradiance (De Smedt et al., 2015),  
484 which allows accounting for the known long-term drift of the GOME-2 instrument spectral response function.



485 Differences in air mass factor calculations consist only in using, over land, a priori profiles provided by  
486 IMAGESv2, the chemical transport model predecessor of MAGRITTE, at the coarser resolution of  $2.0^\circ \times 2.5^\circ$ . For  
487 the GOME-2 instruments, we use the directionally dependent Lambertian-equivalent reflectivity database  
488 produced by Tilstra et al. (2021) instead of the OMI database.

489 A background correction procedure is applied consistently with the one used for TROPOMI. The GOME-2  
490 instruments being whiskbroom scanners, there is no destriping procedure as such but instead a viewing zenith  
491 angle-dependent correction is applied, also relying on the slant columns retrieved in the Equatorial Pacific sector.  
492 This correction may account for example for remaining biases related to the instrumental polarization sensitivity.  
493 For both OMI and GOME-2, the row/VZA dependence does not show any obvious change along the orbit and the  
494 corresponding correction thus relies only on the low latitude measurements.

495 Note that the OMI and GOME-2 glyoxal products are filtered for cloudy scenes using cloud fraction lower than  
496 20% as taken from the  $O_2-O_2$  (Veeffkind et al., 2016) and OCRA (Lutz et al., 2016) cloud products, respectively.  
497 The empirical correction for strong  $NO_2$  absorption signature described in section 3.1 has been applied to those  
498 instruments as well. In the following section, we compare the TROPOMI glyoxal retrievals with the OMI and  
499 GOME-2A/B data sets. The OMI record covers the period 2005-2018, while GOME-2A/B span the periods 2007-  
500 2017 and 2013-2020, respectively. OMI and GOME-2A records were interrupted when their respective quality  
501 was degraded too severely and other instruments were available to continue the morning and afternoon time series.

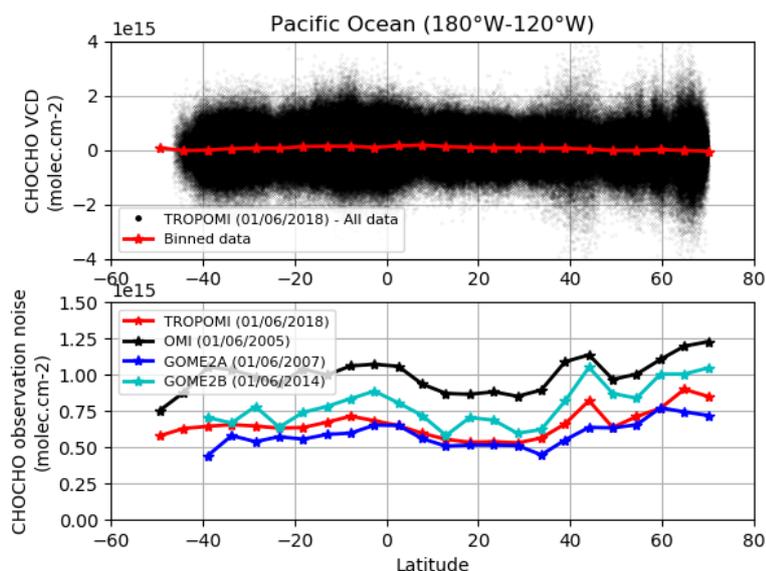
## 502 **4.2. Glyoxal satellite inter-comparison**

### 503 **4.2.1. Comparison of the noise level**

504 As mentioned before, the level of noise in the satellite glyoxal tropospheric column products is large compared to  
505 the real signal. This is illustrated in the upper panel of Figure 6 which shows all individual clear-sky TROPOMI  
506 glyoxal columns retrieved in the Pacific Ocean on June, 1<sup>st</sup> 2018 and plotted as a function of their latitude. The  
507 scatter is significant ( $\sigma \approx 5-7 \times 10^{14}$  molec/cm<sup>2</sup>) with respect to the small glyoxal VCDs averaged in  $5^\circ$ -latitude  
508 bins in this sector. The lower panel compares the standard deviation of the retrievals from TROPOMI, OMI,  
509 GOME-2A and B in the same remote sector for the 1<sup>st</sup> of June of their respective first year of operation. The  
510 scatter in the retrievals is directly linked to the instrumental signal-to-noise ratio, which is documented to be  
511 around 500 for OMI (Schenkeveld et al., 2017), 1000 for GOME-2 (Zara et al., 2018) and 1500 for TROPOMI  
512 (Kleipool et al., 2018). In practice, we see that the CHOCHO observation noise is indeed slightly larger for OMI,  
513 that GOME-2B retrievals are noisier than those from GOME-2A, which have a level of noise similar to  
514 TROPOMI. Considering the very small footprint size of TROPOMI ( $3 \times 7.5$  km<sup>2</sup> and  $3 \times 5.5$  km<sup>2</sup> after August 2019)  
515 compared to the other instruments (GOME-2:  $80 \times 40$  km<sup>2</sup>; OMI:  $13 \times 24$  km<sup>2</sup> at nadir), the TROPOMI observation  
516 noise is remarkably low. More importantly, the much larger amount of TROPOMI data compared to OMI (~15x)  
517 and GOME-2 (~100x) allows maintaining a better time or spatial resolution for a given target noise level. For  
518 example, the random error associated to the daily glyoxal column averaged in an area defined by a circle with a  
519 radius of 50 km will be less than  $0.5 \times 10^{14}$  molec.cm<sup>-2</sup> for TROPOMI, while it will remain larger than  $2.5 \times 10^{14}$   
520 molec/cm<sup>2</sup> and  $4.0 \times 10^{14}$  molec/cm<sup>2</sup> for OMI and GOME-2, respectively.



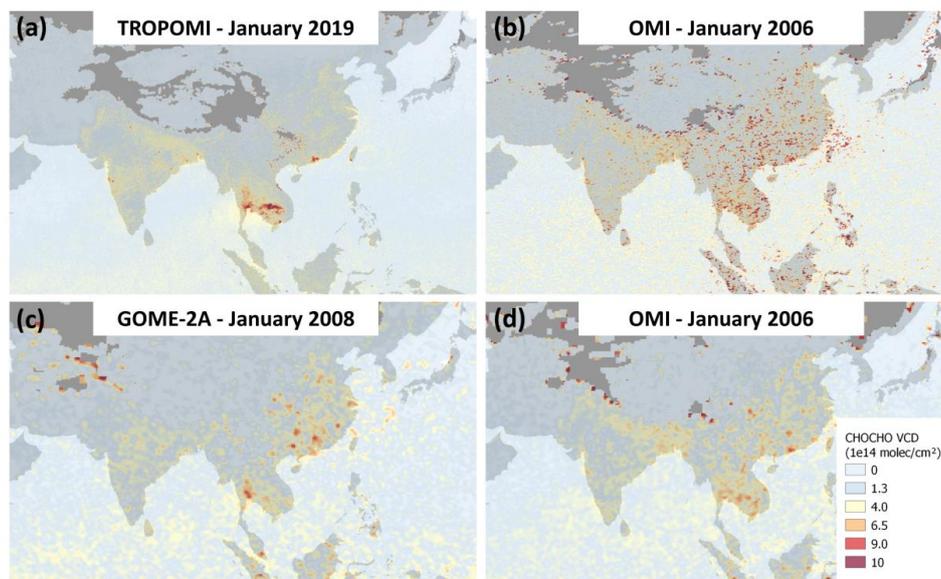
521 This is illustrated in Figure 7, which compares January monthly mean glyoxal VCD fields over Asia at the  
522 resolution of  $0.05^\circ$  for TROPOMI and OMI (upper panels) and  $0.25^\circ$  for GOME-2A and OMI (lower panels) after  
523 one year of their respective operation. At the resolution of  $0.05^\circ$ , the level of noise in the TROPOMI glyoxal map  
524 is very low and many details can be distinguished in the glyoxal spatial distribution. In particular, hot spots of  
525 glyoxal over many megacities are clearly identified (e.g. over Bangkok, New Delhi, Ho Chi Minh City,  
526 Shenzhen...) but also over Cambodia where large fires occur every year from January to March. At this resolution  
527 of  $0.05^\circ$ , the level of noise in the OMI map remains high and prevents distinguishing such details. At the coarser  
528 spatial resolution of  $0.25^\circ$ , the reduction of the noise in the OMI and GOME-2 monthly glyoxal fields appears to  
529 be sufficient to better distinguish the glyoxal spatial distribution but at the cost of a significant smoothing. In the  
530 next section, we will intercompare the four satellite products at low temporal and spatial resolution in order to  
531 minimize the impact of the noise and to identify possible systematic discrepancies.



532

533 **Figure 6** : Illustration of the level of noise in satellite CHOCHO VCD retrievals. The upper panel shows all clear sky  
534 individual glyoxal VCDs retrieved in the Pacific Ocean from TROPOMI observations on June, 1st 2018. The scatter is  
535 very large compared to the low real signal as illustrated by the data binned in  $5^\circ$  latitude bands. The lower panel  
536 compares the standard deviation of the retrievals in the same sector from TROPOMI, OMI, and GOME2A/B on the  
537 1st of June of their respective first year of operation.

538



539

540 **Figure 7: Illustration of the impact of the instrumental signal-to-noise and available amount of data on monthly mean**  
541 **glyoxal VCD fields retrieved from different satellite instruments. The (a) TROPOMI data for January 2019 gridded**  
542 **at a resolution of 0.05°, (b) OMI data for January 2006 gridded at a resolution of 0.05°, (c) GOME-2A data for**  
543 **January 2008 gridded at a resolution of 0.25°, and (d) OMI data for January 2006 gridded at a resolution of 0.25°.**  
544 **Cloudy scenes have been filtered out and a smoothing filter has been applied on the four presented fields based on a**  
545 **spatial mean with the nearest neighbouring grid cells.**

546

#### 4.2.2. Comparison of mean glyoxal fields

547

548 First, Figure 8 and Figure 9 compare seasonal maps of glyoxal VCDs generated from TROPOMI, OMI and  
549 GOME-2A/B data products. In order to reduce the data scatter for each instrument, those maps are based on long  
550 time series as indicated in the figures. Therefore, a one-to-one match is not expected. As can be seen, the  
551 consistency between the four instruments is excellent. Glyoxal patterns are captured similarly for all seasons in  
552 terms of both spatial distribution and VCD values. The largest glyoxal columns are observed in tropical regions,  
553 where biogenic emissions are important, and in regions with important fire events (e.g. Amazonia and Northern  
554 Africa in SON, Thailand/Indochina in MAM, Western US in August,...). At mid-latitudes, the glyoxal columns  
555 follow the seasonal cycle of biogenic activity with maximum values during summertime. Localized hot spots of  
556 glyoxal are visible over megacities corresponding to strong anthropogenic emissions (e.g. Northern China Plain,  
557 Bangkok, Teheran, New Delhi, Sao Paulo...). Also a persistent oceanic glyoxal signal is seen consistently by the  
558 four sensors. Note that a similar signal has been detected from ship- and airborne MAX-DOAS in the equatorial  
559 Pacific and Atlantic Oceans (Behrens et al., 2019; Sinreich et al., 2010; Volkamer et al., 2015). In contrast to  
560 TROPOMI and OMI, the level of noise in the GOME-2 data sets significantly increases over the South Atlantic  
561 Anomaly despite the application of a spike-removal procedure (section 3.1). Overall the GOME-2B maps are  
562 noisier than those from other sensors due to the lower signal-to-noise ratio of the spectra and a shorter time series.  
Compared to the UV, the sensitivity to the surface is larger in the visible, which may introduce interferences with



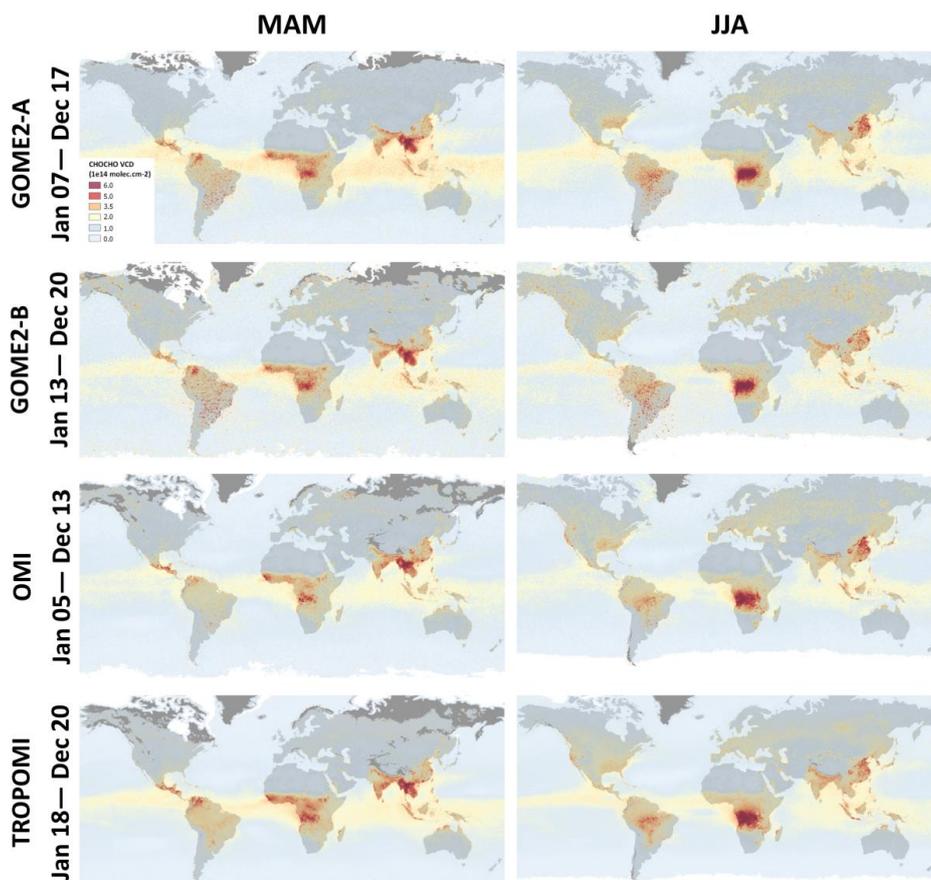
563 the spectral signature of specific ground surfaces, and thus may potentially lead to a bias on the retrieved columns.  
564 A striking example is over the Kara-Bogaz-Gol near the Caspian sea, which is one of the saltiest lakes in the world  
565 and contains large concentrations of sediments (Kosarev et al., 2009). The glyoxal signal detected over that lagoon  
566 is unlikely to be physical and likely originates from interferences with the ground reflectance spectral signature.

