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

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

[1] Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[2] Alfred Wegner Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany
[3] University of Bremen, Institute of Environmental Physics, Bremen, Germany
[4] Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft und Raumfahrt (DLR), Oberpfaffenhofen, Germany
[5] Center for Environmental Remote Sensing, Chiba University, Japan
[6] Max Planck Institute for Chemistry (MPIC), Mainz, Germany
[7] Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences, Vienna, Austria
[8] Indian Institute of Science Education and Research Mohali, Mohali, India
[9] Institute of Atmospheric Physics, Chinese Academy of Sciences (CAS), Beijing, China
[10] European Space Agency, ESRIN, Frascati, Italy

Correspondence to: Christophe Lerot, Christophe.Lerot@aeronomie.be

Abstract. We present the first global glyoxal (CHOCHO) tropospheric column product derived from the TROPOspheric Monitoring Instrument (TROPOMI) on board of the Sentinel-5 Precursor satellite. Atmospheric glyoxal results from the oxidation of other non-methane volatile organic compounds (NMVOCs) and from direct emissions caused by combustion processes. Therefore, this product is a useful indicator of VOC emissions. It is generated with an improved version of the BIRA-IASB scientific retrieval algorithm relying on the Differential Optical Absorption Spectroscopy (DOAS) approach. Among the algorithmic updates, the DOAS fit now includes corrections to mitigate the impact of spectral misfits caused by scene brightness inhomogeneity and strong NO_2 absorption. The product comes along with a full error characterization, which allows providing random and systematic error estimates for every observation. Systematic errors are typically in the range of $1\text{-}3\times10^{14}$ molec/cm² (~30-70% in emission regimes) and originate mostly from a priori data uncertainties as well as spectral interferences with other absorbing species. The latter may be at the origin, at least partly, of an enhanced glyoxal signal over Equatorial Oceans and further investigation is needed to mitigate them. Random errors are large ($>6\times10^{14}$ molec/cm²) but can be reduced by averaging observations in space and/or time. Benefiting from a high signal-to-noise ratio and a large number of small-size observations, TROPOMI provides glyoxal tropospheric column fields with an unprecedented level of details.

Using the same retrieval algorithmic baseline, glyoxal column data sets are also generated from the Ozone Monitoring Instrument (OMI) on Aura and from the Global Ozone Monitoring Experiment-2 (GOME-2) on board of Metop-A and Metop-B. Those four data sets are intercompared over large-scale regions worldwide and show a high level of consistency. The satellite glyoxal columns are also compared to glyoxal columns retrieved from ground-based Multi-Axis (MAX-) DOAS instruments at nine stations in Asia and Europe. In general, the satellite

44 and MAX-DOAS instruments provide consistent glyoxal columns both in terms of absolute values and variability.
45 Correlation coefficients between TROPOMI and MAX-DOAS glyoxal columns range between 0.61 and 0.87.
46 The correlation is only poorer at one mid-latitude station, where satellite data appears low biased during
47 wintertime. The mean absolute glyoxal columns from satellite and MAX-DOAS generally agree well for
48 low/moderate columns with differences less than 1×10^{14} molec/cm². A larger bias is identified at two sites where
49 the MAX-DOAS columns are very large. Despite this systematic bias, the consistency of the satellite and MAX-
50 DOAS glyoxal seasonal variability is high.

1. Introduction

52 Exposure to poor air quality kills millions of people annually (e.g. Vohra et al., 2021; World Health Organization,
53 2016) due to natural and human emissions of a large range of particulate matters and gases, including among
54 others nitrous oxides (NOx) , sulphur dioxide, carbon monoxide, methane and volatile organic compounds
55 (VOCs). The latter, in combination with NOx, play a significant role in the secondary production of tropospheric
56 ozone (Jacob, 2000), which is highly toxic for the respiratory system and also contributes to global warming
57 because of its absorption in the thermal infrared. Global measurements of atmospheric concentrations of the ozone
58 precursors are therefore crucial. The number of VOCs that can be found in the atmosphere is manifold, but only
59 a few of them can be probed using remote sensing techniques. For example, formaldehyde (HCHO) measurements
60 have been used in many studies as a proxy for probing emissions of non-methane VOCs of biogenic, pyrogenic
61 and anthropogenic origin (e.g. Abbot et al., 2003; Barkley et al., 2013; Bauwens et al., 2016; Beekmann and
62 Vautard, 2010; Curci et al., 2010; Jin et al., 2020; Marais et al., 2012; Palmer et al., 2006; Stavrakou et al., 2016;
63 Wells et al., 2020).

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

79 Glyoxal has three absorption bands in the visible spectral range that have been exploited to remotely retrieve
80 information on its atmospheric abundance using the Differential Optical Absorption Spectroscopy method
81 (DOAS, Platt and Stutz, 2008) applied to ground-based (e.g. Benavent et al., 2019; Hoque et al., 2018; Javed et

82   al., 2019; Schreier et al., 2020), air-borne (e.g. Kluge et al., 2020; Volkamer et al., 2015), ship-borne (e.g. Behrens

83   et al., 2019; Sinreich et al., 2010) and space-based instruments. The first global glyoxal tropospheric column

84   observations from space have been realized by Wittrock et al. (2006) using nadir measurements from the

85   SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY) instrument. Based

86   on this pioneering work, different glyoxal data products were derived from the Global Ozone Monitoring

87   Experiment-2 (GOME-2) (Lerot et al., 2010; Vrekoussis et al., 2009) and from the Ozone Monitoring Instrument

88   (OMI) (Alvarado et al., 2014; Chan Miller et al., 2014). All those different products rely on a similar DOAS

89   approach, but generally differ from each other by the choice of the fit settings and of the auxiliary input data.

90   In general, the glyoxal optical depth is very low (< 5e-4), typically one order of magnitude smaller than the NO_2

91   optical depth in the same spectral range. This results in retrievals prone-to-noise, requiring to average many of

92   them to extract meaningful glyoxal signals. With an enhanced spatial resolution resulting in a number of

93   observations more than ten times larger than provided by its predecessor OMI, the TROPOspheric Monitoring

94   Instrument (TROPOMI), operating since 2017, allows observing weak atmospheric absorbers with an

95   unprecedented level of spatio-temporal details. This has been illustrated by Alvarado et al. (2020a) who

96   investigated the large amounts of formaldehyde and glyoxal emitted by the intense North-American wildfires in

97   August 2018 as observed by TROPOMI for several days and over long distances. Theys et al. (2020) have

98   evaluated the respective contributions to the hydroxyl radical production in fresh fire plumes from nitrous acid,

99   VOCs and other sources with the support of different TROPOMI data sets, including the glyoxal data product

100   described here.

101   This work presents the latest version of the BIRA-IASB scientific glyoxal tropospheric column retrieval algorithm

102   that has been applied to three years of TROPOMI measurements, and also to data from the predecessor nadir

103   instruments OMI and GOME-2A/B. The quality of the TROPOMI glyoxal retrievals is investigated with (1) a

104   global intercomparison of the satellite glyoxal data products generated with a common algorithm and (2)

105   comparisons with independent glyoxal measurements from a series of Multi-AXis DOAS (MAX-DOAS)

106   instruments located at nine stations in Asia and Europe.

107   After a brief introduction of the satellite instruments used in this study in Section 2, the retrieval algorithm and its

108   different steps are described in Section 3, with emphasis on the updated and innovative aspects compared to

109   heritage studies. This section also presents the typical random and systematic errors associated to the retrievals

110   and how they are estimated for each individual measurement. Section 4 presents the evaluation of the inter-satellite

111   consistency by comparing both seasonal global spatial patterns as seen from different instruments as well as

112   monthly mean time series and seasonal cycles in a series of selected large-scale regions. Finally, Section 5 presents

113   validation results based on MAX-DOAS data.

114   **2. TROPOMI and other nadir-viewing satellite sensors**

115   TROPOMI was launched on 13 October 2017 on board of the Sentinel-5 precursor platform. It flies on a sun-

116   synchronous Low Earth Orbit (LEO) with an ascending node crossing the equator at the local time of 13:30. In

117   the series of Sentinel missions from the European Union Copernicus programme, it is the first one dedicated to

118   atmospheric composition. The instrument operates in a nadir viewing mode and measures Earthshine radiances

119   and solar irradiances in the ultraviolet (UV), visible, near infrared and short infrared spectral bands. It aims at

120 providing column amounts of a number of key pollutants, such as ozone (O_3), NO_2, SO_2, HCHO, CO, CH_4 as well
121 as cloud and aerosol parameters. TROPOMI offers a quasi-daily global coverage at the unprecedented spatial
122 resolution of 3.5x5.5 km² (3.5x7 km² before August 2019) in the UV-visible spectral range. It is an imager-type
123 instrument using a two-dimensional Charge Coupled Device (CCD) for the light measurements, the detector
124 columns being used for the spectral resolution while the rows are binned to resolve spatially the 2600 km across-
125 track swath into 450 individual ground pixels. The spectral resolution of the instrument is about 0.5 nm and offers
126 a remarkably high signal-to-noise ratio of about 1500 in band 4 (405-500 nm) used in this study. More details on
127 the instrument and its performance can be found in (Kleipool et al., 2018; Ludewig et al., 2020; Schenkeveld et
128 al., 2017; Veefkind et al., 2012). The TROPOMI measurements allow to derive the vertical columns of multiple
129 species, some of them not included among the operational products listed above. Glyoxal is one of them and the
130 details on how its column quantities are retrieved will be described in the next section.

131 The TROPOMI design strongly inherits from past nadir-viewing sensors, and in particular from the Ozone
132 Monitoring Instrument (OMI) that we use to evaluate the TROPOMI glyoxal product presented in this work. OMI
133 (Levelt et al., 2006) is also an imager instrument and flies on an early afternoon orbit since October 2004. The
134 OMI swath, divided into 60 across-track pixels with a size varying from 13x24 km² (at nadir) to 13x150km² (at
135 the edges), allowed a daily global coverage before being limited in 2008 by the so-called row anomaly. The latter
136 consists in a modification of the signal recorded by OMI at specific rows, due to a mechanical obstruction of the
137 field of view, and leads to lower quality spectral measurements (Torres et al., 2018). We also exploit spectral
138 measurements from the Global Ozone Monitoring Experiment-2 (GOME-2) instruments aboard the Metop-A and
139 Metop-B platforms. In contrast to OMI and TROPOMI, the GOME-2 instruments (Munro et al., 2016) fly on
140 early morning LEOs with local equator crossing times around 09:30 and are scanning spectrometers, meaning that
141 across-track pixels are successively sounded. The scan is divided into 24 pixels for a total swath of 1920 km,
142 providing global coverage in 1.5 day. Each pixel has a size of 80x40km². After the launch of Metop-B, the GOME-
143 2A swath was reduced to 960 km in July 2013, leading to ground pixel two times smaller.

3. Description of the Algorithm

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

153 The glyoxal algorithm presented here largely inherits from past developments for predecessor nadir-viewing
154 satellite sensors (Alvarado et al., 2014, 2020; Chan Miller et al., 2014; Lerot et al., 2010; Vrekoussis et al., 2009;
155 Wittrock et al., 2006). Figure 1 illustrates for one day of TROPOMI data the resulting main output of every
156 algorithmic components, which we further describe in the following sections, with emphasis on their specificities.

157 The retrievals are provided with estimates for the random and systematic errors, which are discussed in subsection
158 3.4.