567

568

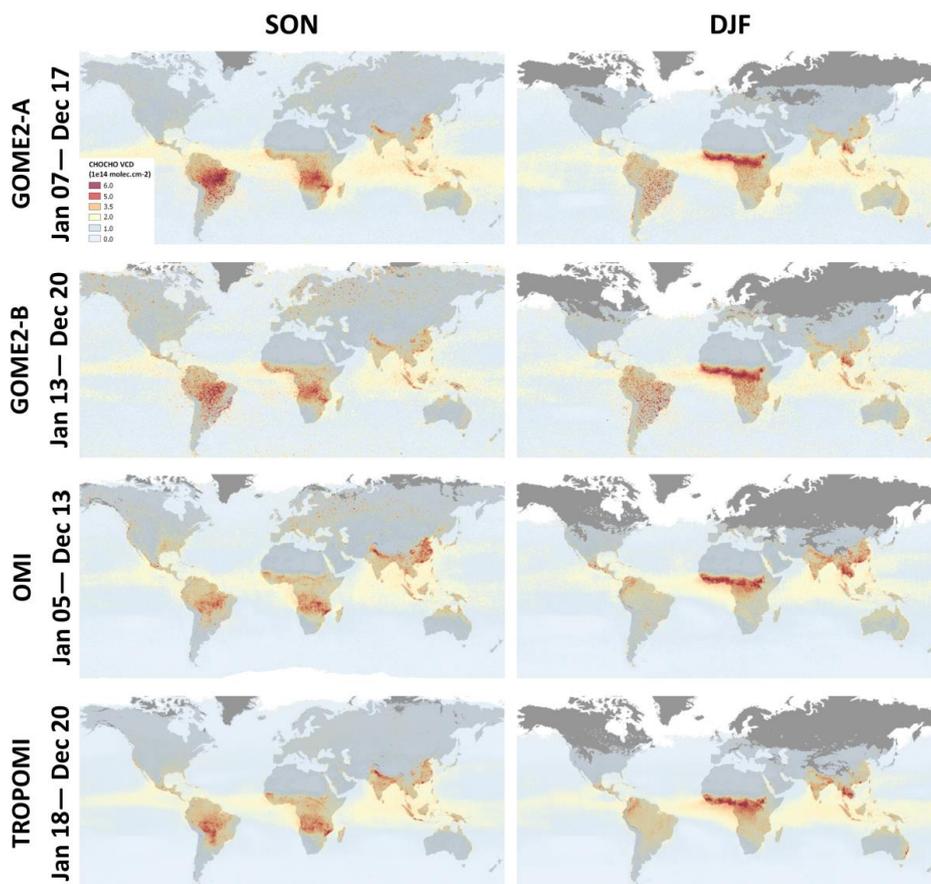
569

570



571

572 **Figure 8: Comparison of long-term averaged global CHOCHO VCDs (in  $10^{14}$  molec/cm<sup>2</sup>) derived from**  
573 **GOME-2A, GOME-2B, OMI and TROPOMI sensors, for the March-April-May period (left panels) and**  
574 **the June-July-August period (right panels).**



575

576 **Figure 9: Comparison of long-term averaged global CHOCHO VCDs (in  $10^{14}$  molec/cm<sup>2</sup>) derived from**  
577 **GOME-2A, GOME-2B, OMI and TROPOMI sensors, for the September-October-November period (left**  
578 **panels) and the December-January-February period (right panels).**

579 For a more detailed investigation of the consistency of the TROPOMI data set with OMI and GOME-2A/B, we  
580 compare complete time-series of monthly median glyoxal columns in selected regions (shown in Figure 10).

581 The red rectangles indicate the regions on which we focus in Figure 11 and Figure 12, while the global statistics  
582 for all highlighted regions are given in Figure 13. Detailed figures are provided for all regions as supplementary  
583 material (Figures S1, S2, S3, S4). Figure 11 compares directly the four full time series, while Figure 12  
584 compares the typical climatological seasonal variations as obtained by combining all available years. The error  
585 bars in the latter figure represent the interannual variability, and the 2-sigma standard deviation of the four  
586 satellite products is indicated as inset.

587 In the Tropics (e.g. Amazonia, Equatorial/North Central Africa), the four data sets are relatively stable over time.  
588 All instruments observe similar seasonal cycles and column values, although OMI appears to be slightly lower

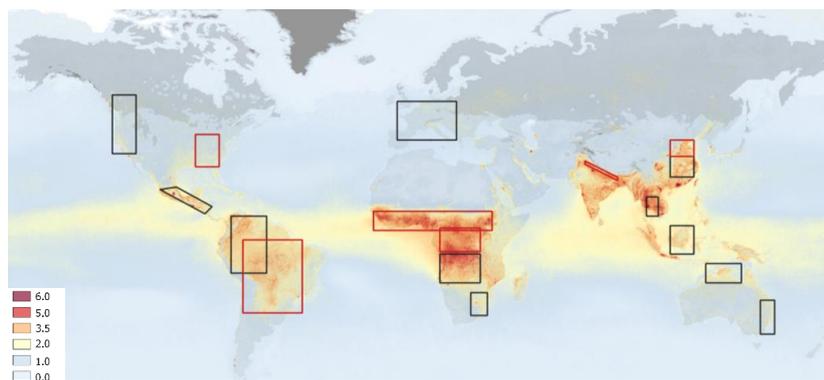


589 than the others, in particular in Equatorial Africa. The inter-annual variability in Amazonia is high compared to  
590 other regions worldwide. Glyoxal is produced in that region to a large extent by fire emissions, which are highly  
591 variable. There is a direct correlation between years with high glyoxal columns and large fire emissions (e.g. 2007,  
592 2010, 2015, 2019) as derived from the GFED database (van Der Werf et al., 2017;  
593 <https://www.globalfiredata.org/>). Interestingly, glyoxal columns measured by the morning GOME-2 instruments  
594 are larger than the OMI columns in the early afternoon during the fire seasons. This is consistent with the diurnal  
595 variation measured in satellite HCHO columns by De Smedt et al. (2015) and would deserve further investigation.  
596 Other regions display a more regular seasonal cycle, consistently seen by the four instruments.

597 In Asia, there are many hot spots, of which the origin is manifolds and strongly depends on the region and season.  
598 In addition to biogenic activities, large emissions due to fires may significantly contribute to the glyoxal columns.  
599 As illustrated in Figure 12, in the Indo-Gangetic Plain, there are typically two fire seasons in April/May and in  
600 October/November (after the Monsoon period) related to agricultural burning of wheat residue (Kumar et al.,  
601 2016), and leading to two maxima in the glyoxal VCD seasonal cycle with a significant interannual variability.  
602 For example, during the COVID-19 Indian lockdown in April/May 2020, fire activity has been reduced leading  
603 to smaller emissions (Levelt et al., 2021). This region is also highly populated, causing large emissions due to  
604 human activities. This is also true in North-East China where glyoxal columns remain significant in winter, while  
605 biogenic emissions are low during that period of the year. Although less variable than fire emissions,  
606 anthropogenic emissions may also change over time. Despite those variable emissions, the four data sets spanning  
607 different time periods show a high level of consistency. In China, it seems that the glyoxal columns as observed  
608 by OMI, GOME-2A and B are slightly reduced after 2014. This would deserve further investigation. On the other  
609 hand, any interpretation based on long-series of OMI data must be treated carefully since the instrument suffers  
610 from an evolving row anomaly (Schenkeveld et al., 2017), which impacts the stability of the product and causes  
611 an increasing number of outliers, especially at mid-latitudes. For example, over the Indo-Gangetic Plain, the OMI  
612 columns deviate regularly from the other instruments after 2014. In general, remnants of noise are also visible in  
613 the GOME-2 time series, which show somewhat less smooth time series than TROPOMI.

614 At mid-latitudes, the lower sun elevation, especially during local wintertime, makes the retrievals more  
615 challenging. Nevertheless, a small maximum is consistently observed during the local summertime. During  
616 wintertime, TROPOMI columns appear slightly lower than those from the other satellites. As mentioned before,  
617 the stronger impact of the row anomaly at mid/high-latitudes leads to a larger number of outliers in the OMI data  
618 set and to a low bias in winter after 2013/2014.

619

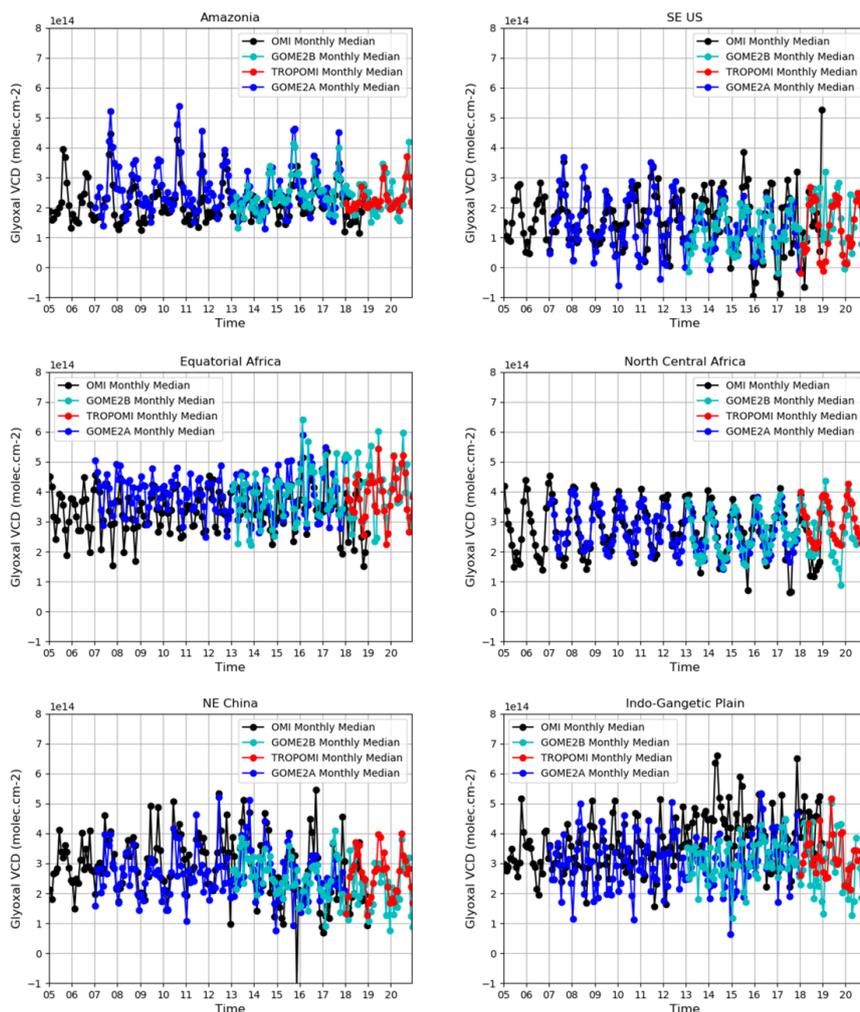


620

621 **Figure 10 : TROPOMI glyoxal VCD distribution (in  $1e14$  molec/cm<sup>2</sup>) averaged on the period January 2018-December**  
622 **2020. The rectangles represent the regions where the glyoxal products from different satellites are intercompared with**  
623 **a specific focus on red regions in Figure 11 and Figure 12.**

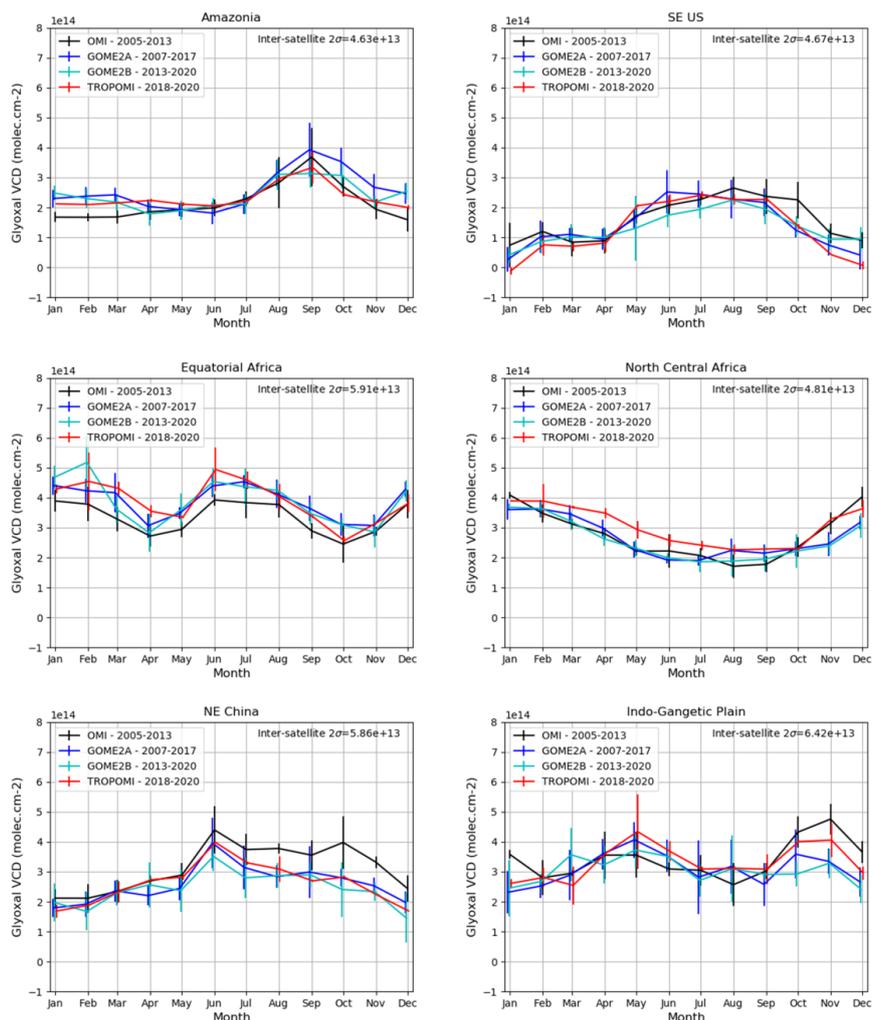
624

625



626

627 **Figure 11: Comparison of the monthly median glyoxal VCD time series from GOME-2A/B, OMI and TROPOMI in a**  
628 **few selected regions worldwide.**



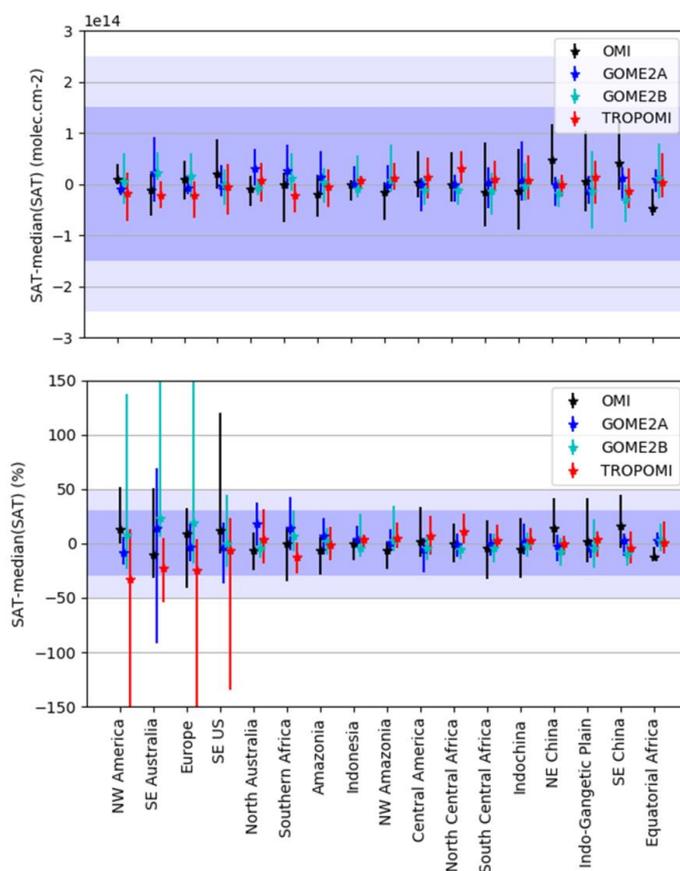
629

630 **Figure 12: Comparison of the climatological seasonal variation of the monthly median glyoxal VCDs from GOME-**  
 631 **2A/B, OMI and TROPOMI in a few selected regions worldwide. The error bars represent the interannual variability**  
 632 **as derived from the full time series.**

633 Figure 13 summarizes for all regions drawn in Figure 10 the absolute and relative deviation of each of the four data  
 634 sets with respect to the median values of the ensemble. The symbols represent the median deviation considering  
 635 all months of the year, while the error bars represent the full range of the monthly deviations. Regions are sorted  
 636 by increasing mean glyoxal vertical column amounts and light and dark blue shaded areas indicate  $2.5 \times 10^{14}$   
 637  $\text{molec}/\text{cm}^2$  (50%) and  $1.5 \times 10^{14}$   $\text{molec}/\text{cm}^2$  (30%) differences as guidelines. Inter-satellite deviations are generally  
 638 less than  $5 \times 10^{13}$   $\text{molec}/\text{cm}^2$  (20%). The large error bars in the relative differences plot for mid-latitude regions are  
 639 caused by local wintertime months during which the glyoxal content is very low, if not negligible, and are therefore  
 640 meaningless. Overall, the inter-satellite consistency of the glyoxal VCD products is excellent. In the next section,



641 we will investigate the product quality with comparisons with independent ground-based MAX-DOAS glyoxal  
642 observations at a few stations in Asia and Europe.



643  
644 **Figure 13:** Median deviation of the glyoxal VCD differences for TROPOMI, OMI, GOME-2A/B against the median  
645 value of the ensemble of the four data sets in the selected regions worldwide drawn in Figure 10. Those are plotted in  
646 absolute values ( $\text{molec}/\text{cm}^2$ ) in the upper panel and in relative values (%) in the lower panel. The error bars indicate  
647 the full range of the deviations considering climatological monthly data. Regions are sorted by increasing median  
648 glyoxal VCD value from left to right. The light and dark blue shaded area indicate differences of  $1.5 \text{ molec}/\text{cm}^2$  (30%)  
649 and  $2.5 \text{ molec}/\text{cm}^2$  (50%).

650

## 651 5. Validation with MAX-DOAS data

### 652 5.1. Description of MAX-DOAS data sets and methodology

653

654 MAX-DOAS instruments measure scattered solar light in the UV-Visible spectral range at different elevation  
655 angles above the horizon and allow retrieving information on trace gases and aerosol extinction in the altitude



656 range below 2-3km of the atmosphere, where the instrumental sensitivity is the highest. In a first approximation,  
 657 vertical columns of boundary layer gases can be estimated from MAX-DOAS measurements using a simple  
 658 geometrical approach (Brinksma et al., 2008; Hönninger et al., 2004). More elaborated approaches exploit a set  
 659 of different elevation angles to derive information on the vertical distribution of the gas concentration with up to  
 660 4 degrees of freedom, resulting in more accurate vertical columns in the 0-4 km altitude range (e.g. Beirle et al.,  
 661 2019; Clémer et al., 2010; Irie et al., 2011; Friedrich et al., 2019).

662 Glyoxal concentrations can be derived from MAX-DOAS measurements in the visible range. However, the  
 663 number of glyoxal MAX-DOAS data sets is very limited, especially those covering a period long enough to allow  
 664 the validation of satellite data during entire seasonal cycles. Moreover, MAX-DOAS retrievals are affected by  
 665 similar difficulties as satellite retrievals (noise, spectral interferences). Here, we collected an ensemble of data  
 666 sets from nine stations located in Asia and Europe (see Table 3) spanning at least one year. Altogether a wide  
 667 range of glyoxal columns and emission regimes are covered by those stations. Unfortunately, the approach to  
 668 retrieve glyoxal from MAX-DOAS has not been homogenized so far, and they cannot be considered as true  
 669 fiducial reference measurements. For example, although the same interfering species have been included in the  
 670 DOAS fits, the reference cross-section data as well as the fitting interval may vary. The design (spectral range,  
 671 spectral resolution, detector type, etc.) and operation mode of the instruments differ substantially, resulting in  
 672 different sensitivities to changes in retrieval settings. Finally, the slant-to-vertical column conversion is performed  
 673 differently from one station to another (see Table 3). Despite those limitations, the comparison of glyoxal  
 674 tropospheric columns from satellites with nine different MAX-DOAS instruments is unprecedented.

675 Among the available MAX-DOAS data sets, three (Xianghe/China, Chiba/Japan and Phimai/Thailand) are long  
 676 enough to allow a comparison with OMI and GOME-2A/B in addition to TROPOMI. The other ones span shorter,  
 677 and more recent periods and will be used only for comparison with the TROPOMI product. The Xianghe station  
 678 has the longest and stable data record, and provide vertical profiles of glyoxal. Therefore we have used this  
 679 reference station to perform a thorough analysis of the satellite product stability and of the impact of applying  
 680 satellite averaging kernels. For the other stations, we performed a more qualitative comparison of the seasonal  
 681 cycles of the glyoxal tropospheric columns. For the data collocation, we select MAX-DOAS data  $\pm 1.5$  hour around  
 682 the satellite overpass time and satellite data within a radius of 100 km (150 km for Phimai) and 20 km around the  
 683 station for GOME-2A/B/OMI and TROPOMI, respectively. Daily median glyoxal columns are computed if both  
 684 satellite and ground-based data are available and finally monthly medians of the daily median columns are  
 685 compared.

686 **Table 3 : List of MAX-DOAS stations used in the study and brief description of the approach to generate the glyoxal**  
 687 **data.**

Station (coordinates) Time range	Institution PI	Retrieval Approach and fit interval	Reference
Xianghe/China (39.75°, 116.96°E) 2010-2020 Uccle/Belgium (50.78°N, 4.35°E) 2017-2020	BIRA-IASB	Profile retrieved using an Optimal Estimation scheme  436-468 nm	(Clémer et al., 2010; Hendrick et al., 2014)
Chiba/Japan (35.63°N, 140.10°E) 2012-2020	CERES	Profile retrieved using a parametrization approach	(Hoque et al., 2018; Irie et al., 2011)



Phimai/Thailand (15.18°N, 140.10°E) 2014-2020 Pantnagar/India (29.03°N, 79.47°E) 2017-2020		436–457 nm	
Mohali/India (30.67°N, 76.73°E) May 2019 - 2020	MPIC/IISERM	Profile retrieved using a parametrization approach  400-460 nm	(Beirle et al., 2019; Kumar et al., 2020)
Athens/Greece (38.05°N, 23.80°E) 2018-2020 Vienna/Austria (48.18°N, 16.39°E) 2018-2020 Bremen/Germany (53.11°N, 8.86°E) 2018-2020	IUP-UB	Columns retrieved using the Geometrical Approximation  436-468 nm	(Alvarado et al., 2020b; Gratsea et al., 2016; Schreier et al., 2020)

688

689

690

691

## 5.2. Validation results

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

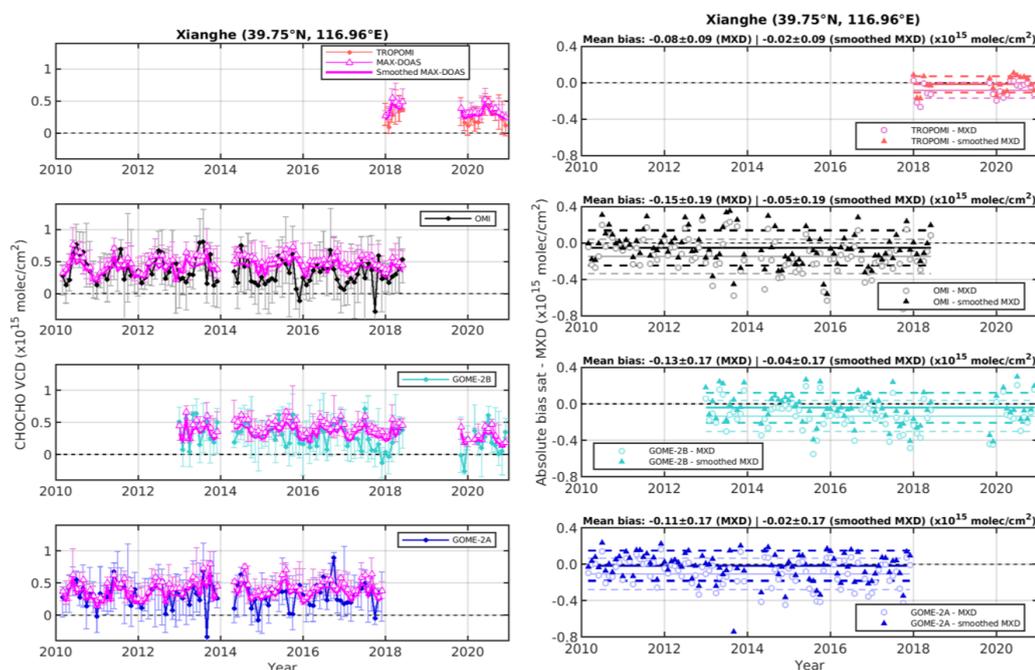
713

714

Figure 14 focuses on the comparison of monthly median glyoxal tropospheric columns retrieved from TROPOMI, OMI, GOME-2A and GOME-2B with columns from the BIRA-IASB MAX-DOAS instrument in Xianghe (China). The left panels compare the full time series for each satellite sensor with the MAX-DOAS data record. The right panels show the corresponding satellite/MAX-DOAS absolute differences. Note that the MAX-DOAS measurements have been interrupted from mid-2018 to mid-2019 due to an instrumental problem. Overall, all four satellite instruments reproduce quite well the seasonal cycle seen by the MAX-DOAS instrument. However for all of them, except for the recent TROPOMI, a degradation appears after a few years of operation. For OMI, the consistency with the MAX-DOAS is excellent before 2013, but the number of outliers increases afterwards and the columns during wintertime become too low. This is attributed to the evolving row anomaly as discussed in section 4.2. The GOME-2A/B data sets also agree quite well with the ground-based data for their first years of operation but then suffer from an increasing number of outliers after 2014 and 2017, respectively. Nonetheless, the quality of the data sets remains very reasonable. The consistency of the TROPOMI time series with the MAX-DOAS is also excellent and is characterized by a smooth temporal variability without any outliers on a monthly basis. The absolute differences shown in the right panels also clearly indicates a reduced scatter compared to the other satellites, despite the fact that a smaller overpass radius of 20 km was used instead of 100 km. This is reflected in the standard deviation of the differences given in the titles of each subpanels. The TROPOMI standard deviation is  $0.9 \times 10^{14}$  molec/cm<sup>2</sup>, while it is larger than  $1.7 \times 10^{14}$  molec/cm<sup>2</sup> for other sensors. On average, there are small negative biases with respect to the MAX-DOAS data for the four satellite time series (also given in the panel titles), ranging between  $-0.8 \times 10^{14}$  molec/cm<sup>2</sup> for TROPOMI and  $-1.5 \times 10^{14}$  molec/cm<sup>2</sup> for OMI. For this particular station, we investigated the impact of applying the satellite averaging kernels to smooth the MAX-DOAS glyoxal profiles. This process allows simulating MAX-DOAS columns which would be retrieved from the satellite algorithm, considering its own a priori profile information. The comparison of the satellite columns with the smoothed MAX-DOAS data therefore removes differences due to imperfect satellite a priori profile



715 information. As shown in Figure 14, smoothing the MAX-DOAS columns reduces the satellite/MAX-DOAS bias  
 716 to values ranging from  $-0.2 \times 10^{14}$  molec/cm<sup>2</sup> (TROPOMI) to  $-0.5 \times 10^{14}$  molec/cm<sup>2</sup> (OMI).