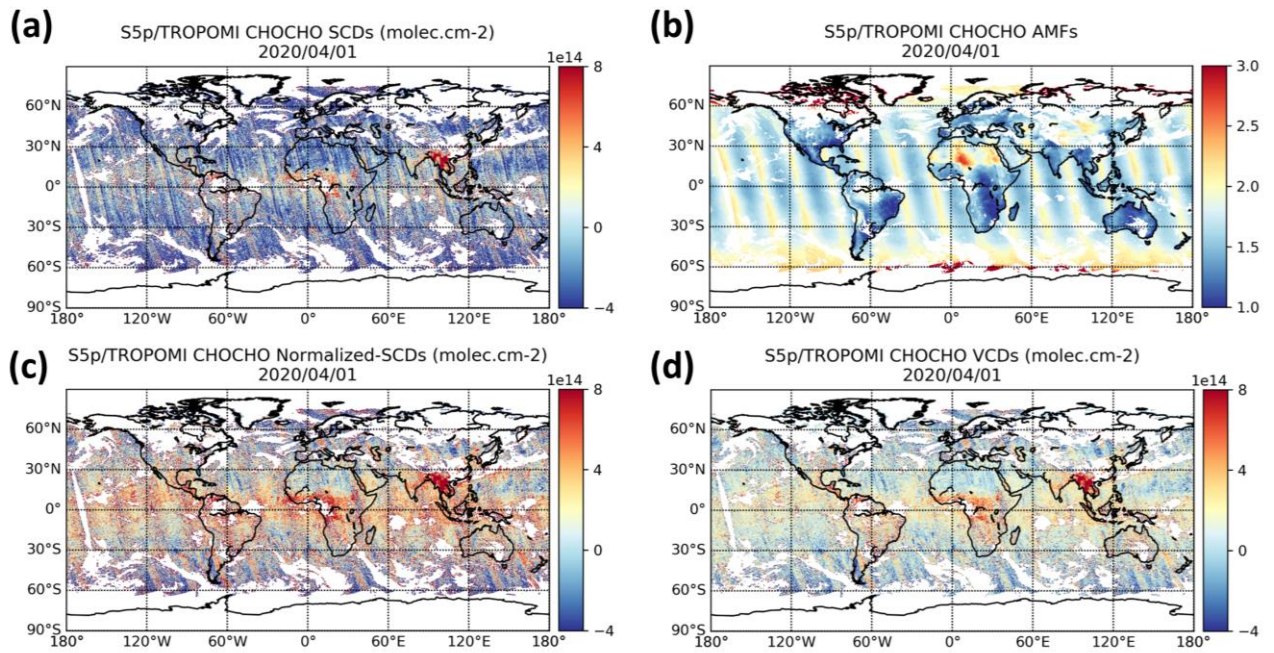



Figure 1: Illustration of the different glyoxal algorithmic steps for one day of TROPOMI data (01/04/2020). (a) Glyoxal SCDs retrieved from the DOAS spectral fits; (b) Glyoxal AMFs; (c) Glyoxal normalized-SCDs resulting from the background correction procedure; (d) Final retrieved glyoxal VCDs. The data has been gridded at the spatial resolution of 0.25°.

3.1. DOAS fit

165 To exploit the glyoxal absorption bands, we use a fitting window from 435 to 460 nm encompassing the two most

166 intense bands, which has shown in the past to provide reliable results (Barkley et al., 2017; Lerot et al., 2010).

167 This has been confirmed by sensitivity tests carried out by Alvarado et al. (2014) and Chan Miller et al. (2014).

168 Owing to its low optical depth ($<5x10^{-4}$), any poorly fitted feature in the radiance measurements may affect the

169 retrieved glyoxal SCD. It is therefore crucial to account for any physical or instrumental effect in order to optimise

170 the fit quality as much as possible. Different aspects of the algorithm contribute to achieve this.

171 The wavelength grids of the measured spectra are recalibrated before the actual DOAS fits with a cross-correlation

172 procedure (Danckaert et al., 2017; De Smedt et al., 2018) during which the position of the lines in the measured

173 irradiance spectrum is fitted to an external solar atlas (Chance and Kurucz, 2010), convolved to the satellite

174 spectral resolution. This recalibration procedure is done once per orbit and separately for every detector row of

175 the instrument.

176 Although the DOAS fit generally uses an irradiance as the reference spectrum, it is common practice, in the case

177 of weak tropospheric absorbers, to replace it by a mean radiance spectrum recorded in a remote region where the

178 concentration of the gas of interest is low (e.g. De Smedt et al., 2018). This allows reducing the presence of

179 systematic biases caused by spectral interferences and/or instrumental limitations. In particular, the use of one

180 separate mean radiance spectrum per detector row minimizes the presence of so-called stripes in the product

181 typical of imager-type instruments such as OMI or TROPOMI. Here we compute those mean radiance spectra on

182 a daily basis by averaging for each row all spectra located within the equatorial Pacific Ocean (15°S-15°N;

183 120°W-180°W).

184 The selected settings for the DOAS fits rely on the aforementioned past studies and are summarized in Table 1.

185 The latest available cross-sections for species absorbing in the selected fitting window are included in the fit, i.e.

186 O_3, NO_2, O_2-O_2, water vapour and liquid water in addition to glyoxal. Note that the water vapour cross-section

187 has been generated for a temperature of 293K and a pressure of 1013hPa using the HITRAN2012 database

188 (Rothman et al., 2013) as we found that the latest HITRAN2016 version (Gordon et al., 2017) led to poorer fit

189 quality. Sensitivity tests have shown that the retrieved glyoxal SCDs are significantly impacted by the choice of

190 the H_2O cross-section but also of its temperature. Effective water vapour temperatures (computed as the mean of

191 temperature profiles weighted by typical H_2O concentration profiles) are generally close to our selected value in

192 regions with high water vapour content. This high sensitivity nevertheless points to the importance of having

193 accurate water vapour cross-section, especially in regards of its possible influence on glyoxal fields over oceans

194 (Chan Miller et al., 2014) (see section 4.2.3). The temperature dependence of the NO_2 absorption is taken into

195 account by including a second cross-section, taken as the difference between NO_2 cross-sections reported at 2

196 temperatures (220 and 294K) as proposed by Alvarado et al. (2014) and Chan Miller et al. (2014) for their

197 respective OMI glyoxal products. Consistently with Alvarado et al. (2014), we found that fitting the liquid water

198 optical depth in the glyoxal fitting window performs as well as fixing it to a value previously determined in a

199 larger spectral interval as proposed in the past (Lerot et al., 2010) with a limited presence of systematically

200 negative glyoxal columns over remote oceans. Vibrational Raman scattering on remote ocean water also

201 introduces some spectral structures caused by the filling-in of Fraunhofer lines. However, (Peters et al., 2014)

202 have shown the simultaneous fit of the liquid water cross-section and of an intensity offset (see below) efficiently

203 considers all remote ocean-related structures. A number of additional cross-sections are included in the fit to

204 consider (1) Inelastic scattering (Ring effect) introduces high-frequency structures that are treated as a pseudo-

205 absorber (Chance and Spurr, 1997); (2) Intensity offsets in the spectra, caused for example by residual straylight,

206 are corrected for by fitting the inverse of the reference spectrum (Danckaert et al., 2017); (3) heterogeneity of the

207 scene brightness may also introduce high frequency structures, which are considered with pseudo-cross-sections

208 (more details hereafter). All those cross-sections are generated at the instrumental spectral resolution by using the

209 key data Instrumental Spectral Response Functions provided for all individual detector rows. During the DOAS

210 procedure, the earthshine radiance spectrum is further aligned with the reference, by allowing it to be shifted and

211 stretched in wavelength. In addition, the DOAS fit procedure includes a spike removal scheme as described in

212 Richter et al. (2011) enabling to filter out from the fit individual corrupted radiance measurements, and hence to

213 reduce the noise in the product.

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

Fitting interval	435-460 nm
Absorption cross-sections	
Glyoxal	Volkamer et al. (2005)

Ozone	Serdyuchenko et al. (2014), 223K, I_0 effect-corrected (Aliwell et al., 2002)
NO$_2$	Vandaele et al. (1998), 220K and 294K, I_0 effect-corrected
O$_4$ (O$_2$-O$_2$)	Thalman and Volkamer (2013), 293K
H$_2$O (vapour)	Rothman et al. (2013), 293K
H$_2$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)
Other parameters	
Polynomial	3rd order
Intensity offset correction	1st-order offset (additional cross-section taken as the inverse of the reference spectrum)
Earthshine wavelength shift	1st-order shift
Reference spectrum (E$_0$)	Daily average of radiances, per detector row, selected in equatorial Pacific (Lat: [-15° 15°], Long: [180°-240°])

215

3.1.1. Scene heterogeneity

216

217   Any intensity variation within the probed scene taking place perpendicularly to the instrumental slit (i.e. along

218   track) leads to perturbations of the instrumental spectral slit function (ISRF) (Noël et al., 2012; Voors et al., 2006).

219   Richter et al. (2018) have shown that those perturbations lead to a degradation of the NO$_2$ DOAS spectral fit

220   quality and to systematic biases on the retrieved slant columns. Such abrupt intensity changes occur for example

221   along the coasts, mountains or cloud edges. Glyoxal retrievals are also affected by such scene heterogeneity as

222   illustrated in Figure 2 over the Horn of Africa and Middle East. This figure shows in the panel (a) that the root

223   mean square (RMS) of the DOAS fit residuals is systematically higher along the coasts but also over land where

224   contamination by broken clouds or abrupt elevation changes cause discontinuities in brightness fields. The stripes

225   visible in this figure are due to the smaller pixel size (and hence lower signal-to-noise ratio) on the edges of the

226   across-track field of view. The panel (c) shows that there are some collocated artificial patterns (positive/negative

227   biases) in the mean retrieved glyoxal slant column field. The latter result from spectral interferences with the

228   signature introduced by the ISRF distortion. Richter et al. (2018) showed that those spectral interferences can be

229   significantly reduced with additional cross-sections in the DOAS fit scaling the possible scene heterogeneity

230   signature. Those cross-sections are generated with a statistical analysis of the fit residuals for many observations

231   in the Pacific Ocean as a function of the level of scene heterogeneity. The latter can be computed using radiance

232   measurements at higher spatial resolution available in the TROPOMI level-1 data at a limited number of

233   wavelengths. It ranges between -1 and +1 and is close to/deviates from 0 for homogenous/heterogeneous scenes,

234   the sign indicating the part of the ground pixel that dominates the scene brightness. Following this approach, two

235   additional cross-sections corresponding to the systematic residuals of scenes with an heterogeneity factor

236   larger/smaller than +/- 0.08 have been added to the DOAS baseline and both the fit residuals and the identified

237   glyoxal biases have been reduced as illustrated in the right panels (b) and (d) of Figure 2. This effect is particularly

visible along coasts and mountains but also over lands where some pseudo-noise caused by persistent broken clouds is also largely reduced. Although significantly correlated, including the two heterogeneity cross-sections leads to a further improvement of the fit quality, likely due to a slit function perturbation that depends on the radiance distribution within the nominal ground pixel. Note that a third cross-section derived from the mean residuals of homogeneous scenes is also added, which explains why the fit RMS are also reduced (but less drastically) in homogeneous scenes. This cross-section has no impact on the retrieved glyoxal SCDs and allowed mostly isolating systematic residuals due to scene heterogeneity only for the pseudo cross-sections creation.

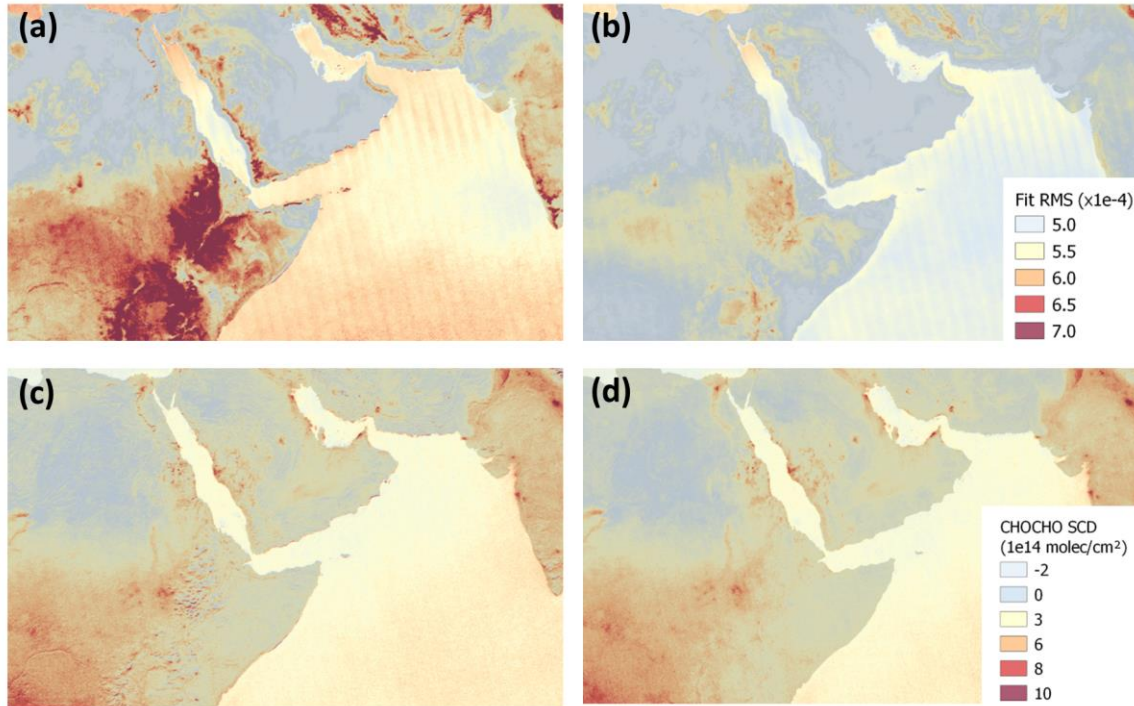



Figure 2 : Impact of scene brightness heterogeneity on glyoxal retrievals in the fitting window 435-460 nm over the Horn of Africa and Middle-East. The panels (a) and (b) show mean fit residuals RMS for the year 2019 without and with (left and right) pseudo-cross sections to correct for spectral signatures introduced by scene heterogeneity. The panels (c) and (d) show the corresponding mean glyoxal slant column densities. Only observations with cloud fractions less than 20% are considered.

3.1.2. Empirical correction for strong NO_2 absorption

The DOAS approach assumes that the wavelength dependence of the effective light path within the fit interval can be neglected. Although this assumption is generally reasonable, it may fail in case of strong absorption by one (or more) species, of which the slant column density becomes dependent on the wavelength (Puķīte et al., 2010). In that case, fitting the optical depth of that species by a simple scaling of its cross-section is inaccurate and the fit quality is degraded. Puķīte et al. (2010) have shown that fitting additional cross-sections resulting from a Taylor expansion of the wavelength-dependent slant column corrects for its variability within the fit window. As mentioned before, the high sensitivity of glyoxal retrievals to potential sources of misfit was a motivation to further investigate its sensitivity to extreme NO_2 concentration levels.

260   For this purpose, synthetic spectra were generated at a spectral resolution of 0.5 nm with the radiative transfer

261   model SCIATRAN (Rozanov et al., 2005) for a satellite nadir-viewing geometry and two different solar zenith

262   angles. In those simulations, inelastic scattering was neglected and a large range of tropospheric NO_2 columns

263   was covered by scaling the NO_2 a priori profile. The TROPOMI DOAS baseline described above was then applied

264   to those simulated spectra in order to retrieve CHOCHO SCDs and evaluate the error as a function of the NO_2

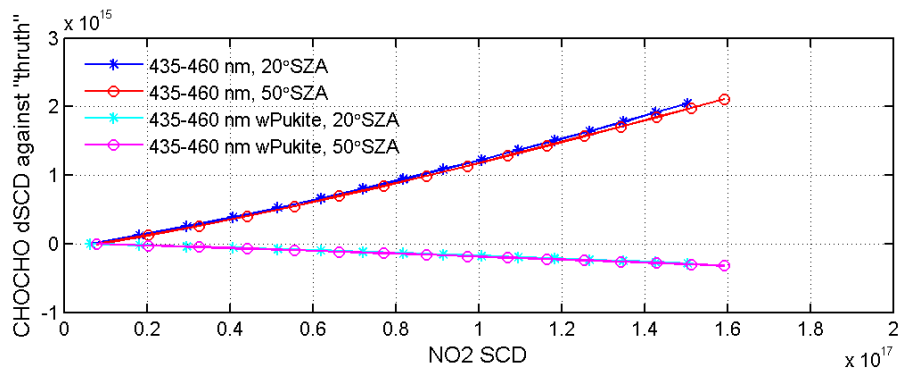
265   SCD as illustrated in Figure 3. Results clearly point to a CHOCHO SCD error increasing with the NO_2 SCD. Note

266   that the exact error magnitude may change slightly depending on the NO_2 vertical distribution and on the actual

267   atmospheric content. On the other hand, adding two so-called Pukite cross-sections (Puķīte et al., 2010) resulting

268   from a first order expansion of the NO_2 slant column around the wavelength and the vertical optical depth

269   significantly reduces the errors.

270

271   **Figure 3 : Absolute error (in molec/cm²) on the retrieved CHOCHO SCD as a function of the NO_2 SCD for simulated**

272   **spectra in a nadir-viewing satellite geometry and for two solar zenith angles. The reference "true" CHOCHO SCD is**

273   **taken as the value retrieved for the lowest NO2 SCD scenario. The error increases with the NO_2 SCD when Pukite**

274   **cross-sections are not included in the fit, but remains small otherwise.**

275   On this basis, the impact of adding the Pukite cross-sections to the DOAS baseline has been investigated using

276   one month of TROPOMI data. A wintertime period was chosen (December 2019) to favour the number of

277   observations with large NO_2 concentrations, in particular in China but also in other megacities in the Northern

278   Hemisphere. Figure 4 (upper panel) displays the monthly mean NO_2 SCDs in December 2019, and (middle panel)

279   the mean impact on the retrieved CHOCHO SCDs of introducing the Pukite terms in the DOAS spectral fit

280   baseline. The CHOCHO SCD differences caused by the Pukite terms are also plotted as a function of the NO_2

281   SCDs to better visualize the correlation (lower panel). For regions with enhanced NO_2 concentrations ($>2x10^{16}$

282   molec/cm²) (e.g. China, India, Teheran), the Pukite cross-sections lead to a systematic reduction of the CHOCHO

283   SCDs, consistent with the closed-loop tests described above. A small improvement of the fit quality is found (not

284   shown). Unexpectedly, the impact of those additional cross-sections on the CHOCHO SCDs can also be non-

285   negligible in regions with low NO_2 columns: positive differences are for example observed over equatorial oceans,

286   but also over South America and Africa. The correlation plot of Figure 4 clearly shows these two regimes. While

287   the impact of the Pukite cross-sections on the glyoxal retrievals is understood and reliable for large NO_2 SCDs,

288   their influence at low NO_2 SCD is more questionable and likely results from spectral interferences occurring

289   between the different fitted spectra (e.g. with the Ring signature), which introduces additional noise in the product.

290   To avoid this, rather than fitting additional cross-sections, we introduce an empirical correction applied to the

291   glyoxal SCDs. This correction consists in subtracting from the glyoxal SCD a NO_2-SCD dependent value, directly

292   prescribed from the linear regression fit through the sensitivity test results for all observations worldwide from

293   December 2019, with NO_2 SCDs larger than $2x10^{16}$ molec/cm² as illustrated in Figure 4 (c). It is worth noting that

294   the regression fit results agree well with the glyoxal SCD errors estimated from the simulations presented above

295   (Figure 3). For extreme pollution conditions such as what can be found in China during Wintertime, this correction

296   may lead to glyoxal column reduction up to 30%.

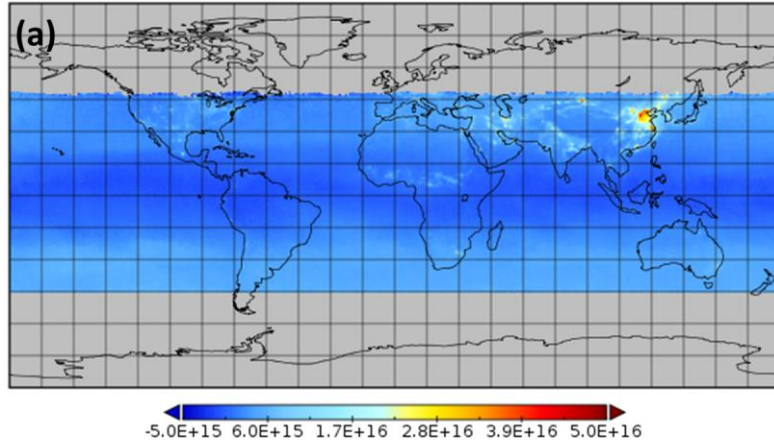
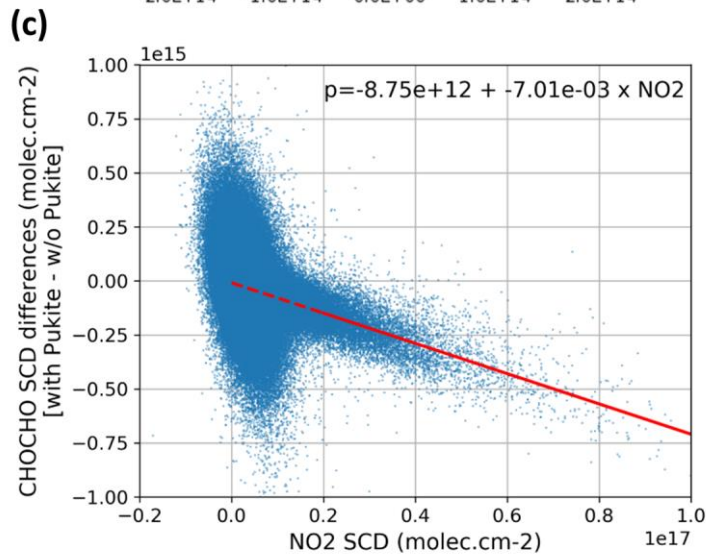
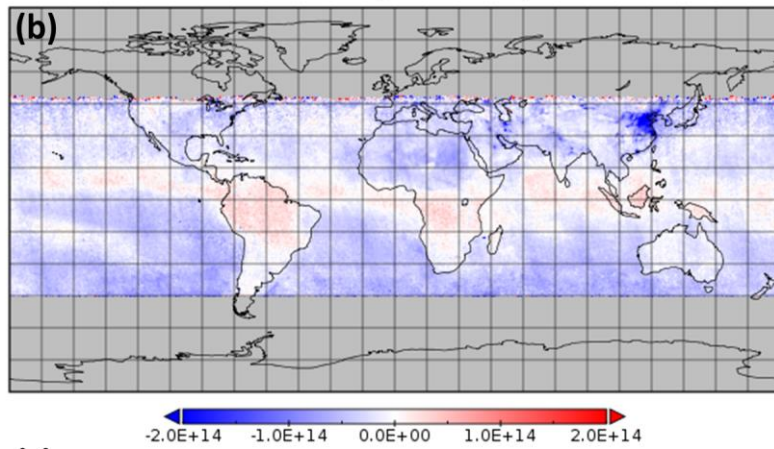



Figure 4 : (a) Monthly mean NO₂ SCDs retrieved from TROPOMI data in December 2019. Panel (b) illustrates the CHOCHO SCD absolute differences (molec/cm²) due to the incorporation of the Pukite et al. (2010) cross-sections in the DOAS spectral fit and panel (c) shows the correlation between those differences and the NO₂ SCDs. The red line corresponds to a linear regression fit through all points with NO₂ SCD larger than 2×10^{16} molec/cm².