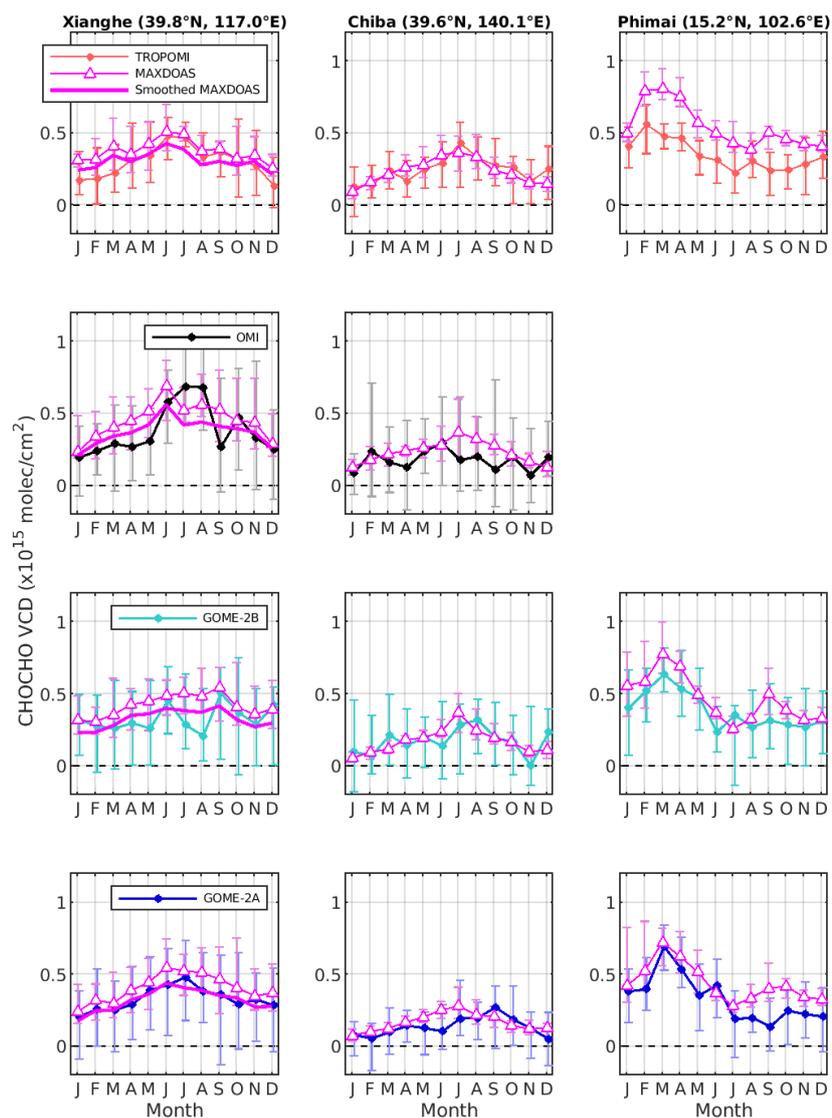


717  
 718 **Figure 14 : Comparison of the monthly median glyoxal tropospheric vertical columns retrieved from satellite and**  
 719 **MAX-DOAS (MXD) instruments in Xianghe (China). The four left panels compare the time series from TROPOMI,**  
 720 **OMI and GOME-2A/B with the MXD time series. MXD columns are also shown when smoothed with the satellite**  
 721 **averaging kernels. The error bars represent the 25 and 75% percentiles. The four right panels show the corresponding**  
 722 **time series of the satellite-MD absolute differences. Both original and smoothed MXD data are shown. Mean bias and**  
 723 **standard deviation of the differences are given in the panel titles and are also represented in the right panels with the**  
 724 **full and dashed coloured lines.**

725  
 726 In Figure 15, we compare the median satellite and MAX-DOAS seasonal cycles of the glyoxal tropospheric  
 727 columns at three stations (Xianghe, Chiba and Phimai) where the time series present a good overlap with the OMI  
 728 and GOME-2A and B records, in addition to TROPOMI. In Xianghe, the seasonal cycle of the smoothed MAX-  
 729 DOAS columns is also shown, illustrating again the reduction of the satellite/MAX-DOAS bias when the a priori  
 730 profile error component is removed. Note that the OMI and GOME-2B seasonal cycles are computed using data  
 731 until end of 2013 and 2016 to limit the impact of the increasing number of outliers. In each comparison panel, the  
 732 MAX-DOAS cycle is always computed using the same time range as the satellite instrument. Overall, the seasonal  
 733 patterns are consistently captured by the satellite and MAX-DOAS instruments. In Xianghe, the GOME-2A and



734 TROPOMI cycles follow closely the MAX-DOAS curves, although TROPOMI slightly underestimates the MAX-  
735 DOAS columns during winter months. OMI and GOME-2B also reproduce the general seasonal pattern but show  
736 a somewhat more scattered curve, likely due to their slightly less stable time series. In Chiba where the glyoxal  
737 signal is mostly driven by biogenic emissions, the agreement between the satellites and the MAX-DOAS  
738 measurements is excellent both in terms of variability and absolute values. Again, OMI shows a larger scatter (as  
739 also indicated by the larger error bars representing the inter-annual variability). In Phimai, where pyrogenic  
740 emissions are responsible for large glyoxal columns especially in the first few months of the year, the seasonal  
741 variability seen by the satellites and the MAX-DOAS is very consistent. A negative bias larger than for other  
742 stations is nevertheless observed. This can be related to other studies that identified larger biases in NO<sub>2</sub> or HCHO  
743 DOAS products for elevated column conditions (e.g. De Smedt et al., 2021; Verhoelst et al., 2021; Vigouroux et  
744 al., 2020). Possible causes for such biases are the different air masses probed by the satellite and ground-  
745 based instruments, their different vertical sensitivity as well as the a priori vertical profile information  
746 used in the retrieval algorithms.



747

748 **Figure 15** : Comparison of the monthly median glyoxal tropospheric vertical column seasonal cycle as retrieved from  
 749 **TROPOMI, OMI, GOME-2A/B and MXD** in Xianghe (China), Chiba (Japan) and Phimai (Thailand). The columns  
 750 correspond to the three stations and the rows to the different satellites. In Xianghe, MXD data smoothed with the  
 751 satellite averaging kernels are also shown. The error bars represent the interannual variability (25% and 75%  
 752 percentiles based on the full time series available). Note that the comparison of with the MAX-DOAS data in Phimai is  
 753 not shown as the latter starts in 2014 when OMI is degraded.

754

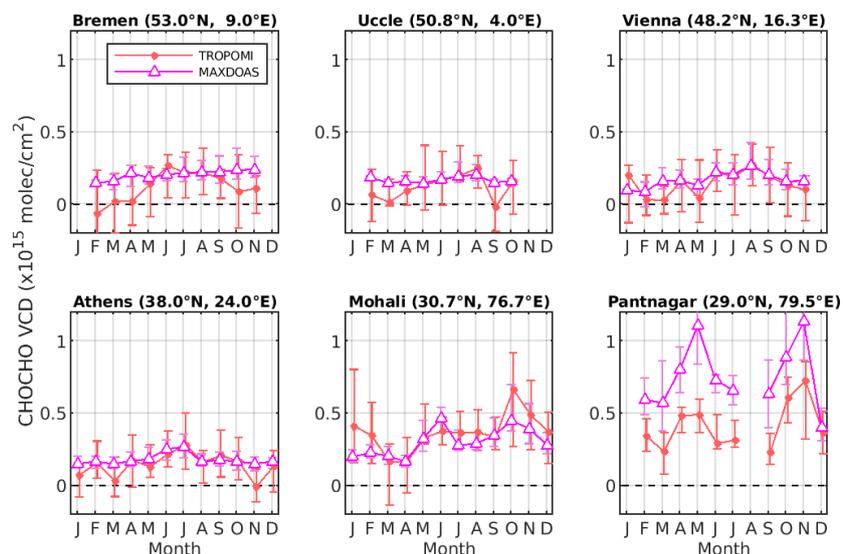


755 In Figure 16, we compare again the seasonal cycle of glyoxal VCDs retrieved from TROPOMI with that from  
756 more recent MAX-DOAS time series at six different stations. Four of them are located at mid-latitude in Europe  
757 and show relatively low glyoxal columns, while larger average values are measured at the two other stations, in  
758 Northern India (Mohali and Pantnagar). In Vienna/Austria and Athens/Greece, TROPOMI and MAX-DOAS  
759 glyoxal columns agree very well and show consistent seasonal dependencies with maximum and minimum values  
760 during summertime and wintertime, respectively. On the other hand, at the higher latitude stations of  
761 Bremen/Germany and Uccle/Belgium, the consistency of the seasonal variations seen from space and from the  
762 ground is somewhat poorer. While the glyoxal columns agree well during summertime, the satellite columns tend  
763 to underestimate MAX-DOAS values in winter, the latter showing almost no seasonal variation. Satellite glyoxal  
764 retrievals at those latitudes are challenging in winter because of the low sun elevation causing a reduced sensitivity  
765 to the lowermost atmospheric layers. As mentioned in section 3.2, observations with solar zenith angles larger  
766 than  $70^\circ$  are filtered for this reason, which explains the gap between November and January at those two stations.  
767 In Uccle, we have also tested the impact of smoothing the MAX-DOAS columns with the satellite averaging  
768 kernels (similarly as for Xianghe), which turned out to be very small. The absence of any seasonal dependence in  
769 the cities of Brussels (Uccle) and Bremen, in contrast to that observed (although limited) in Vienna and Athens,  
770 is to some extent puzzling. One should keep in mind however that the glyoxal retrievals from MAX-DOAS  
771 measurements are also challenging and it cannot be excluded that errors in ground-based data might also partly  
772 contribute to the observed differences.

773 In Mohali and Pantnagar, glyoxal columns are much larger and the seasonal variability is driven by fire emissions  
774 and meteorological factors such as the monsoon. At those two stations, the glyoxal seasonal variability is very  
775 well reproduced by TROPOMI. In terms of absolute values, the TROPOMI columns agree reasonably well in  
776 Mohali but, they significantly underestimate the (large) MAX-DOAS columns in Pantnagar. The reason why the  
777 systematic satellite/ground-based bias is so different between those two stations is unclear. MAX-DOAS columns  
778 are clearly higher in Pantnagar than in Mohali pointing either to possible local differences in air quality, not  
779 reflected in the satellite data, or to inconsistencies in the ground-based data sets. Although the agreement is  
780 excellent in Mohali, the typical behaviour is an underestimation of the columns by the satellites, as discussed  
781 before. Note also that those sites are significantly contaminated by aerosols, which are neglected in the satellite  
782 retrievals (apart from the stringent cloud filtering). MAX-DOAS data have also been analysed using very different  
783 approaches, which may also cause differences. This calls for a more detailed analysis, which would require an  
784 homogenization of the MAX-DOAS data treatment, a more sophisticated approach for the computation of the  
785 satellite AMFs (e.g. with an explicit aerosol treatment) and possibly some independent information on the glyoxal  
786 vertical distribution. This being said, the nice consistency in the glyoxal column seasonal variability by the  
787 different systems is remarkable in itself. Table 4 provides an overview of the correlation coefficient between the  
788 satellite and the MAX-DOAS glyoxal columns at all considered stations. For stations where the analysis was  
789 possible for all satellite sensors, the correlation coefficient was found to be significantly better for TROPOMI  
790 than for the other instruments. It is also clear that correlation coefficients are better for sites characterised by large  
791 and highly variable glyoxal columns (e.g. Asian stations). Apart from the Bremen station where the negative bias  
792 during winter leads to a low correlation coefficient, all other values are quite reasonable (between 0.61 and 0.87)  
793 for TROPOMI. Table 4 also gives the mean bias as derived from the comparison of the satellite and MAX-DOAS  
794 glyoxal column seasonal cycle as well as the standard deviation of the differences. As discussed above, the mean



795 differences are generally lower than  $1 \times 10^{14}$  molec/cm<sup>2</sup>, except for high columns where differences are noticeably  
 796 higher.



797

798 **Figure 16 :** Comparison of the monthly median glyoxal tropospheric vertical column seasonal cycle as retrieved from  
 799 **TROPOMI** and **MXD** at four European stations (**Bremen**, **Uccle**, **Vienna**, **Athens**) and at two Indian stations (**Mohali**,  
 800 **Pantnagar**). The error bars represent the interannual variability (25% and 75% percentiles based on the full time  
 801 series available).

802

803 **Table 4 :** Correlation coefficients between the satellite and MAX-DOAS monthly median glyoxal tropospheric  
 804 vertical columns as well as mean absolute difference and associated standard deviation at nine stations.