3.2. Air Mass Factor computation

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

Typically, the sensitivity of nadir-viewing UV-Visible instruments is reduced in the lowermost atmospheric layers because of Rayleigh scattering. However, this sensitivity depends strongly on the observation geometry, on the surface reflectivity and altitude and on the presence of clouds. For example, the sensitivity is generally further reduced for low sun elevation. For this reason, retrievals with solar zenith angles larger than 70° are filtered out. We use a pre-computed five-dimensional look-up table of Box-AMFs spanning all observation conditions (see Table 2) and from which appropriate values are linearly interpolated for every TROPOMI observation. This interpolation uses as input the observation angles provided in the level-1 data, surface elevation taken from the GMTED2010 topography (Danielson and Gesch, 2011) and surface albedo extracted from the OMI minimum Lambertian Equivalent Reflectivity climatology (Kleipool et al., 2008). The spatial resolution of the latter database (0.5°x0.5°) is coarse compared to the TROPOMI footprint and neglects anisotropy, which may introduce significant errors (Lorente et al., 2018). However, at the time of writing, it is the only database available at the S5p overpass time although new Lambertian Equivalent Reflectivity climatologies relying on past works (e.g. Loyola et al., 2020; Tilstra et al., 2021, 2017) are currently being prepared. On the other hand, the level of noise in glyoxal retrievals generally requires averaging in space and/or time which in turn will reduce part of those error sources. We also neglect the impact of clouds and aerosols on the radiative transfer. Instead we apply a stringent cloud filtering approach: only observations with an effective cloud fraction (as retrieved in the same spectral range and provided in the TROPOMI operational NO_2 product (van Geffen et al., 2019) lower than 20% are conserved. This approach is motivated by the fact that glyoxal slant columns tend to be biased high over bright scenes because of poorly understood residual spectral interferences (e.g. with the Ring signature). Similarly, scenes covered by snow and ice are also discarded.

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

332

Typical monthly-dependent a priori glyoxal vertical profile shapes necessary to perform the AMF computations have been calculated at the different satellite overpass times with the global Chemistry Transport Model MAGRITTE developed at BIRA-IASB, which inherits from the IMAGES model (Bauwens et al., 2016; Müller and Brasseur, 1995; Stavrakou et al., 2009b, 2013). This model runs at 1°x1° resolution and calculates the distribution of 182 chemical compounds, of which 141 species undergo transport. The modelled troposphere is vertically divided in 40 levels between the surface and the lower stratosphere and meteorological fields are provided by the ECMWF ERA-5 analyses. The chemical mechanism and deposition scheme have been recently updated (Müller et al., 2018, 2019). Anthropogenic NMVOC emissions are provided by the EDGAR 4.3.2 inventory (Huang et al., 2017) for the year 2012. Biomass burning emissions are obtained from the Global Fire Emission Database version 4 (GFED4s) (Van Der Werf et al., 2017). The emissions of isoprene and monoterpenes are calculated using the MEGAN-MOHYCAN model (Guenther et al., 2012; Müller et al., 2008). The model also incorporates biogenic emissions of methanol, methyl-butenol, ethylene, ethanol, acetaldehyde, formaldehyde and acetone, as well as oceanic emissions of methanol, acetone, acetaldehyde and alkyl nitrates (Müller et al., 2019). The global source of glyoxal in the model amounts to 47 Tg/yr (in 2013), 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 atmospheric degradation of isoprene, acetylene, aromatics and monoterpenes, respectively (Müller et al., 2019).

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

3.3. Background correction

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

Smedt et al., 2018). The principle of this background correction is to add offset values to the retrieved SCDs to ensure that the resulting mean VCD in a clean remote region match an a priori known tropospheric glyoxal column. Here we use the Pacific Ocean as reference sector with a constant reference VCD of $1x10^{14}$ molec/cm^2. This value was chosen according to independent measurements performed in this region (Sinreich et al., 2010) since current global models fail to reproduce remote sensing glyoxal levels observed over oceans (Fu et al., 2008; Myriokefalitakis et al., 2008; Stavrakou et al., 2009b). There is nevertheless an uncertainty related to this reference value, which impacts the overall level of the product. This error component is further discussed in section 3.4.3 and is taken into account to estimate the total glyoxal VCD error. As all intermediate variables (SCD, corrected-SCD, AMF) are provided in the product, a user could recompute glyoxal VCDs using a different reference Pacific value.

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

1. First, a destriping procedure such as proposed in Boersma et al. (2007) is applied consisting in an offset correction determined separately for each instrumental row, and relying on clear sky observations from the Equatorial Pacific Ocean (15°S-15°N, 165°E-220°E). The offset corrections are added to all glyoxal SCDs worldwide, considering their respective row.

2. Additionally to the high frequency stripes, a broadband row-dependent structure, of which the shape also depends on the latitude, was identified as illustrated in Figure 5, panel (a). This figure compares the row-dependence of mean uncorrected VCDs in the Pacific Ocean at Equatorial and Northern mid-latitudes. The two curves are somehow anti-correlated, meaning that the destriping correction based on equatorial latitudes only as applied in step 1 is not sufficient and even reinforces the mid-latitude structure. The second step of the background correction aims thus at reducing this broadband row-dependent structure at all latitudes while maintaining the mean latitudinal distribution of the measured background glyoxal columns. For this, Pacific Ocean measurements (40°S-40°N, 165°E-220°E) are binned per 20° in latitude and in groups of 15 rows in a 2-dimensional matrix. For this step, we use reference VCDs depending on the latitude and resulting from the averaging of the binned VCDs along the row dimension. A corresponding 2-dimensional matrix of SCD offset corrections is then computed in order, once applied to the binned VCDs, the corrected values match the reference VCDs. Interpolation through this correction matrix provides offsets to be applied to all SCDs retrieved worldwide. For satellite pixels out of the latitude range, no extrapolation is performed but instead, the nearest neighbour correction values are taken.

3. Finally, the overall level of the product is adjusted with a single offset correction to ensure that the mean of all clear-sky VCDs within the full reference sector (40°S-40°N, 165°E-220°E) is equal to $1x10^{14}$ molec/cm^2. Panel (b) of Figure 5 shows how the identified row dependence in the VCDs at different latitudes has been reduced. The general level of the columns has also been adjusted.

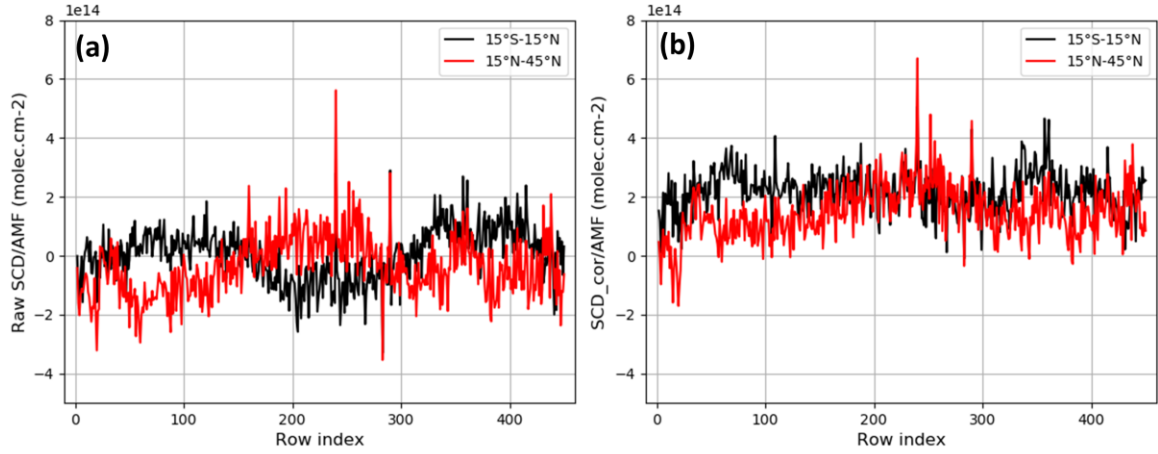



Figure 5 : Row-dependence of the glyoxal vertical columns of S5p orbit #5877 (December, 1st 2018) averaged in an equatorial latitude band and in a Northern mid-latitude band. (a) No background correction is applied; (b) a latitude-dependent background correction is applied.

3.4. Uncertainty estimates

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

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

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 $N_{s,0}$, M_0, $N_{v,0,ref}$ are the slant columns, AMF, and the reference vertical column within the reference sector, the total glyoxal vertical column error can be written as

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

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

422 may lead to conservative systematic error estimates and to an underestimation of the product scatter, depending

423 on the time and spatial resolution of interest. In particular, uncertainties associated to the input parameters needed

424 for the AMF calculation are directly related to the resolution of the used databases and may appear as random at

425 coarser resolution. This has been discussed by Vigouroux et al. (2020) who attributed part of the scatter in

426 formaldehyde vertical column TROPOMI/MAX-DOAS differences to a random component of the AMF errors.

3.4.1. Slant column uncertainties

428 As mentioned above, the radiance measurement noise directly propagates into the glyoxal slant column retrieval

429 and leads to large random errors $\sigma_{N_{s,rand}}$ (or precision) due to the low glyoxal optical depth. Those are easily

430 estimated using the fit residuals RMS and the covariance matrix of the cross-sections included in the fit (Danckaert

431 et al., 2017). In the visible spectral range, the TROPOMI signal-to-noise ratio is about 1600 over dark scenes.

432 This leads to a glyoxal VCD precision (i.e. $\sigma_{N_{s,rand}}$/AMF) in the range of 6-10x10^{14} molec/cm² as illustrated in

433 Figure 6, panel (d). This range of values is consistent with the scatter observed in the retrieved glyoxal SCDs in

434 regions without any significant glyoxal source. Over bright scenes, for example covered by clouds or snow, those

435 errors significantly drop because of the increased signal-to-noise ratio. For individual observations, random errors

436 dominate and averaging is needed to extract meaningful glyoxal signals.

437 There are also systematic errors associated to the DOAS spectral fit that are mainly dominated by absorption

438 cross-section uncertainties, by interferences with other species (O$_4$, liquid water, Ring …), or by other effects such

439 as residual stray light. Those contributions are difficult to assess and can only be estimated from sensitivity tests

440 (Lerot et al., 2010). In general, this error term can be as high as $2–3\times10^{14}$ molec/cm². However, the use of a

441 radiance as reference in the DOAS fit and the application of a background correction removes a large part of the

442 systematic error in the slant column fit (see section 3). As those corrections are not always sufficient to eliminate

443 completely the SCD systematic errors due to local conditions (local pollution, residual clouds,…), we set $\sigma_{N_{s,syst}}$

444 to 1×10^{14} molec/cm².

3.4.2. AMF uncertainties

446 The errors on the air mass factor depend on the input parameter uncertainties and on the sensitivity of the air mass

447 factor to each of them. This contribution can be broken down into the squared sum (Boersma et al., 2011; Lerot

448 et al., 2010; De Smedt et al., 2018) as

$$\sigma_{M,syst}^2 = \left(\frac{\partial M}{\partial A_s} \cdot \sigma_{A_s}\right)^2 + \left(\frac{\partial M}{\partial S} \cdot \sigma_S\right)^2 + (0.15M)^2 \tag{3}$$

449 where σ_{A_s} and σ_S are typical uncertainties on the surface albedo and profile shape, respectively.

450 The contribution of each parameter to the total air mass factor error depends on the observation conditions.

451 Therefore, a small table of air mass factor derivatives spanning all observation conditions was computed using

452 VLIDORT, considering glyoxal box profile shapes with different effective heights.

453 The AMF error component related to the surface reflectivity (1st term of Eq. (3)) is calculated using an estimated

454 uncertainty on the albedo σ_{A_s} of 0.02 (Kleipool et al., 2008). Note that this uncertainty can be occasionally larger,

455 in particular at high latitudes where snow falls may cause abrupt changes in scene albedo. The uncertainty

associated to the a priori profile shapes (the smoothing error) used in the retrieval is more difficult to assess, especially due to the scarcity of independent glyoxal profile measurements. For every observation, an effective height corresponding to the a priori glyoxal profile used in the AMF calculation is derived and used to extract the appropriate AMF derivative and σ_S is taken equal to 50hPa. The latter value corresponds to the typical standard deviation of the model profile shape effective heights over polluted regions.