	Correlation coefficient Mean bias $\pm$ standard deviation ( $10^{14}$ molec/cm <sup>2</sup> )								
	Xianghe	Chiba	Phimai	Bremen	Uccle	Vienna	Athens	Mohali	Pantnagar
<b>TROPOMI</b>	0.87 -0.8 $\pm$ 0.6	0.80 0.1 $\pm$ 0.6	0.85 -2.0 $\pm$ 0.8	0.13 -0.9 $\pm$ 0.9	0.67 -0.5 $\pm$ 0.7	0.73 -0.3 $\pm$ 0.6	0.61 -0.4 $\pm$ 0.6	0.70 0.6 $\pm$ 0.9	0.78 -3.5 $\pm$ 1.5
<b>OMI (until 2013)</b>	0.70 -0.7 $\pm$ 1.3	0.32 -0.6 $\pm$ 0.8	N/A						
<b>GOME-2B (until 2016)</b>	0.37 -0.9 $\pm$ 0.9	0.66 0.0 $\pm$ 0.7	0.88 -0.8 $\pm$ 0.8						
<b>GOME-2A</b>	0.92 -0.8 $\pm$ 0.4	0.58 -0.1 $\pm$ 0.9	0.86 -1.1 $\pm$ 0.8						



## 805 6. Conclusions

806 We presented the first global TROPOMI glyoxal tropospheric column product derived from three years (2018-  
807 2020) of visible radiance measurements. The DOAS-based algorithm, which relies largely on previous  
808 developments for heritage satellite nadir-viewing instruments, has been further improved in different aspects. In  
809 particular, the use of additional pseudo cross-sections in the DOAS spectral fit allows mitigating the effect of the  
810 instrumental spectral response function perturbations in case of scene brightness inhomogeneity, which otherwise  
811 would lead to systematic biases in the retrieved glyoxal columns. This helps removing artefacts along the coasts  
812 and reducing pseudo-noise in regions covered by persistent broken clouds. The glyoxal slant columns are also  
813 empirically corrected for biases caused by the NO<sub>2</sub> misfit in case of strong absorption. Finally, the background  
814 correction procedure has been optimized for the TROPOMI characteristics and the a priori glyoxal vertical  
815 distribution, essential to the AMF computation, is now provided by the CTM MAGRITTE, an updated version of  
816 the IMAGES model, running at the higher spatial resolution of 1°x1°. The glyoxal column retrievals have been  
817 fully characterized with an error budget considering the different error components introduced in each of the  
818 algorithm modules. This allows extending the glyoxal column data product with total random and systematic error  
819 estimates provided for every observation, with corresponding averaging kernels and a priori profiles.

820 Glyoxal tropospheric columns have also been derived from data of the OMI, GOME-2A and GOME-2B satellite  
821 instruments using retrieval baselines similar to the TROPOMI algorithm. An extensive inter-comparison of those  
822 four data sets emphasised their excellent consistency with absolute mean glyoxal column differences found to be  
823 generally lower than  $0.5 \times 10^{14}$  molec/cm<sup>2</sup>. This demonstrates that glyoxal retrievals respond in the same manner  
824 to our selection of settings for all nadir-viewing satellite instruments. Because of this sensitivity, the retrievals  
825 may be easily impacted by spectral features caused by instrumental degradation. We have shown that the stability  
826 of the OMI and GOME-2 data records is somewhat degraded after a few years of operations. Glyoxal retrievals  
827 are characterized by a high level-of-noise, requiring significant spatio-temporal averaging to extract meaningful  
828 signals. With both a much larger number of observations and a finer spatial resolution, TROPOMI outperforms  
829 by far the previous instruments in its ability to provide high quality and detailed glyoxal fields.

830 Satellite observations have also been compared with a few independent MAX-DOAS data sets from stations  
831 located in Asia and Europe. Owing to the scarcity of MAX-DOAS glyoxal data sets, especially covering several  
832 seasons, this validation exercise is therefore unprecedented. Based on a thorough analysis at the Xianghe station  
833 (China), where a 10-year time series of MAX-DOAS data is available, and on the comparison of seasonal cycles  
834 at other stations, we conclude that satellite and MAX-DOAS instruments observe consistent glyoxal signals and  
835 have similar intra-annual variations. This is reflected by the strong correlation coefficients, ranging between 0.61  
836 and 0.87 for TROPOMI, with the exception of one mid-latitude station where the correlation is poorer. In general,  
837 the satellite and MAX-DOAS columns also agree in absolute values with differences less than  $1 \times 10^{14}$  molec/cm<sup>2</sup>,  
838 at least for stations with moderate columns. In Xianghe, we showed that the application of the satellite averaging  
839 kernels to the MAX-DOAS data further reduces the mean differences. There are however two stations  
840 (Phimai/Thailand and Pantnagar/India) where the satellite/MAX-DOAS bias is more significant, despite an  
841 excellent agreement between the seasonal variations. The origin of this bias is not fully understood, but it is not  
842 uncommon to have such biases in UV-Visible satellite retrievals for strongly polluted sites. In addition, we have  
843 indications that the satellite observations are low-biased during wintertime at mid-high latitudes where both the



844 glyoxal signal is weak and the sensitivity to the boundary layer is reduced. The comparisons of OMI, GOME-2  
845 and MAX-DOAS glyoxal columns also show reasonable agreement and similar intra-annual variability. Both the  
846 correlation coefficients and the scatter of the satellite/ground differences were however less good than those of  
847 TROPOMI. This points again to the better performance of TROPOMI for the detection of glyoxal from space and  
848 to its enhanced capability at providing information on VOC emissions. For future work, it would be beneficial to  
849 dedicate more efforts in the homogenization of the MAX-DOAS glyoxal retrievals in terms of both spectral  
850 analysis and slant-to-vertical column conversion in order to strengthen their potential for the validation of satellite  
851 data sets such as the one presented in this work.

#### 852 **Data availability**

853 Access to TROPOMI glyoxal tropospheric column data is possible via the GLYRETRO website  
854 (<https://glyretro.aeronomie.be/>), OMI glyoxal data can be obtained on request from the authors. Information to  
855 download the GOME-2/Metop-A and GOME-2/Metop-B glyoxal data records is provided at  
856 [https://acsaf.org/datarecord\\_access.php](https://acsaf.org/datarecord_access.php).

#### 857 **Author contributions**

858 CL is the main contributor to the study and led the writing of this paper. FH performed the validation exercise,  
859 with support from MVR and LMAA. MVR, LMAA, AR, IDS, NT, JV, HY and JVG contributed to algorithm  
860 and/or code development. TS and JFM provides the a priori modelled glyoxal profiles. PV and DL are responsible  
861 for the production of the GOME-2 glyoxal operational data records. MVR, FH, LMAA, SFS, HI, VK, TW, VS,  
862 TiW and PW contributed to operating the MAX-DOAS instruments, and to producing and providing glyoxal data.  
863 CR supervised the study. All co-authors have been involved into the discussion of results and the writing of this  
864 article.

#### 865 **Competing interests**

866 The authors have the following competing interests: Thomas Wagner is chief-executive editor of AMT. Andreas  
867 Richter is executive editor of AMT. Diego Loyola, Andreas Richter, Michel Van Roozendael and Thomas  
868 Wagner act as associate editors for AMT.  
869

#### 870 **Acknowledgments**

871 This work contains modified Copernicus Sentinel-5 Precursor satellite data (2018-2020). It has been supported  
872 by the European Space Agency via the preparation of the Level-2 Prototype Processor of the future Copernicus  
873 Sentinel-5 satellite (contract #4000118463/16/NL/AI) and via the GLYRETRO project, part of the Sentinel-  
874 5p+Innovation programme (contract #4000127610/19/I-NS). EUMETSAT, the Belgian Federal Science Policy  
875 Office (BELSPO) and the German Aerospace Center (DLR) are acknowledged for their respective financial  
876 support of the GOME-2 algorithmic developments through the AC SAF Continuous Development and Operations  
877 Phase (CDOP-3) and the ProDEx B-ACSAF contribution to the ACSAF. The Vienna MAX-DOAS instrument is  
878 part of the VINDOBONA project, which is funded by the Austrian Science Fund (FWF): I 2296-N29, the German  
879 Science Foundation (DFG): Ri1800/6-1, and A1 Telekom Austria. IISER Mohali Atmospheric Chemistry Facility  
880 is gratefully acknowledged for supporting the MAX-DOAS operations in Mohali, India. We thank Caroline Fayt  
881 and Christian Herman from BIRA-IASB for maintaining the Uccle and Xianghe MAX-DOAS instruments.



882

883 **References**

- 884 Abbot, D. S., Palmer, P. I., Martin, R. V., Chance, K. V., Jacob, D. J. and Guenther, A.: Seasonal and  
885 interannual variability of North American isoprene emissions as determined by formaldehyde column  
886 measurements from space, *Geophys. Res. Lett.*, 30(17), n/a-n/a, doi:10.1029/2003GL017336, 2003.
- 887 Aliwell, S. R., Van Roozendaal, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P.,  
888 Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J. C., Pfeilsticker, K. and Pundt, I.: Analysis for BrO in  
889 zenith-sky spectra: An intercomparison exercise for analysis improvement, *J. Geophys. Res. Atmos.*, 107(14),  
890 ACH 10-1, doi:10.1029/2001JD000329, 2002.
- 891 Alvarado, L. M. A., Richter, A., Vrekoussis, M., Wittrock, F., Hilboll, A., Schreier, S. F. and Burrows, J. P.: An  
892 improved glyoxal retrieval from OMI measurements, *Atmos. Meas. Tech.*, 7(12), 4133–4150, doi:10.5194/amt-  
893 7-4133-2014, 2014.
- 894 Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Kalisz Hedegaard, A. B., Schneising, O. and  
895 Burrows, J. P.: Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus  
896 Sentinel-5 Precursor satellite during the 2018 Canadian wildfires, *Atmos. Chem. Phys.*, 20(4), 2057–2072,  
897 doi:10.5194/acp-20-2057-2020, 2020a.
- 898 Alvarado, L. M. A., Richter, A., Lerot, C.: GLYoxal Retrievals from TROPOMI (GLYRETRO), S5p+I -  
899 Validation Report, version 1.1, S5p+I\_CHOCHO\_BIRA\_VR, available from:  
900 <https://glyretro.aeronomie.be/index.php/documents>, 2020b.
- 901 Azam, F. and Richter, A.: GOME2 on MetOp Follow-on analysis of GOME2 in orbit degradation Final Report,  
902 2015.
- 903 Barkley, M. P., Smedt, I. De, Van Roozendaal, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D.,  
904 Guenther, A., Paulot, F., Marais, E. and Mao, J.: Top-down isoprene emissions over tropical South America  
905 inferred from SCIAMACHY and OMI formaldehyde columns, *J. Geophys. Res. Atmos.*, 118(12), 6849–6868,  
906 doi:10.1002/jgrd.50552, 2013.
- 907 Barkley, M. P., González Abad, G., Kurosu, T. P., Spurr, R., Torbatian, S. and Lerot, C.: OMI air-quality  
908 monitoring over the Middle East, *Atmos. Chem. Phys.*, 17, 4687–4709, doi:10.5194/acp-17-4687-2017, 2017.
- 909 Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendaal, M., van der Werf, G. R.,  
910 Wiedinmyer, C., Kaiser, J. W., Sindelarova, K. and Guenther, A.: Nine years of global hydrocarbon emissions  
911 based on source inversion of OMI formaldehyde observations, *Atmos. Chem. Phys.*, 16(15), 10133–10158,  
912 doi:10.5194/acp-16-10133-2016, 2016.
- 913 Beekmann, M. and Vautard, R.: A modelling study of photochemical regimes over Europe: robustness and  
914 variability, *Atmos. Chem. Phys.*, 10(20), 10067–10084, doi:10.5194/acp-10-10067-2010, 2010.
- 915 Behrens, L. K., Hilboll, A., Richter, A., Peters, E., Alvarado, L. M. A., Kalisz Hedegaard, A. B., Wittrock, F.,  
916 Burrows, J. P. and Vrekoussis, M.: Detection of outflow of formaldehyde and glyoxal from the African



- 917 continent to the Atlantic Ocean with a MAX-DOAS instrument, *Atmos. Chem. Phys.*, 19(15), 10257–10278,  
918 doi:10.5194/acp-19-10257-2019, 2019.
- 919 Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y. and Wagner, T.: The Mainz profile algorithm  
920 (MAPA), *Atmos. Meas. Tech.*, 12(3), 1785–1806, doi:10.5194/amt-12-1785-2019, 2019.
- 921 Benavent, N., Garcia-Nieto, D., Wang, S. and Saiz-Lopez, A.: MAX-DOAS measurements and vertical profiles  
922 of glyoxal and formaldehyde in Madrid, Spain, *Atmos. Environ.*, 199, 357–367,  
923 doi:10.1016/j.atmosenv.2018.11.047, 2019.
- 924 Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinkma, E. J., Van Der A, R. J., Sneep, M., Van Den Oord, G.  
925 H. J., Levelt, P. F., Stammes, P., Gleason, J. F. and Bucseles, E. J.: Near-real time retrieval of tropospheric NO<sub>2</sub>  
926 from OMI, *Atmos. Chem. Phys.*, 7(8), 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.
- 927 Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V.,  
928 Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y. and Brunner, D.: An improved  
929 tropospheric NO<sub>2</sub> column retrieval algorithm for the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 4(9),  
930 1905–1928, doi:10.5194/amt-4-1905-2011, 2011.
- 931 Brinkma, E. J., Pinardi, G., Volten, H., Braak, R., Richter, A., Schönhardt, A., van Roozendaal, M., Fayt, C.,  
932 Hermans, C., Dirksen, R. J., Vlemmix, T., Berkhout, A. J. C., Swart, D. P. J., Oetjen, H., Wittrock, F., Wagner,  
933 T., Ibrahim, O. W., de Leeuw, G., Moerman, M., Curier, R. L., Celarier, E. A., Cede, A., Knap, W. H.,  
934 Veefkind, J. P., Eskes, H. J., Allaart, M., Rothe, R., PETERS, A. J. M. and Levelt, P. F.: The 2005 and 2006  
935 DANDELIONS NO<sub>2</sub> and aerosol intercomparison campaigns, *J. Geophys. Res.*, 113(D16), D16S46,  
936 doi:10.1029/2007JD008808, 2008.
- 937 Cao, H., Fu, T. M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., De Smedt, I., Zhang, Q., Van  
938 Roozendaal, M., Hendrick, F., Chance, K., Li, J., Zheng, J. and Zhao, Y.: Adjoint inversion of Chinese non-  
939 methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal,  
940 *Atmos. Chem. Phys.*, 18(20), 15017–15046, doi:10.5194/acp-18-15017-2018, 2018.
- 941 Chan, A. W. H., Chan, M. N., Surratt, J. D., Chhabra, P. S., Loza, C. L., Crouse, J. D., Yee, L. D., Flagan, R.  
942 C., Wennberg, P. O. and Seinfeld, J. H.: Role of aldehyde chemistry and NO<sub>x</sub> concentrations in secondary  
943 organic aerosol formation, *Atmos. Chem. Phys. Discuss.*, 10(4), 10219–10269 [online] Available from:  
944 <http://www.atmos-chem-phys-discuss.net/10/10219/2010/>, 2010.
- 945 Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J. and Chance, K.: Glyoxal  
946 retrieval from the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 7(11), 3891–3907, doi:10.5194/amt-7-  
947 3891-2014, 2014.
- 948 Chan Miller, C., Jacob, D. J., Marais, E. A., Yu, K., Travis, K. R., Kim, P. S., Fisher, J. A., Zhu, L., Wolfe, G.  
949 M., Hanisco, T. F., Keutsch, F. N., Kaiser, J., Min, K.-E., Brown, S. S., Washenfelder, R. A., González Abad,  
950 G. and Chance, K.: Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism,  
951 constraints from SENEX aircraft observations, and interpretation of OMI satellite data, *Atmos. Chem. Phys.*,  
952 17(14), 8725–8738, doi:10.5194/acp-17-8725-2017, 2017.