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

3.4.3. Background correction uncertainties

Although the background correction is designed to overcome systematic features/deficiencies of the slant column fitting, some errors are also associated to this procedure. In particular, systematic errors on the reference slant columns and their air mass factors are propagated into the computed correction values. Also, there is an uncertainty related to the reference glyoxal vertical column value in the reference sector as mentioned in section 3.3. The three last terms of Eq. (2) represent the total background correction uncertainty in which $\sigma_{N_{s,0}}$ is the systematic slant column error fixed to 1×10^{14} molec/cm² (see above section 3.4.1), and M_0 and σ_{M_0} are the air mass factors and their associated errors within the reference sector. In practice, those quantities are treated similarly as the reference slant columns (i.e. binned in latitude and row bins – see section 3.3). $\sigma_{N_{v,0,ref}}$ represents the error associated to the reference value $N_{v,0,ref}$ and is fixed to 5×10^{13} molec/cm². Using a different reference value would directly impact the overall level of glyoxal VCDs worldwide, with some small modulations related to the ratio of the AMFs over Pacific and in other regions following Eq. (1).

3.4.4. Total systematic uncertainties

Figure 6, panel (c) shows the estimated mean VCD systematic errors for the month of June 2018 when all systematic error sources are combined together using Eq. (2). Note that the conversion of the AMF error into an absolute vertical column error (2nd term of the equation) requires this error to be multiplied by the corresponding vertical column. Because of the high level of noise in the product, using the retrieved column for this would lead to a strong overestimation of the systematic error. To circumvent this, we use instead pre-computed climatological glyoxal noise-free VCDs.

Total glyoxal VCD systematic errors are generally in the range $1-3\times10^{14}$ molec/cm², corresponding to about 30-70% for emission regimes (columns larger than 2×10^{14} molec/cm²). Figure 7 shows the zonally averaged total systematic error along with its different components for one S5p orbit passing over Africa. In general, the three components contribute similarly to the total error for emission conditions. On contrary, the AMF error becomes smaller in background conditions while the two other terms dominate. Note that pixels strongly contaminated by

clouds (cloud fraction > 20%) or covered by snow/ice are discarded. Systematic errors are expected to be large for those pixels mainly due to spectral interference effects (see section 3.2) and also because the information content on glyoxal is reduced in case of cloud shielding. Figure 6, panel (b) shows monthly mean AMFs for the same month. Small AMFs are generally caused by a priori profiles peaking near the surface, which makes the retrieval more sensitive to albedo uncertainties and to a lesser extent to the a priori profile shape uncertainties. This explains the anti-correlation between the AMFs and the systematic errors. In contrast, large AMFs are caused either by bright surface or by background a priori profiles. For such cases, systematic errors are smaller. Note that satellite column averaging kernels, defined as the Box-AMF divided by the total AMF (Eskes and Boersma, 2003), are provided for every observation. They can be used to remove the smoothing error component when comparing the satellite data to any other external data.

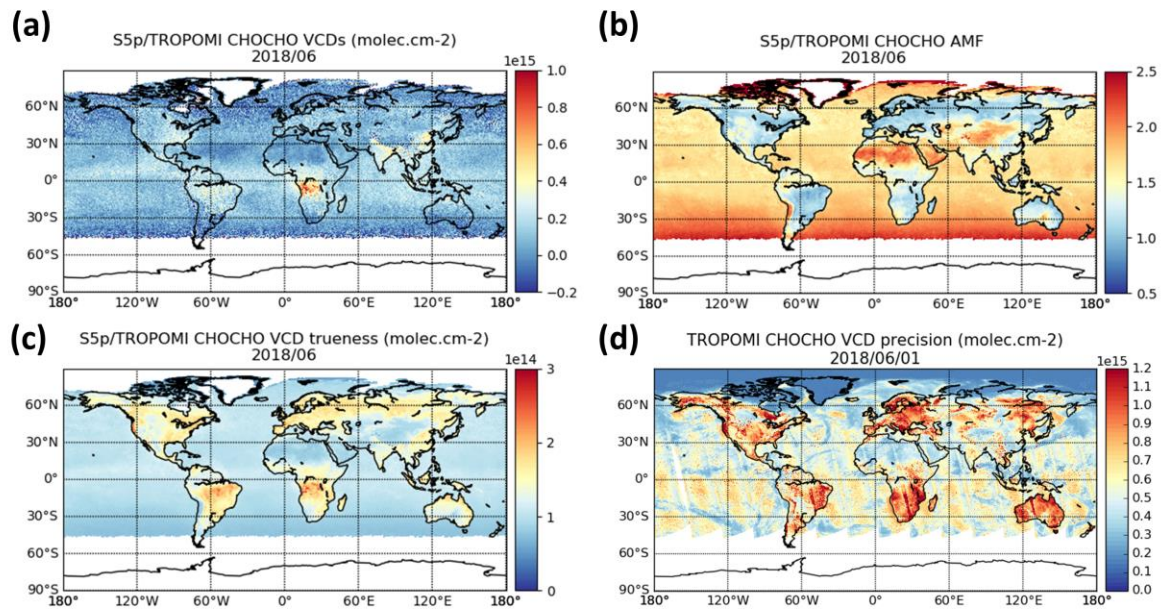



Figure 6: (a) TROPOMI June 2018 monthly means of glyoxal tropospheric columns, glyoxal air mass factors (panel (b)) and glyoxal tropospheric column systematic errors (panel (c)). Scenes contaminated by clouds or Ice/snow have been filtered out. Panel (d) shows glyoxal tropospheric column random errors for one single day, in which all observations have been kept to illustrate the impact of the scene brightness.

508

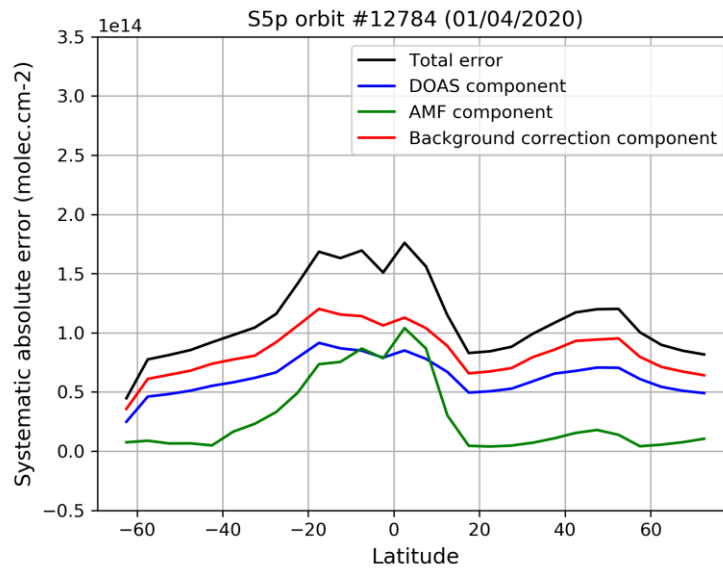



Figure 7 : Zonal mean of the total glyoxal VCD systematic error along with its different components for one single S5p orbit passing over Africa on April, 1st 2020.

509
510

4.   Comparison with other satellite instruments

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

511
512

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

513
514
515
516
517

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

518
519
520
521
522
523
524
525
526
527
528

Differences in air mass factor calculations consist only in using, over land, a priori profiles provided by IMAGESv2, the chemical transport model predecessor of MAGRITTE, at the coarser resolution of 2.0°x2.5°. For the GOME-2 instruments, we use the directionally dependent Lambertian-equivalent reflectivity database produced by Tilstra et al. (2021) instead of the OMI database.

529
530
531
532

533   A background correction procedure is applied consistently with the one used for TROPOMI. The GOME-2

534   instruments being whiskbroom scanners, there is no destriping procedure as such but instead a viewing zenith

535   angle-dependent correction is applied, also relying on the slant columns retrieved in the Equatorial Pacific sector.

536   This correction may account for example for remaining biases related to the instrumental polarization sensitivity.

537   For both OMI and GOME-2, the row/VZA dependence does not show any obvious change along the orbit and the

538   corresponding correction thus relies only on the low latitude measurements.

539   Note that the OMI and GOME-2 glyoxal products are filtered for cloudy scenes using cloud fraction lower than

540   20% as taken from the O_2-O_2 (Veefkind et al., 2016) and OCRA (Lutz et al., 2016) cloud products, respectively.

541   The empirical correction for strong NO_2 absorption signature described in section 3.1 has been applied to those

542   instruments as well. In the following section, we compare the TROPOMI glyoxal retrievals with the OMI and

543   GOME-2A/B data sets. The OMI record covers the period 2005-2018, while GOME-2A/B span the periods 2007-

544   2017 and 2013-2020, respectively. OMI and GOME-2A records were interrupted when their respective quality

545   was degraded too severely and other instruments were available to continue the morning and afternoon time series.

4.2. Glyoxal satellite inter-comparison

4.2.1. Comparison of the noise level

548   As mentioned before, the level of noise in the satellite glyoxal tropospheric column products is large compared to

549   the real signal. This is illustrated in the upper panel of Figure 8 which shows all individual clear-sky TROPOMI

550   glyoxal columns retrieved in the Pacific Ocean on June, 1st 2018 and plotted as a function of their latitude. The

551   scatter is significant ($\sigma \approx$ 5-7x10^{14} molec/cm^2) with respect to the small glyoxal VCDs averaged in 5°-latitude

552   bins in this sector. The lower panel compares the standard deviation of the retrievals from TROPOMI, OMI,

553   GOME-2A and B in the same remote sector for the 1st of June of their respective first year of operation. The

554   scatter in the retrievals is directly linked to the instrumental signal-to-noise ratio, which is documented to be

555   around 500 for OMI (Schenkeveld et al., 2017), 1000 for GOME-2 (Zara et al., 2018) and 1500 for TROPOMI

556   (Kleipool et al., 2018). In practice, we see that the CHOCHO observation noise is indeed slightly larger for OMI,

557   that GOME-2B retrievals are noisier than those from GOME-2A, which have a level of noise similar to

558   TROPOMI. Considering the very small footprint size of TROPOMI (3x7.5 km² and 3x5.5 km² after August 2019)

559   compared to the other instruments (GOME-2: 80x40 km²; OMI: 13x24 km² at nadir), the TROPOMI observation

560   noise is remarkably low. More importantly, the much larger amount of TROPOMI data compared to OMI (~15x)

561   and GOME-2 (~100x) allows maintaining a better time or spatial resolution for a given target noise level. For

562   example, the random error associated to the daily glyoxal column averaged in an area defined by a circle with a

563   radius of 50 km will be less than 0.5x10^{14} molec.cm^{-2} for TROPOMI, while it will remain larger than 2.5x10^{14}

564   molec/cm²and 4.0x10^{14} molec/cm² for OMI and GOME-2, respectively.

565   This is illustrated in Figure 9, which compares January monthly mean glyoxal VCD fields over Asia at the

566   resolution of 0.05° for TROPOMI and OMI (upper panels) and 0.25° for GOME-2A and OMI (lower panels) after

567   one year of their respective operation. At the resolution of 0.05°, the level of noise in the TROPOMI glyoxal map

568   is very low and many details can be distinguished in the glyoxal spatial distribution. In particular, hot spots of

569   glyoxal over many megacities are clearly identified (e.g. over Bangkok, New Delhi, Ho Chi Minh City,

Shenzhen…) but also over Cambodia where large fires occur every year from January to March. At this resolution of 0.05°, the level of noise in the OMI map remains high and prevents distinguishing such details. At the coarser spatial resolution of 0.25°, the reduction of the noise in the OMI and GOME-2 monthly glyoxal fields appears to be sufficient to better distinguish the glyoxal spatial distribution but at the cost of a significant smoothing. In the next section, we will intercompare the four satellite products at low temporal and spatial resolution in order to minimize the impact of the noise and to identify possible systematic discrepancies.