- 953 Chance, K. and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere  
954 measurements in the ultraviolet, visible, and near infrared, *J. Quant. Spectrosc. Radiat. Transf.*, 111(9), 1289–  
955 1295, doi:10.1016/j.jqsrt.2010.01.036, 2010.
- 956 Chance, K. V. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for  
957 rotational Raman scattering, and the Fraunhofer spectrum, *Appl. Opt.*, 36(21), 5224, doi:10.1364/ao.36.005224,  
958 1997.
- 959 Clémer, K., Van Roozendaal, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Spurr, R., Wang, P. and De  
960 Mazière, M.: Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS  
961 measurements in Beijing, *Atmos. Meas. Tech.*, 3(4), 863–878, doi:10.5194/amt-3-863-2010, 2010.
- 962 Coburn, S., Ortega, I., Thalman, R., Blomquist, B., Fairall, C. W. and Volkamer, R.: Measurements of diurnal  
963 variations and eddy covariance (EC) fluxes of glyoxal in the tropical marine boundary layer: description of the  
964 Fast LED-CE-DOAS instrument, *Atmos. Meas. Tech. Discuss.*, 7(10), 6245–6285, doi:10.5194/amt-7-3579-  
965 2014, 2014.
- 966 Curci, G., Palmer, P. I., Kurosu, T. P., Chance, K. and Visconti, G.: Estimating European volatile organic  
967 compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument,  
968 *Atmos. Chem. Phys.*, 10(23), 11501–11517, doi:10.5194/acp-10-11501-2010, 2010.
- 969 Danckaert, T., Fayt, C., van Roozendaal, M., De Smedt, I., Letocard, V., Merlaud, A. and Pinardi, G.: QDOAS  
970 software user manual. [online] Available from: [http://uv-  
971 vis.aeronomie.be/software/QDOAS/QDOAS\\_manual.pdf](http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf), 2017.
- 972 Danielson, J. J. and Gesch, D. B.: Global multi-resolution terrain elevation data 2010 (GMTED2010)., 2011.
- 973 De Smedt, I., Müller, J.-F., Stavrou, T., van der A, R., Eskes, H. and Van Roozendaal, M.: Twelve years of  
974 global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, *Atmos.  
975 Chem. Phys.*, 8(16), 4947–4963 [online] Available from: <http://www.atmos-chem-phys.net/8/4947/2008/>, 2008.
- 976 De Smedt, I., Stavrou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen,  
977 C., Vigouroux, C., Hermans, C., Fayt, C., Veeffkind, P., Müller, J.-F. and Van Roozendaal, M.: Diurnal, seasonal  
978 and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2  
979 observations, *Atmos. Chem. Phys.*, 15(21), 12519–12545, doi:10.5194/acp-15-12519-2015, 2015.
- 980 De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernelle, S., Van Roozendaal, M., Richter, A.,  
981 Hilboll, A., Peters, E., Pedernana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma,  
982 K. F. and Veeffkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and  
983 from the QA4ECV project, *Atmos. Meas. Tech.*, 11(4), 2395–2426, doi:10.5194/amt-11-2395-2018, 2018.
- 984 De Smedt, I., Pinardi, G., Vigouroux, C., Compernelle, S., Bais, A., Boersma, F., Chan, K.-L., Donner, S.,  
985 Eichmann, K.-U., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Lerot, C., Liu, C., Loyola, D., Piter, A.,  
986 Richter, A., Rivera Cárdenas, C., Romahn, F., George Ryan, R., Sinha, V., Vlietinck, J., Wagner, T., Wang, T.,  
987 Yu, H. and Van Roozendaal, M.: Comparative assessment of TROPOMI and OMI formaldehyde 1 observations  
988 against MAX-DOAS network column measurements. 2., doi:10.5194/acp-2021-378, n.d.



- 989 DiGangi, J. P., Henry, S. B., Kammrath, A., Boyle, E. S., Kaser, L., Schnitzhofer, R., Graus, M., Turnipseed,  
990 A., Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin, R. L., Kim, S., Nakashima, Y., Wolfe, G. M.,  
991 Kajii, Y., Apel, E. C. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A., Keutsch, F. N., Park, J.-H. J.-H.,  
992 Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin III, R. L., Kim, S., Nakashima, Y., Wolfe, G. M., Kajii,  
993 Y., Apel, E. C. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A. and Keutsch, F. N.: Observations of  
994 glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, *Atmos. Chem.*  
995 *Phys.*, 12(20), 9529–9543, doi:10.5194/acp-12-9529-2012, 2012.
- 996 Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, *Atmos. Chem.*  
997 *Phys.*, 3(5), 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.
- 998 EUMETSAT: GOME-2 Product Guide (EUM/OPS-EPS/MAN/07/0445), , (March) [online] Available from:  
999 <http://www.eumetsat.int>, 2011.
- 1000 Fu, T.-M. T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M. and Henze, D. K.: Global budgets of  
1001 atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, *J.*  
1002 *Geophys. Res.*, 113(D15), D15303, doi:10.1029/2007JD009505, 2008.
- 1003 van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., Maasakkers, J. D. and Veeffkind, J. P.: TROPOMI ATBD  
1004 of the total and tropospheric NO<sub>2</sub> data products, S5p/TROPOMI, (1.4.0), 1–76 [online] Available from:  
1005 <https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-NO2-data-products>, 2019.
- 1006 Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V.,  
1007 Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J. M., Gamache, R. R., Hodges, J. T., Jacquemart, D.,  
1008 Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M. A. H., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G.,  
1009 Barbe, A., Császár, A. G., Devi, V. M., Furtenbacher, T., Harrison, J. J., Hartmann, J. M., Jolly, A., Johnson, T.  
1010 J., Karman, T., Kleiner, I., Kyuberis, A. A., Loos, J., Lyulin, O. M., Massie, S. T., Mikhailenko, S. N.,  
1011 Moazzen-Ahmadi, N., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Polyansky, O. L., Rey, M., Rotger,  
1012 M., Sharpe, S. W., Sung, K., Starikova, E., Tashkun, S. A., Auwera, J. Vander, Wagner, G., Wilzewski, J.,  
1013 Weislo, P., Yu, S. and Zak, E. J.: The HITRAN2016 molecular spectroscopic database, *J. Quant. Spectrosc.*  
1014 *Radiat. Transf.*, 203, 3–69, doi:10.1016/j.jqsrt.2017.06.038, 2017.
- 1015 Gratsea, M., Vrekoussis, M., Richter, A., Wittrock, F., Schönhardt, A., Burrows, J., Kazadzis, S., Mihalopoulos,  
1016 N. and Gerasopoulos, E.: Slant column MAX-DOAS measurements of nitrogen dioxide, formaldehyde, glyoxal  
1017 and oxygen dimer in the urban environment of Athens, *Atmos. Environ.*, 135, 118–131,  
1018 doi:10.1016/J.ATMOENV.2016.03.048, 2016.
- 1019 Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K. and Wang, X.: The  
1020 Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated  
1021 framework for modeling biogenic emissions, *Geosci. Model Dev.*, 5(6), 1471–1492, doi:10.5194/gmd-5-1471-  
1022 2012, 2012.
- 1023 Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N.  
1024 M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M.  
1025 E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S.



- 1026 H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R. and Wildt, J.: The formation, properties and impact of  
1027 secondary organic aerosol: current and emerging issues, *Atmos. Chem. Phys.*, 9(14), 5155–5236 [online]  
1028 Available from: <http://www.atmos-chem-phys.net/9/5155/2009/>, 2009.
- 1029 Hendrick, F., Clémer, K., Wang, P., De Mazière, M., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G.,  
1030 Stavrou, T., Vlemmix, T. and Van Roozendael, M.: Four years of ground-based MAX-DOAS observations of  
1031 HONO and NO<sub>2</sub> in the Beijing area, *Atmos. Chem. Phys.*, 14(2), 765–781, doi:10.5194/acp-14-765-2014, 2014.
- 1032 Hönninger, G., von Friedeburg, C. and Platt, U.: Multi axis differential optical absorption spectroscopy (MAX-  
1033 DOAS), *Atmos. Chem. Phys.*, 4(1), 231–254, doi:10.5194/acp-4-231-2004, 2004.
- 1034 Hoque, H. M. S., Irie, H. and Damiani, A.: First MAX-DOAS Observations of Formaldehyde and Glyoxal in  
1035 Phimai, Thailand, *J. Geophys. Res. Atmos.*, 123(17), 9957–9975, doi:10.1029/2018JD028480, 2018.
- 1036 Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean,  
1037 M., Schaaf, E. and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic  
1038 compounds: A global gridded data set for 1970–2012, *Atmos. Chem. Phys.*, 17(12), 7683–7701,  
1039 doi:10.5194/acp-17-7683-2017, 2017.
- 1040 Irie, H., Takashima, H., Kanaya, Y., Boersma, K. F., Gast, L., Wittrock, F., Brunner, D., Zhou, Y. and Van  
1041 Roozendael, M.: Eight-component retrievals from ground-based MAX-DOAS observations, *Atmos. Meas.  
1042 Tech.*, 4(6), 1027–1044, doi:10.5194/amt-4-1027-2011, 2011.
- 1043 Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press. [online] Available from:  
1044 <https://press.princeton.edu/books/hardcover/9780691001852/introduction-to-atmospheric-chemistry> (Accessed  
1045 19 March 2021), 2000.
- 1046 Javed, Z., Liu, C., Khokhar, M., Tan, W., Liu, H., Xing, C., Ji, X., Tanvir, A., Hong, Q., Sandhu, O., Rehman,  
1047 A., Javed, Z., Liu, C., Khokhar, M. F., Tan, W., Liu, H., Xing, C., Ji, X., Tanvir, A., Hong, Q., Sandhu, O. and  
1048 Rehman, A.: Ground-Based MAX-DOAS Observations of CHOCHO and HCHO in Beijing and Baoding,  
1049 China, *Remote Sens.*, 11(13), 1524, doi:10.3390/rs11131524, 2019.
- 1050 Jin, X., Fiore, A., Boersma, K. F., Smedt, I. De and Valin, L.: Inferring Changes in Summertime Surface Ozone-  
1051 NO<sub>x</sub>-VOC Chemistry over U.S. Urban Areas from Two Decades of Satellite and Ground-Based Observations,  
1052 *Environ. Sci. Technol.*, 54(11), 6518–6529, doi:10.1021/acs.est.9b07785, 2020.
- 1053 Kaiser, J., Wolfe, G. M., Min, K. E., Brown, S. S., Miller, C. C., Jacob, D. J., deGouw, J. A., Graus, M.,  
1054 Hanisco, T. F., Holloway, J., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Washenfelder, R. A. and  
1055 Keutsch, F. N.: Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor  
1056 speciation, *Atmos. Chem. Phys.*, 15(13), 7571–7583, doi:10.5194/acp-15-7571-2015, 2015.
- 1057 Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat,  
1058 R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J.,  
1059 Vacanti, G., Vonk, F. and Veefkind, P.: Pre-launch calibration results of the TROPOMI payload on-board the  
1060 Sentinel-5 Precursor satellite, *Atmos. Meas. Tech.*, 11(12), 6439–6479, doi:10.5194/amt-11-6439-2018, 2018.