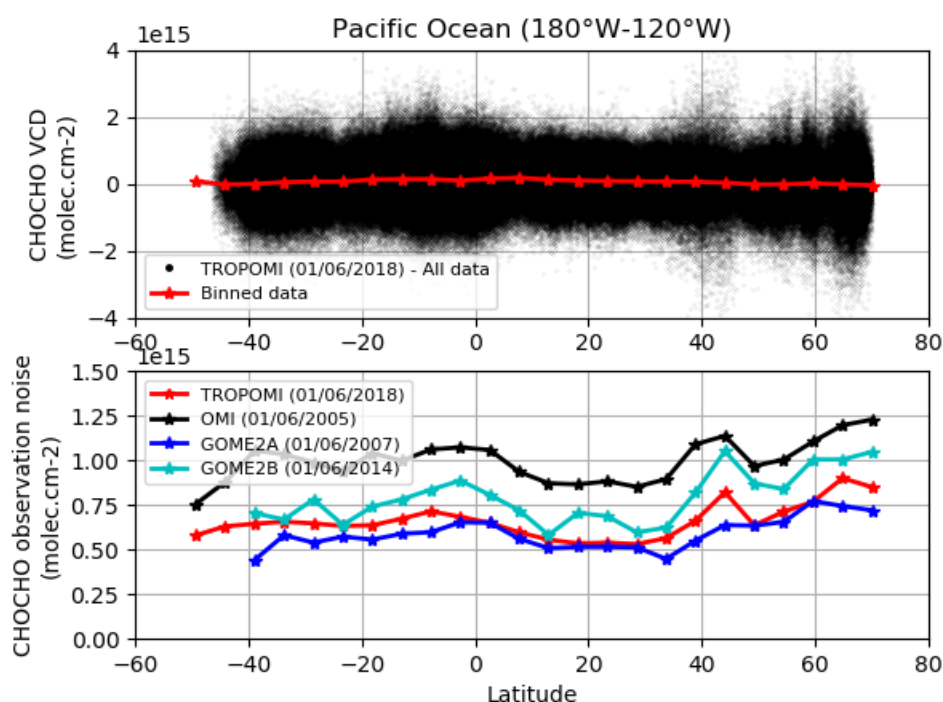



Figure 8 : Illustration of the level of noise in satellite CHOCHO VCD retrievals. The upper panel shows all clear sky individual glyoxal VCDs retrieved in the Pacific Ocean from TROPOMI observations on June, 1st 2018. The scatter is very large compared to the low real signal as illustrated by the data binned in 5° latitude bands. The lower panel compares the standard deviation of the retrievals in the same sector from TROPOMI, OMI, and GOME2A/B on the 1st of June of their respective first year of operation.

583

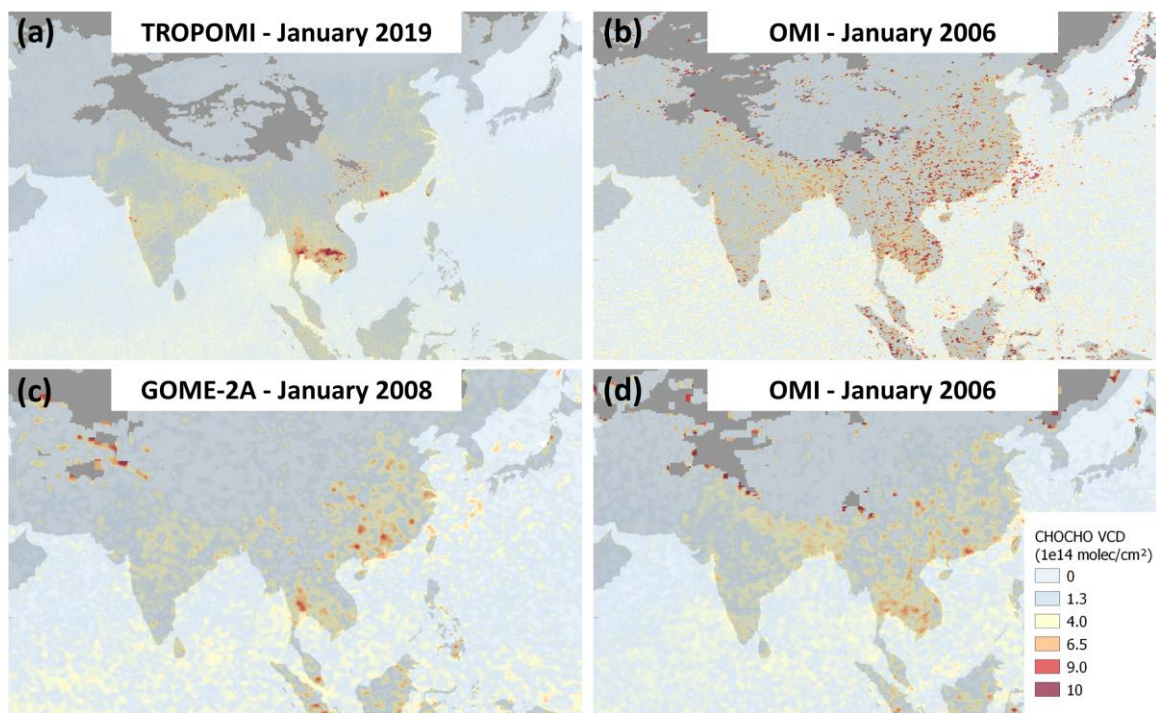



Figure 9: Illustration of the impact of the instrumental signal-to-noise and available amount of data on monthly mean glyoxal VCD fields retrieved from different satellite instruments. The (a) TROPOMI data for January 2019 gridded at a resolution of 0.05°, (b) OMI data for January 2006 gridded at a resolution of 0.05°, (c) GOME-2A data for January 2008 gridded at a resolution of 0.25°, and (d) OMI data for January 2006 gridded at a resolution of 0.25°. Cloudy scenes have been filtered out and a smoothing filter has been applied on the four presented fields based on a spatial mean with the nearest neighbouring grid cells.

4.2.2. Comparison of mean glyoxal fields

First, Figure 10 and Figure 11 compare seasonal maps of glyoxal VCDs generated from TROPOMI, OMI and GOME-2A/B data products. In order to reduce the data scatter for each instrument, those maps are based on long time series as indicated in the figures. Therefore, a one-to-one match is not expected. As can be seen, the consistency between the four instruments is excellent. Glyoxal patterns are captured similarly for all seasons in terms of both spatial distribution and VCD values. The largest glyoxal columns are observed in tropical regions, where biogenic emissions are important, and in regions with important fire events (e.g. Amazonia and Northern Africa in SON, Thailand/Indochina in MAM, Western US in August,…). At mid-latitudes, the glyoxal columns follow the seasonal cycle of biogenic activity with maximum values during summertime. Localized hot spots of glyoxal are visible over megacities corresponding to strong anthropogenic emissions (e.g. Northern China Plain, Bangkok, Teheran, New Delhi, Sao Paulo…).

In contrast to TROPOMI and OMI, the level of noise in the GOME-2 data sets significantly increases over the South Atlantic Anomaly despite the application of a spike-removal procedure (section 3.1). Overall the GOME-2B maps are 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 the spectral signature of specific ground surfaces, and thus may potentially lead to a bias on the

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

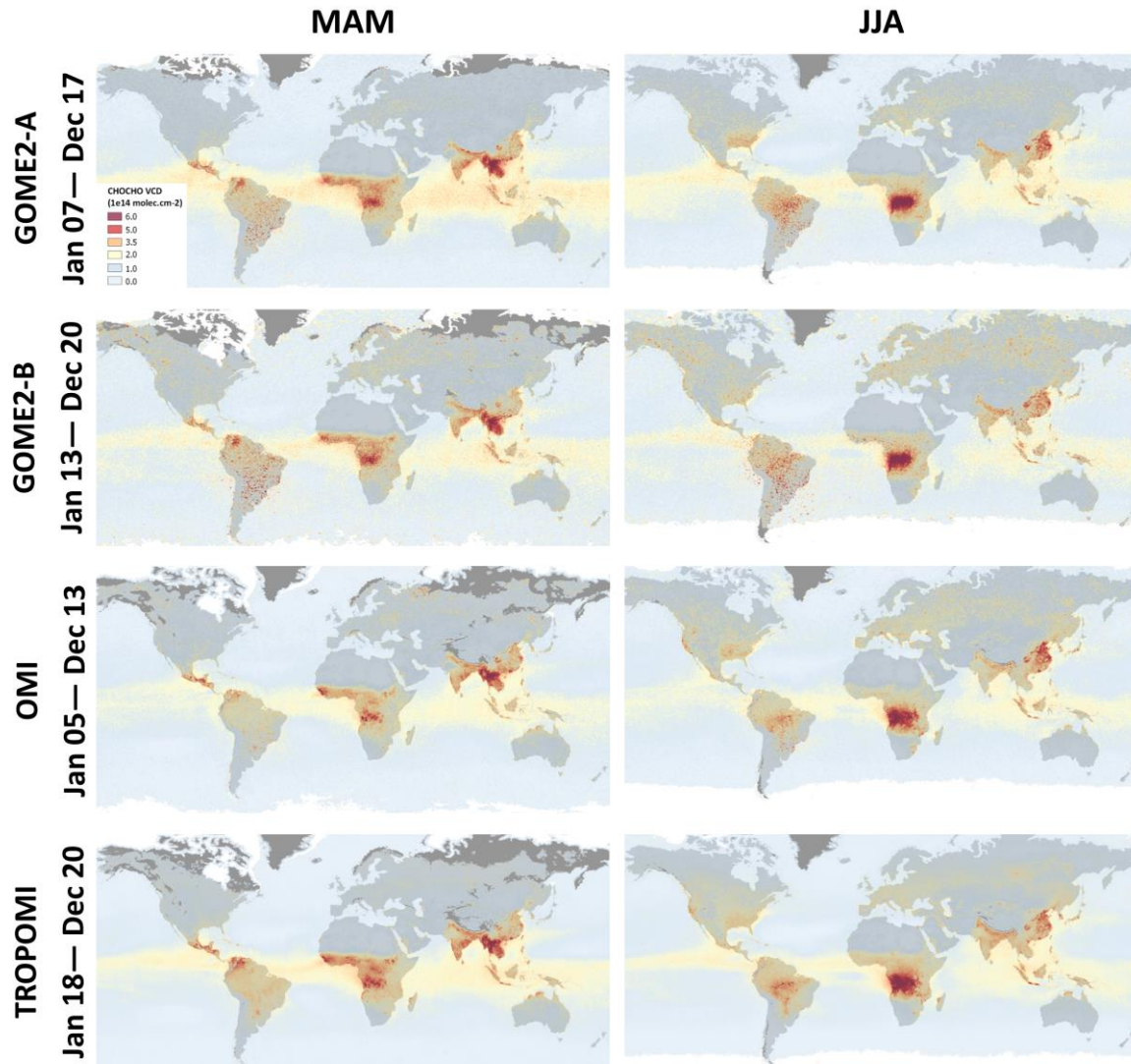



Figure 10: Comparison of long-term averaged global CHOCHO VCDs (in 10^{14} molec/cm²) derived from GOME-2A, GOME-2B, OMI and TROPOMI sensors, for the March-April-May period (left panels) and the June-July-August period (right panels).

614

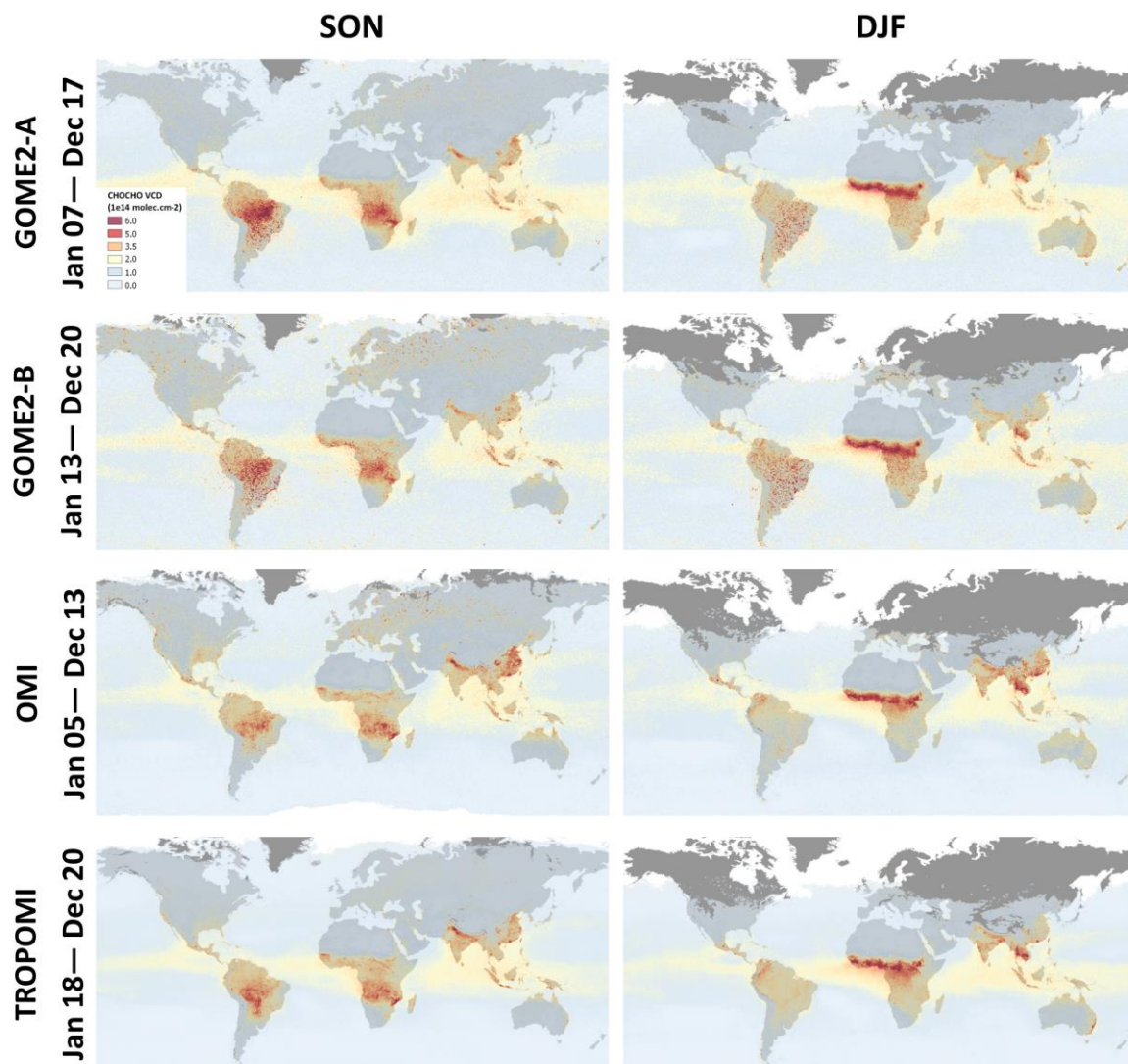



Figure 11: Comparison of long-term averaged global CHOCHO VCDs (in 10^{14} molec/cm²) derived from GOME-2A, GOME-2B, OMI and TROPOMI sensors, for the September-October-November period (left panels) and the December-January-February period (right panels).

For a more detailed investigation of the consistency of the TROPOMI data set with OMI and GOME-2A/B, we compare complete time-series of monthly median glyoxal columns in selected regions (shown in Figure 12). The red rectangles indicate the regions on which we focus in Figure 13 and Figure 14, while the global statistics for all highlighted regions are given in Figure 15. Detailed figures are provided for all regions as supplementary material (Figures S1, S2, S3, S4). Figure 13 compares directly the four full time series, while Figure 14 compares the typical climatological seasonal variations as obtained by combining all available years. The error bars in the latter figure represent the interannual variability, and the 2-sigma standard deviation of the four satellite products is indicated as inset.

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

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

In Asia, there are many hot spots, of which the origin is manifolds and strongly depends on the region and season. In addition to biogenic activities, large emissions due to fires may significantly contribute to the glyoxal columns. As illustrated in Figure 14, in the Indo-Gangetic Plain, there are typically two fire seasons in April/May and in October/November (after the Monsoon period) related to agricultural burning of wheat residue (Kumar et al., 2016), and leading to two maxima in the glyoxal VCD seasonal cycle with a significant interannual variability. For example, during the COVID-19 Indian lockdown in April/May 2020, fire activity has been reduced leading to smaller emissions (Levelt et al., 2021). This region is also highly populated, causing large emissions due to human activities. This is also true in North-East China where glyoxal columns remain significant in winter, while biogenic emissions are low during that period of the year. Although less variable than fire emissions, anthropogenic emissions may also change over time. Despite those variable emissions, the four data sets spanning different time periods show a high level of consistency. In China, it seems that the glyoxal columns as observed by OMI, GOME-2A and B are slightly reduced after 2014. This would deserve further investigation. On the other hand, any interpretation based on long-series of OMI data must be treated carefully since the instrument suffers from an evolving row anomaly (Schenkeveld et al., 2017), which impacts the stability of the product and causes an increasing number of outliers, especially at mid-latitudes. For example, over the Indo-Gangetic Plain, the OMI columns deviate regularly from the other instruments after 2014. In general, remnants of noise are also visible in the GOME-2 time series, which show somewhat less smooth time series than TROPOMI.

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

659

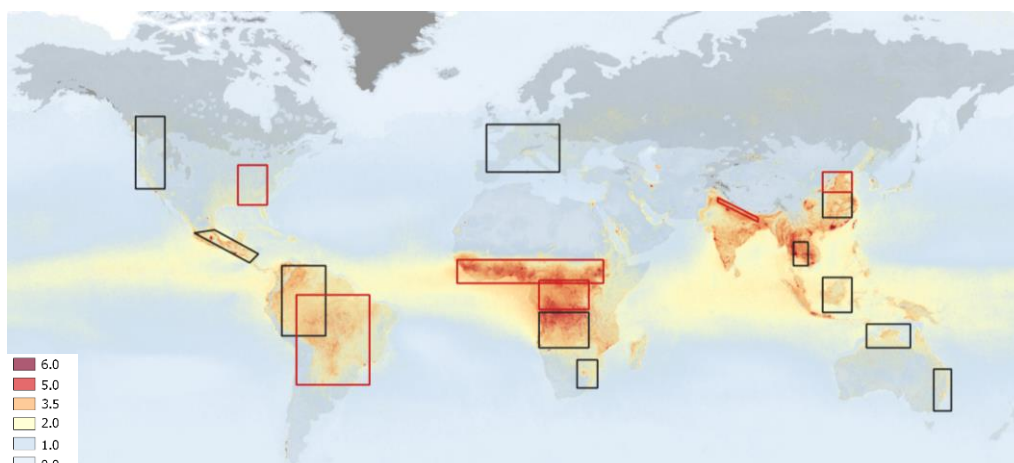



Figure 12 : TROPOMI glyoxal VCD distribution (in 1e14 molec/cm²) averaged on the period January 2018-December 2020. The rectangles represent the regions where the glyoxal products from different satellites are intercompared with a specific focus on red regions in Figure 13 and Figure 14.

663

664

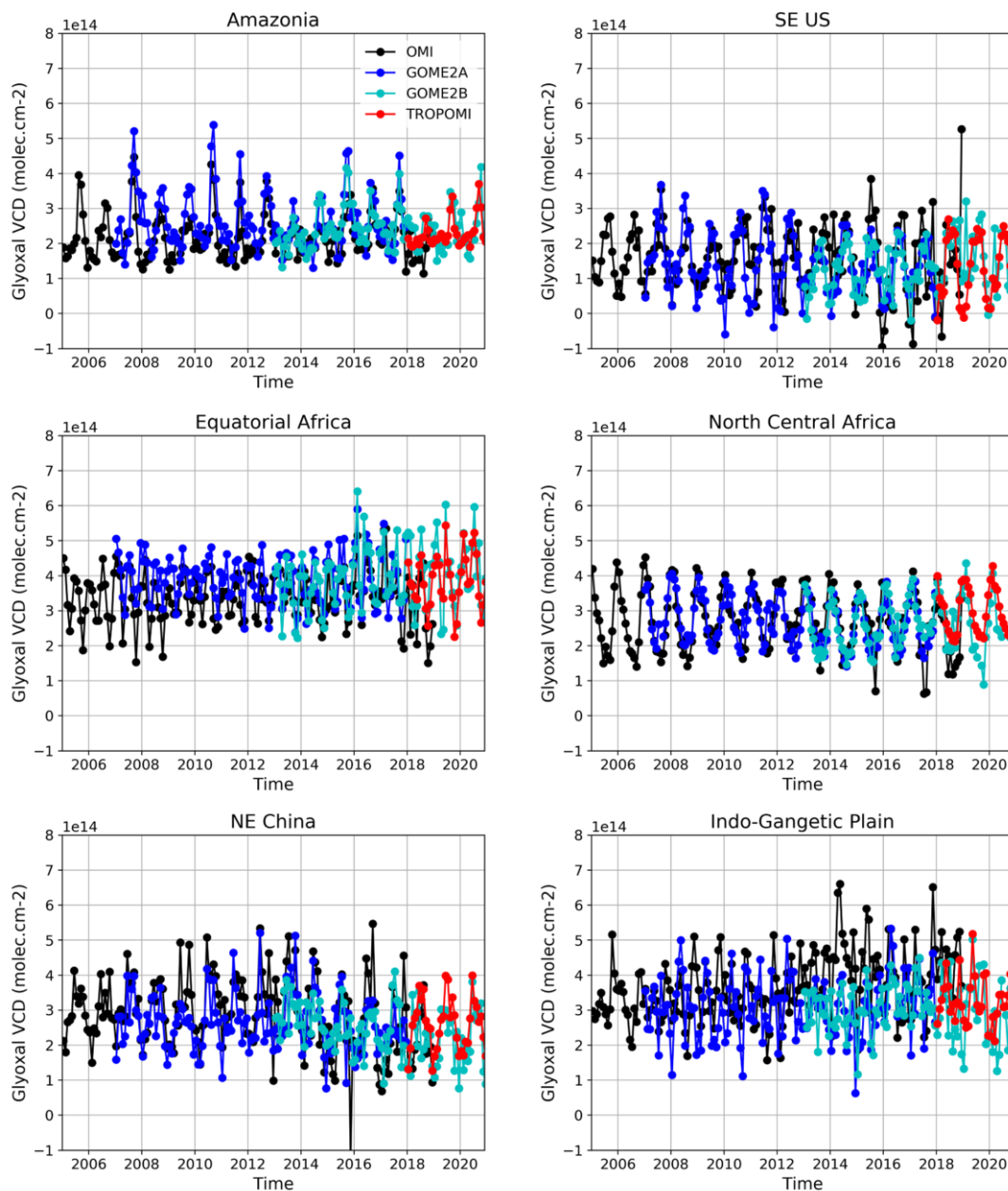



Figure 13: Comparison of the monthly median glyoxal VCD time series from GOME-2A/B, OMI and TROPOMI in a few selected regions worldwide.