- 1061 Kleipool, Q. L., Dobber, M. R., de Haan, J. F. and Levelt, P. F.: Earth surface reflectance climatology from 3  
1062 years of OMI data, *J. Geophys. Res.*, 113(D18), D18308, doi:10.1029/2008JD010290, 2008.
- 1063 Kluge, F., Hüneke, T., Knecht, M., Lichtenstern, M., Rotermund, M., Schlager, H., Schreiner, B. and  
1064 Pfeilsticker, K.: Profiling of formaldehyde, glyoxal, methylglyoxal, and CO over the Amazon: normalized  
1065 excess mixing ratios and related emission factors in biomass burning plumes, *Atmos. Chem. Phys.*, 20(20),  
1066 12363–12389, doi:10.5194/acp-20-12363-2020, 2020.
- 1067 Knote, C., Hodzic, A., Jimenez, J. L., Volkamer, R., Orlando, J. J., Baidar, S., Brioude, J., Fast, J., Gentner, D.  
1068 R., Goldstein, A. H., Hayes, P. L., Knighton, W. B., Oetjen, H., Setyan, A., Stark, H., Thalman, R., Tyndall, G.,  
1069 Washenfelder, R., Waxman, E. and Zhang, Q.: Simulation of semi-explicit mechanisms of SOA formation from  
1070 glyoxal in aerosol in a 3-D model, *Atmos. Chem. Phys.*, 14, 6213–6239, doi:10.5194/acp-14-6213-2014, 2014.
- 1071 Kosarev, A. N., Kostianoy, A. G. and Zonn, I. S.: Kara-Bogaz-Gol bay: Physical and chemical evolution, *Aquat.*  
1072 *Geochemistry*, 15(1–2), 223–236, doi:10.1007/s10498-008-9054-z, 2009.
- 1073 Kumar, V., Sarkar, C. and Sinha, V.: Influence of post-harvest crop residue fires on surface ozone mixing ratios  
1074 in the N.W. IGP analyzed using 2years of continuous in situ trace gas measurements, *J. Geophys. Res.*, 121(7),  
1075 3619–3633, doi:10.1002/2015JD024308, 2016.
- 1076 Kumar, V., Beirle, S., Dörner, S., Mishra, A. K., Donner, S., Wang, Y., Sinha, V. and Wagner, T.: Long term  
1077 MAX-DOAS measurements of NO<sub>2</sub>, HCHO and aerosols and evaluation of corresponding satellite data  
1078 products over Mohali in the Indo-Gangetic plain, *Atmos. Chem. Phys.*, (2), 1–62, doi:10.5194/acp-2020-404,  
1079 2020.
- 1080 Lerot, C., Stavrou, T., De Smedt, I., Müller, J. F., Van Roozendael, M., Müller, J. F. and Van Roozendael,  
1081 M.: Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global  
1082 model, *Atmos. Chem. Phys.*, 10(24), 12059–12072, doi:10.5194/acp-10-12059-2010, 2010.
- 1083 Levelt, P. F., Van Den Oord, G. H. J., Dobber, M. R., Mälkki, A., Visser, H., De Vries, J., Stammes, P.,  
1084 Lundell, J. O. V and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote Sens.*, 44(5),  
1085 1093, doi:10.1109/TGRS.2006.872333, 2006.
- 1086 Levelt, P., Stein Zweers, D.C., Aben, I., Bauwens, M., Borsdorff, T., De Smedt, I., Eskes, H.J., Lerot, C.,  
1087 Loyola, D.G., Romahn, F., Stavrou, T., Theys, N., Van Roozendael, M., Veefkind, J.P., Verhoelst, T., Air  
1088 Quality Impacts of COVID-19 Lockdown Measures detected from space using high spatial resolution  
1089 observations of multiple trace gases from Sentinel-5P/TROPOMI, in preparation for *Atmos. Chem. Phys.*, 2021.
- 1090 Li, J., Mao, J., Min, K.-E., Washenfelder, R. A., Brown, S. S., Kaiser, J., Keutsch, F. N., Volkamer, R., Wolfe,  
1091 G. M., Hanisco, T. F., Pollack, I. B., Ryerson, T. B., Graus, M., Gilman, J. B., Lerner, B. M., Warneke, C., de  
1092 Gouw, J. A., Middlebrook, A. M., Liao, J., Welti, A., Henderson, B. H., McNeill, V. F., Hall, S. R., Ullmann,  
1093 K., Donner, L. J., Paulot, F. and Horowitz, L. W.: Observational constraints on glyoxal production from  
1094 isoprene oxidation and its contribution to organic aerosol over the Southeast United States, *J. Geophys. Res.*  
1095 *Atmos.*, 121(16), 9849–9861, doi:10.1002/2016JD025331, 2016.
- 1096 Liu, Z., Wang, Y., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., Shao, M., Chang, C.-C., Liu, S.-C.,



- 1097 Wang, H. and Chen, C.: Exploring the missing source of glyoxal (CHOCHO) over China, *Geophys. Res. Lett.*,  
1098 39(10), L10812, doi:10.1029/2012GL051645, 2012.
- 1099 Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley,  
1100 M., De Smedt, I., Van Roozendaal, M., Wang, Y., Wagner, T., Beirle, S., Lin, J. T., Krotkov, N., Stammes, P.,  
1101 Wang, P., Eskes, H. J. and Krol, M.: Structural uncertainty in air mass factor calculation for NO<sub>2</sub> and HCHO  
1102 satellite retrievals, *Atmos. Meas. Tech.*, 10(3), 759–782, doi:10.5194/amt-10-759-2017, 2017.
- 1103 Lorente, A., Boersma, K. F., Stammes, P., Tilstra, L. G., Richter, A., Yu, H., Kharbouche, S. and Müller, J.-P.:  
1104 The importance of surface reflectance anisotropy for cloud and NO<sub>2</sub> retrievals from GOME-2 and OMI, *Atmos.*  
1105 *Meas. Tech.*, 11(7), 4509–4529, doi:10.5194/amt-11-4509-2018, 2018.
- 1106 Loyola, D. G., Xu, J., Heue, K. P. and Zimmer, W.: Applying FP-ILM to the retrieval of geometry-dependent  
1107 effective Lambertian equivalent reflectivity (GE-LER) daily maps from UVN satellite measurements, *Atmos.*  
1108 *Meas. Tech.*, 13(2), 985–999, doi:10.5194/amt-13-985-2020, 2020.
- 1109 Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E.,  
1110 Rozemeijer, N., Vonk, F. and Veefkind, P.: In-flight calibration results of the TROPOMI payload on board the  
1111 Sentinel-5 Precursor satellite, *Atmos. Meas. Tech.*, 13(7), 3561–3580, doi:10.5194/amt-13-3561-2020, 2020.
- 1112 Lutz, R., Loyola, D., García, S. G. and Romahn, F.: OCRA radiometric cloud fractions for GOME-2 on MetOp-  
1113 A/B, *Atmos. Meas. Tech.*, 9(5), 2357–2379, doi:10.5194/amt-9-2357-2016, 2016.
- 1114 Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet,  
1115 D. B., Barkley, M. P., Paulot, F. and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of  
1116 formaldehyde columns, *Atmos. Chem. Phys.*, 12(14), 6219–6235, doi:10.5194/acp-12-6219-2012, 2012.
- 1117 Mason, J. D., Cone, M. T. and Fry, E. S.: Ultraviolet (250–550 nm) absorption spectrum of pure water, *Appl.*  
1118 *Opt.*, 55(25), 7163, doi:10.1364/AO.55.007163, 2016.
- 1119 Michaela Friedrich, M., Rivera, C., Stremme, W., Ojeda, Z., Arellano, J., Bezanilla, A., García-Reynoso, J. A.  
1120 and Grutter, M.: NO<sub>2</sub> vertical profiles and column densities from MAX-DOAS measurements in Mexico City,  
1121 *Atmos. Meas. Tech.*, 12(4), 2545–2565, doi:10.5194/amt-12-2545-2019, 2019.
- 1122 Müller, J.-F. and Brasseur, G.: IMAGES: A three-dimensional chemical transport model of the global  
1123 troposphere, *J. Geophys. Res.*, 100(D8), 16445, doi:10.1029/94JD03254, 1995.
- 1124 Müller, J.-F., Stavrou, T., Bauwens, M., Compennolle, S. and Peeters, J.: Chemistry and deposition in the  
1125 Model of Atmospheric composition at Global and Regional scales using Inversion Techniques for Trace gas  
1126 Emissions (MAGRITTE v1.0). Part B. Dry deposition, *Geosci. Model Dev. Discuss.*, 1–49, doi:10.5194/gmd-  
1127 2018-317, 2018.
- 1128 Müller, J.-F., Stavrou, T. and Peeters, J.: Chemistry and deposition in the Model of Atmospheric composition  
1129 at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE v1.1) – Part 1:  
1130 Chemical mechanism, *Geosci. Model Dev.*, 12(6), 2307–2356, doi:10.5194/gmd-12-2307-2019, 2019.
- 1131 Müller, J. F., Stavrou, T., Wallens, S., De Smedt, I., Van Roozendaal, M., Potosnak, M. J., Rinne, J., Munger,



- 1132 B., Goldstein, A. and Guenther, A. B.: Global isoprene emissions estimated using MEGAN, ECMWF analyses  
1133 and a detailed canopy environment model, *Atmos. Chem. Phys.*, 8(5), 1329–1341, doi:10.5194/acp-8-1329-  
1134 2008, 2008.
- 1135 Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M.,  
1136 Holdak, A., Kokhanovsky, A., Livschitz, J. and Eisinger, M.: The GOME-2 instrument on the Metop series of  
1137 satellites: instrument design, calibration, and level 1 data processing – an overview, *Atmos. Meas. Tech.*, 9(3),  
1138 1279–1301, doi:10.5194/amt-9-1279-2016, 2016.
- 1139 Myriokefalitakis, S., Vrekoussis, M., Tsigaridis, K., Wittrock, F., Richter, A., Brühl, C., Volkamer, R., Burrows,  
1140 J. P. and Kanakidou, M.: The influence of natural and anthropogenic secondary sources on the glyoxal global  
1141 distribution, *Atmos. Chem. Phys.*, 8(16), 4965–4981 [online] Available from: [http://www.atmos-chem-](http://www.atmos-chem-phys.net/8/4965/2008/)  
1142 [phys.net/8/4965/2008/](http://www.atmos-chem-phys.net/8/4965/2008/), 2008.
- 1143 Noël, S., Bramstedt, K., Bovensmann, H., Gerilowski, K., Burrows, J. P., Standfuss, C., Dufour, E. and  
1144 Veihelmann, B.: Quantification and mitigation of the impact of scene inhomogeneity on Sentinel-4 UVN UV-  
1145 VIS retrievals, *Atmos. Meas. Tech.*, 5(6), 1319–1331, doi:10.5194/amt-5-1319-2012, 2012.
- 1146 Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore,  
1147 A. and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to  
1148 formaldehyde retrievals from the Global Ozone Monitoring Experiment, *J. Geophys. Res. Atmos.*, 106(D13),  
1149 14539–14550, doi:10.1029/2000JD900772, 2001.
- 1150 Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C.,  
1151 Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B. and Sumner, A. L.: Quantifying the seasonal and  
1152 interannual variability of North American isoprene emissions using satellite observations of the formaldehyde  
1153 column, *J. Geophys. Res.*, 111(D12), D12315, doi:10.1029/2005JD006689, 2006.
- 1154 Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy: Principles and Applications*, Springer-  
1155 Verlag, 2008.
- 1156 Puķīte, J., Kühl, S., Deutschmann, T., Platt, U. and Wagner, T.: Extending differential optical absorption  
1157 spectroscopy for limb measurements in the UV, *Atmos. Meas. Tech.*, 3(3), 631–653, doi:10.5194/amt-3-631-  
1158 2010, 2010.
- 1159 Richter, A.: qa4ecv\_no2\_inhomogeneous\_scenes\_171221, in QA4ECV technical note, Treatment of  
1160 inhomogeneous scenes., 2018.
- 1161 Richter, A. and Burrows, J. P.: Tropospheric NO<sub>2</sub> from GOME measurements, *Adv. Sp. Res.*, 29(11), 1673–  
1162 1683, doi:10.1016/S0273-1177(02)00100-X, 2002.
- 1163 Richter, A., Begoin, M., Hilboll, A. and Burrows, J. P.: An improved NO<sub>2</sub> retrieval for the GOME-2 satellite  
1164 instrument, *Atmos. Meas. Tech.*, 4(6), 1147–1159, doi:10.5194/amt-4-1147-2011, 2011.
- 1165 Richter, A., Hilboll, A., Sanders, A., Peters, E. and Burrows, J. P.: Inhomogeneous scene effects in OMI NO<sub>2</sub>  
1166 observations, EGU General Assembly 2018. [online] Available from: <http://www.doas->



- 1167 [bremen.de/posters/egu\\_2018\\_richter.pdf](https://bremen.de/posters/egu_2018_richter.pdf) (Accessed 19 May 2021), 2018.
- 1168 Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., Birk, M., Bizzocchi, L.,  
1169 Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coudert, L. H., Devi, V. M., Drouin, B.  
1170 J., Fayt, A., Flaud, J. M., Gamache, R. R., Harrison, J. J., Hartmann, J. M., Hill, C., Hodges, J. T., Jacquemart,  
1171 D., Jolly, A., Lamouroux, J., Le Roy, R. J., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T.,  
1172 Mikhailenko, S., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V., Perrin, A.,  
1173 Polovtseva, E. R., Richard, C., Smith, M. A. H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G. C.,  
1174 Tyuterev, V. G. and Wagner, G.: The HITRAN2012 molecular spectroscopic database, *J. Quant. Spectrosc.*  
1175 *Radiat. Transf.*, 130, 4–50, doi:10.1016/j.jqsrt.2013.07.002, 2013.
- 1176 Rozanov, A., Rozanov, V., Buchwitz, M., Kokhanovsky, A. and Burrows, J. P.: SCIATRAN 2.0 - A new  
1177 radiative transfer model for geophysical applications in the 175–2400 nm spectral region, in *Advances in Space*  
1178 *Research*, vol. 36, pp. 1015–1019, Elsevier Ltd., 2005.
- 1179 Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Pepijn  
1180 Veeffkind, J. and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*,  
1181 10, doi:10.5194/amt-10-1957-2017, 2017.
- 1182 Schreier, S. F., Richter, A., Peters, E., Ostendorf, M., Schmalwieser, A. W., Weihs, P. and Burrows, J. P.: Dual  
1183 ground-based MAX-DOAS observations in Vienna, Austria: Evaluation of horizontal and temporal NO<sub>2</sub>,  
1184 HCHO, and CHOCHO distributions and comparison with independent data sets, *Atmos. Environ. X*, 5, 100059,  
1185 doi:10.1016/J.AEAOA.2019.100059, 2020.
- 1186 Serdyuchenko, A., Gorshchev, V., Weber, M., Chehade, W. and Burrows, J. P.: High spectral resolution ozone  
1187 absorption cross-sections &ndash; Part 2: Temperature dependence, *Atmos. Meas. Tech.*, 7(2), 625–636,  
1188 doi:10.5194/amt-7-625-2014, 2014.
- 1189 Sinreich, R., Coburn, S., Dix, B. and Volkamer, R.: Ship-based detection of glyoxal over the remote tropical  
1190 Pacific Ocean, *Atmos. Chem. Phys.*, 10(23), 11359–11371, doi:10.5194/acp-10-11359-2010, 2010.
- 1191 Stavrou, T., Müller, J.-F., De Smedt, I., Van Roozendaal, M., Van Der Werf, G. R., Giglio, L. and Guenther,  
1192 A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-  
1193 based formaldehyde columns. [online] Available from: [www.atmos-chem-phys.net/9/1037/2009/](http://www.atmos-chem-phys.net/9/1037/2009/) (Accessed 19  
1194 March 2021a), 2009.
- 1195 Stavrou, T., Müller, J.-F., De Smedt, I., Van Roozendaal, M., Kanakidou, M., Vrekoussis, M., Wittrock, F.,  
1196 Richter, A. and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne  
1197 measurements and inverse modelling, *Atmos. Chem. Phys.*, 9(21), 8431–8446 [online] Available from:  
1198 <http://www.atmos-chem-phys.net/9/8431/2009/>, 2009b.
- 1199 Stavrou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T. and Zhang, Q.: Key  
1200 chemical NO<sub>x</sub> sink uncertainties and how they influence top-down emissions of nitrogen oxides, *Atmos. Chem.*  
1201 *Phys.*, 13(17), 9057–9082, doi:10.5194/acp-13-9057-2013, 2013.
- 1202 Stavrou, T., Müller, J.-F., Bauwens, M., Smedt, I. De, Lerot, C., Roozendaal, M. Van, Coheur, P.-F.,



- 1203 Clerbaux, C., Boersma, K. F., A. R. van der, Song, Y., Jeong, S.-J., Huang, X., Song, Y., Li, M., Li, J., Zhu, T.,  
1204 Yamaji, K., Werf, G. R. van der, Huang, X., Li, M., Li, J., Song, Y., Fu, T.-M., Levelt, P. F., Smedt, I. De,  
1205 Smedt, I. De, Boersma, K. F., Lerot, C., Müller, J.-F., Stavrou, T., Stavrou, T., Müller, J.-F., Smedt, I. De,  
1206 Roozendael, M. Van, Werf, G. van der, Giglio, L., Guenther, A., Stavrou, T., Guenther, A., Karl, T., Harley,  
1207 P., Wiedinmyer, C., Palmer, P. I., Geron, C., Li, M., Andreae, M. O., Merlet, P., Akagi, S., Kurokawa, J., Sun,  
1208 L., Jin, X., Holloway, T., Safieddine, S., Chaudhry, Z., Shi, Z., Tao, J. Z., Wang, Z., Han, D., Li, S., Su, L.,  
1209 Chen, L., Deng, X., Dee, D. P., Randerson, J., Chen, Y., Werf, G., Rogers, B., Morton, D., Kaiser, J.,  
1210 Wiedinmyer, C., Kudo, S., Inomata, S., Warneke, C., Yokelson, R. J., Korontzi, S., McCarty, J., Loboda, T.,  
1211 Kumar, S., Justice, C., Fu, T.-M., Stavrou, T., Castellanos, P., Boersma, K. F., Werf, G. R. van der, Razavi,  
1212 A., Stavrou, T., Lin, J., Castellanos, P., Boersma, K. F., Torres, O., Haan, J. F. de, Barkley, M. P., Roberts,  
1213 G., Wooster, M. J., Lagoudakis, E., Stavrou, T., Miller, C. C., Jacob, D. J., et al.: Substantial Underestimation  
1214 of Post-Harvest Burning Emissions in the North China Plain Revealed by Multi-Species Space Observations,  
1215 *Sci. Reports*, Publ. online 31 August 2016; | doi:10.1038/srep32307, 6, 615–619, doi:10.1038/SREP32307, 2016.
- 1216 Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O<sub>2</sub>–O<sub>2</sub> collision pairs  
1217 between 340 and 630 nm and at atmospherically relevant pressure, *Phys. Chem. Chem. Phys.*, 15(37), 15371,  
1218 doi:10.1039/c3cp50968k, 2013.
- 1219 Theys, N., Volkamer, R., Müller, J. F., Zarzana, K. J., Kille, N., Clarisse, L., De Smedt, I., Lerot, C.,  
1220 Finkenzeller, H., Hendrick, F., Koenig, T. K., Lee, C. F., Knote, C., Yu, H. and Van Roozendael, M.: Global  
1221 nitrous acid emissions and levels of regional oxidants enhanced by wildfires, *Nat. Geosci.*, 13(10), 681–686,  
1222 doi:10.1038/s41561-020-0637-7, 2020.
- 1223 Tilstra, L., Tuinder, O., Wang, P. and Stammes, P.: Directionally dependent Lambertian-equivalent reflectivity  
1224 (DLER) of the Earth’s surface measured by the GOME-2 satellite instruments, *Atmos. Meas. Tech. Discuss.*, 1–  
1225 29, doi:10.5194/amt-2020-502, 2021.
- 1226 Tilstra, L. G., Tuinder, O. N. E., Wang, P. and Stammes, P.: Surface reflectivity climatologies from UV to NIR  
1227 determined from Earth observations by GOME-2 and SCIAMACHY, *J. Geophys. Res. Atmos.*,  
1228 doi:10.1002/2016JD025940, 2017.
- 1229 Torres, O., Bhartia, P. K., Jethva, H. and Ahn, C.: Impact of the ozone monitoring instrument row anomaly on  
1230 the long-term record of aerosol products, *Atmos. Meas. Tech.*, 11, 2701–2715, doi:10.5194/amt-11-2701-2018,  
1231 2018.
- 1232 Valks, P., Hao, N. and Lerot, C.: Algorithm Theoretical Basis Document for GOME-2 glyoxal column data  
1233 records, SAF/AC/DLR/ATBD/GLY/01; Iss. 1/B., 2020.
- 1234 Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérianne, M. F., Jenouvrier, A.  
1235 and Coquart, B.: Measurements of the NO<sub>2</sub> absorption cross-section from 42 000 cm<sup>-1</sup> to 10 000 cm<sup>-1</sup> (238–  
1236 1000 nm) at 220 K and 294 K, *J. Quant. Spectrosc. Radiat. Transf.*, 59(3–5), 171–184, doi:10.1016/S0022-  
1237 4073(97)00168-4, 1998.
- 1238 Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F.,  
1239 Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors,



- 1240 R., Kruijzinga, B., Vink, R., Visser, H. and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES  
1241 mission for global observations of the atmospheric composition for climate, air quality and ozone layer  
1242 applications, *Remote Sens. Environ.*, 120, 70–83, doi:10.1016/j.rse.2011.09.027, 2012.
- 1243 Veefkind, J. P., Haan, J. F. de, Sneep, M. and Levelt, P. F.: Improvements to the OMI O<sub>2</sub>–O<sub>2</sub> operational cloud  
1244 algorithm and comparisons with ground-based radar–lidar observations, *Atmos. Meas. Tech.*, 9(12), 6035–6049,  
1245 doi:10.5194/AMT-9-6035-2016, 2016.
- 1246 Verhoelst, T., Compernelle, S., Pinardi, G., Lambert, J. C., Eskes, H. J., Eichmann, K. U., Fjæraa, A. M.,  
1247 Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A.,  
1248 Folkert Boersma, K., Bogner, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter De  
1249 La Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagiozidis, D., Kivi,  
1250 R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., PETERS, A. J. M., Pommereau, J. P.,  
1251 Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Cárdenas, C.  
1252 R., De Miguel, L. S., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendaal, M., Pepijn Veefkind, J.,  
1253 Wagner, T., Wittrock, F., Yela González, M. and Zehner, C.: Ground-based validation of the Copernicus  
1254 Sentinel-5P TROPOMI NO<sub>2</sub> measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global  
1255 networks, *Atmos. Meas. Tech.*, 14(1), 481–510, doi:10.5194/amt-14-481-2021, 2021.
- 1256 Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I.,  
1257 Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger,  
1258 J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale,  
1259 D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendaal, M., Wang, P. and Winkler, H.:  
1260 TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-  
1261 transform infrared stations, *Atmos. Meas. Tech.*, 13(7), 3751–3767, doi:10.5194/amt-13-3751-2020, 2020.
- 1262 Vohra, K., Vodonos, A., Schwartz, J., Marais, E. A., Sulprizio, M. P. and Mickley, L. J.: Global mortality from  
1263 outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem, *Environ. Res.*,  
1264 195, 110754, doi:10.1016/j.envres.2021.110754, 2021.
- 1265 Volkamer, R., Spietz, P., Burrows, J. and Platt, U.: High-resolution absorption cross-section of glyoxal in the  
1266 UV–vis and IR spectral ranges, *J. Photochem. Photobiol. A Chem.*, 172(1), 35–46 [online] Available from:  
1267 [http://www.colorado.edu/chemistry/volkamer/publications/articles/Volkamer etal \(2005\) HR cross section](http://www.colorado.edu/chemistry/volkamer/publications/articles/Volkamer%20etal%20(2005)%20HR%20cross%20section%20glyoxal.pdf)  
1268 [glyoxal.pdf](http://www.colorado.edu/chemistry/volkamer/publications/articles/Volkamer etal (2005) HR cross section glyoxal.pdf), 2005.
- 1269 Volkamer, R., San Martini, F., Molina, L. T., Salcedo, D., Jimenez, J. L. and Molina, M. J.: A missing sink for  
1270 gas-phase glyoxal in Mexico City: Formation of secondary organic aerosol, *Geophys. Res. Lett.*, 34(19),  
1271 L19807, doi:10.1029/2007GL030752, 2007.
- 1272 Volkamer, R., Baidar, S., Campos, T. L., Coburn, S., DiGangi, J. P., Dix, B., Eloranta, E. W., Koenig, T. K.,  
1273 Morley, B., Ortega, I., Pierce, B. R., Reeves, M., Sinreich, R., Wang, S., Zondlo, M. A. and Romashkin, P. A.:  
1274 Aircraft measurements of BrO, IO, glyoxal, NO<sub>2</sub>, H<sub>2</sub>O, O<sub>2</sub>–O<sub>2</sub> and aerosol extinction profiles in the t, *Atmos.*  
1275 *Meas. Tech.*, 8(5), 2121–2148, doi:10.5194/amt-8-2121-2015, 2015.
- 1276 Voors, R., Dobber, M., Dirksen, R. and Levelt, P.: Method of calibration to correct for cloud-induced



- 1277 wavelength shifts in the Aura satellite's Ozone Monitoring Instrument, *Appl. Opt.*, 45(15), 3652–3658,  
1278 doi:10.1364/AO.45.003652, 2006.
- 1279 Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P.: Temporal and spatial variability of glyoxal as  
1280 observed from space, *Atmos. Chem. Phys.*, 9(13), 4485–4504, doi:10.5194/acp-9-4485-2009, 2009.
- 1281 Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what  
1282 can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmos. Chem. Phys.*, 10(21), 10145–  
1283 10160, doi:10.5194/acp-10-10145-2010, 2010.
- 1284 Wagner, T., Beirle, S. and Deutschmann, T.: Three-dimensional simulation of the Ring effect in observations of  
1285 scattered sun light using Monte Carlo radiative transfer models, *Atmos. Meas. Tech.*, 2(1), 113–124,  
1286 doi:10.5194/amt-2-113-2009, 2009.
- 1287 Wells, K. C., Millet, D. B., Payne, V. H., Deventer, M. J., Bates, K. H., de Gouw, J. A., Graus, M., Warneke, C.,  
1288 Wisthaler, A. and Fuentes, J. D.: Satellite isoprene retrievals constrain emissions and atmospheric oxidation,  
1289 *Nature*, 585(7824), 225–233, doi:10.1038/s41586-020-2664-3, 2020.
- 1290 van der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van  
1291 Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J. and Kasibhatla, P. S.: Global fire emissions  
1292 estimates during 1997–2016, *Earth Syst. Sci. Data*, 9(2), 697–720, doi:10.5194/essd-9-697-2017, 2017.
- 1293 Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle,  
1294 S., Platt, U. and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space,  
1295 *Geophys. Res. Lett.*, 33(16), L16804, doi:10.1029/2006GL026310, 2006.
- 1296 World Health Organization: Ambient air pollution: A global assessment of exposure and burden of disease.  
1297 [online] Available from: <https://apps.who.int/iris/handle/10665/250141> (Accessed 19 March 2021), 2016.
- 1298 Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., Van Geffen, J. H. G. M., Beirle, S., Wagner, T.,  
1299 Van Roozendaal, M., Marchenko, S., Lamsal, L. N. and Eskes, H. J.: Improved slant column density retrieval of  
1300 nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty  
1301 characterization, and trends, *Atmos. Meas. Tech. Discuss.*, 11(7), 1–47, doi:10.5194/amt-11-4033-2018, 2018.
- 1302 Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R. and Wang, P.: An improved tropospheric NO<sub>2</sub> retrieval for  
1303 OMI observations in the vicinity of mountainous terrain, *Atmos. Meas. Tech.*, 2(2), 401–416, doi:10.5194/amt-  
1304 2-401-2009, 2009.
- 1305
- 1306
- 1307
- 1308
- 1309