668

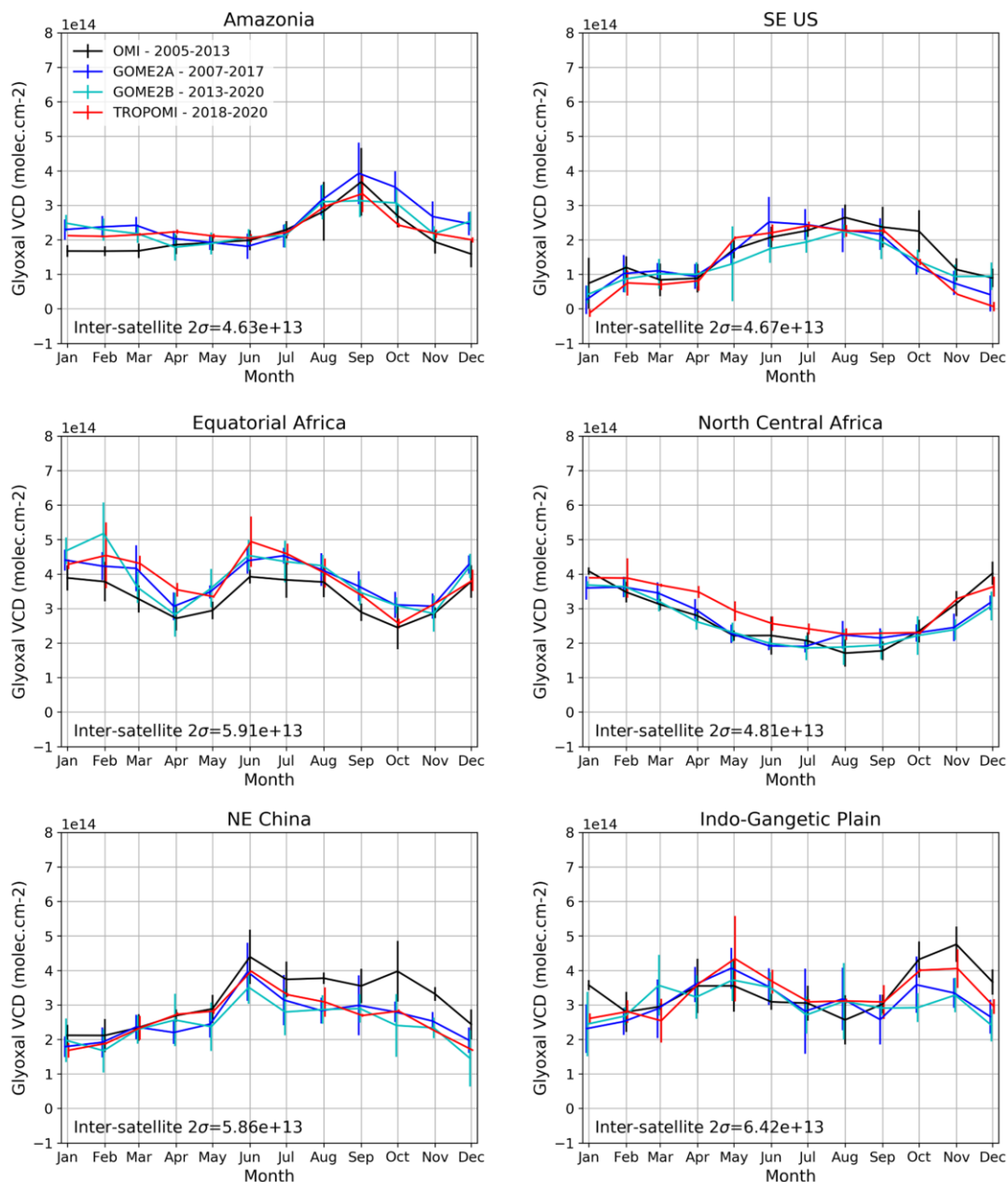



Figure 14: Comparison of the climatological seasonal variation of the monthly median glyoxal VCDs from GOME-2A/B, OMI and TROPOMI in a few selected regions worldwide. The error bars represent the interannual variability as derived from the full time series.

Figure 15 summarizes for all regions drawn in Figure 12 the absolute and relative deviation of each of the four data sets with respect to the median values of the ensemble. The symbols represent the median deviation considering all months of the year, while the error bars represent the full range of the monthly deviations. Regions are sorted by increasing mean glyoxal vertical column amounts and light and dark blue shaded areas indicate 2.5×10^{14} molec/cm² (50%) and 1.5×10^{14} molec/cm² (30%) differences as guidelines. Inter-satellite deviations are generally less than 5×10^{13} molec/cm² (20%). The large error bars in the relative differences plot for mid-latitude regions are caused by local wintertime months during which the glyoxal content is very low, if not negligible, and are therefore meaningless. Overall, the inter-satellite consistency of the glyoxal VCD products is excellent. In the section 5, we

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

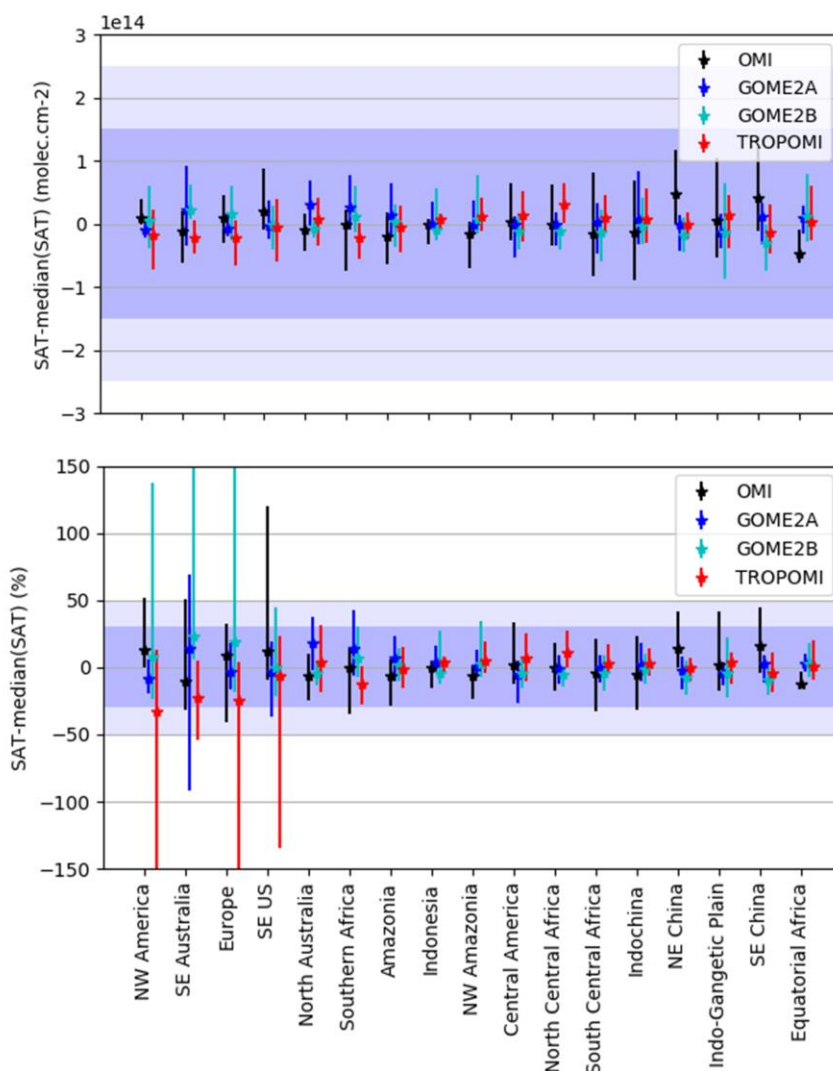



Figure 15: Median deviation of the glyoxal VCD differences for TROPOMI, OMI, GOME-2A/B against the median value of the ensemble of the four data sets in the selected regions worldwide drawn in Figure 12. Those are plotted in absolute values (molec/cm²) in the upper panel and in relative values (%) in the lower panel. The error bars indicate the full range of the deviations considering climatological monthly data. Regions are sorted by increasing median glyoxal VCD value from left to right. The light and dark blue shaded area indicate differences of 1.5 molec/cm² (30%) and 2.5 molec/cm² (50%).

4.2.3. Glyoxal over equatorial oceans

A persistent equatorial oceanic glyoxal signal is seen consistently by the four sensors. The origin and the magnitude of the enhanced glyoxal concentrations over oceans remains nevertheless unclear. A similar feature has been observed from space in previous studies (Lerot et al., 2010; Vrekoussis et al., 2009; Wittrock et al.,

694   2006), while it was much less pronounced in others (Alvarado et al., 2014; Chan Miller et al., 2014). Over the
695   past years, glyoxal measurements have been realized with ship-borne MAX-DOAS. While Sinreich et al. (2010)
696   measured glyoxal concentrations up to 100 ppt in the marine boundary layer of the Equatorial Pacific Ocean, most
697   other studies (Behrens et al., 2019b; Lawson et al., 2015; Mahajan et al., 2014; Volkamer et al., 2015) reported
698   lower concentrations inconsistent with the satellite elevated glyoxal columns. However, Volkamer et al. (2015)
699   also reported elevated glyoxal concentrations measured with an airborne MAX-DOAS in the free troposphere,
700   which might explain the larger satellite glyoxal signal. Remaining spectral interferences may also contribute, at
701   least partly, to this signal. In particular, its spatial correlation with high water vapour concentration regions and
702   the high sensitivity of glyoxal retrievals to the water vapour cross-section as discussed in section 3.1 and by Chan
703   Miller et al. (2014) call for a careful assessment of any future new data release or for future investigation on fit
704   strategy to mitigate this interference (e.g. Kluge et al., 2020).

705

5.   Validation with MAX-DOAS data

5.1.   Description of MAX-DOAS data sets and methodology

708

709   MAX-DOAS instruments measure scattered solar light in the UV-Visible spectral range at different elevation
710   angles above the horizon and allow retrieving information on trace gases and aerosol extinction in the altitude
711   range below 2-3km of the atmosphere, where the instrumental sensitivity is the highest. In a first approximation,
712   vertical columns of boundary layer gases can be estimated from MAX-DOAS measurements using a simple
713   geometrical approach (Brinksma et al., 2008; Hönninger et al., 2004). More elaborated approaches exploit a set
714   of different elevation angles to derive information on the vertical distribution of the gas concentration with up to
715   4 degrees of freedom, resulting in more accurate vertical columns in the 0-4 km altitude range (e.g. Beirle et al.,
716   2019; Clémer et al., 2010; Irie et al., 2011; Friedrich et al., 2019).

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

730   Among the available MAX-DOAS data sets, three (Xianghe/China, Chiba/Japan and Phimai/Thailand) are long
731   enough to allow a comparison with OMI and GOME-2A/B in addition to TROPOMI. The other ones span shorter,

732 and more recent periods and will be used only for comparison with the TROPOMI product. The Xianghe station
733 has the longest and stable data record, and provide vertical profiles of glyoxal. Therefore we have used this
734 reference station to perform a thorough analysis of the satellite product stability and of the impact of applying
735 satellite averaging kernels. For the other stations, we performed a more qualitative comparison of the seasonal
736 cycles of the glyoxal tropospheric columns. For the data colocation, we select MAX-DOAS data ±1.5 hour around
737 the satellite overpass time and satellite data within a radius of 100 km (150 km for Phimai) and 20 km around the
738 station for GOME-2A/B/OMI and TROPOMI, respectively. Daily median glyoxal columns are computed if both
739 satellite and ground-based data are available and finally monthly medians of the daily median columns are
740 compared.

741 **Table 3 : List of MAX-DOAS stations used in the study and brief description of the approach to generate the glyoxal**
742 **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 Phimai/Thailand (15.18°N, 140.10°E) 2014-2020 Pantnagar/India (29.03°N, 79.47°E) 2017-2020	CERES	Profile retrieved using a parametrization approach 436–457 nm	(Hoque et al., 2018; Irie et al., 2011)
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)

743

744

745 **5.2. Validation results**

746

747 Figure 16 focuses on the comparison of monthly median glyoxal tropospheric columns retrieved from TROPOMI,
748 OMI, GOME-2A and GOME-2B with columns from the BIRA-IASB MAX-DOAS instrument in Xianghe
749 (China). The left panels compare the full time series for each satellite sensor with the MAX-DOAS data record.
750 The right panels show the corresponding satellite/MAX-DOAS absolute differences. Note that the MAX-DOAS
751 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 satisfactory 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 TROPOMI time series also agrees well with the MAX-DOAS data 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 0.9×10^{14} molec/cm² , while it is larger than 1.7×10^{14} molec/cm² 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 -0.8×10^{14} molec/cm² for TROPOMI and -1.5×10^{14} molec/cm² 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 information. As shown in Figure 16, smoothing the MAX-DOAS columns reduces the satellite/MAX-DOAS bias to values ranging from -0.2×10^{14} molec/cm² (TROPOMI) to -0.5×10^{14} molec/cm² (OMI).

772

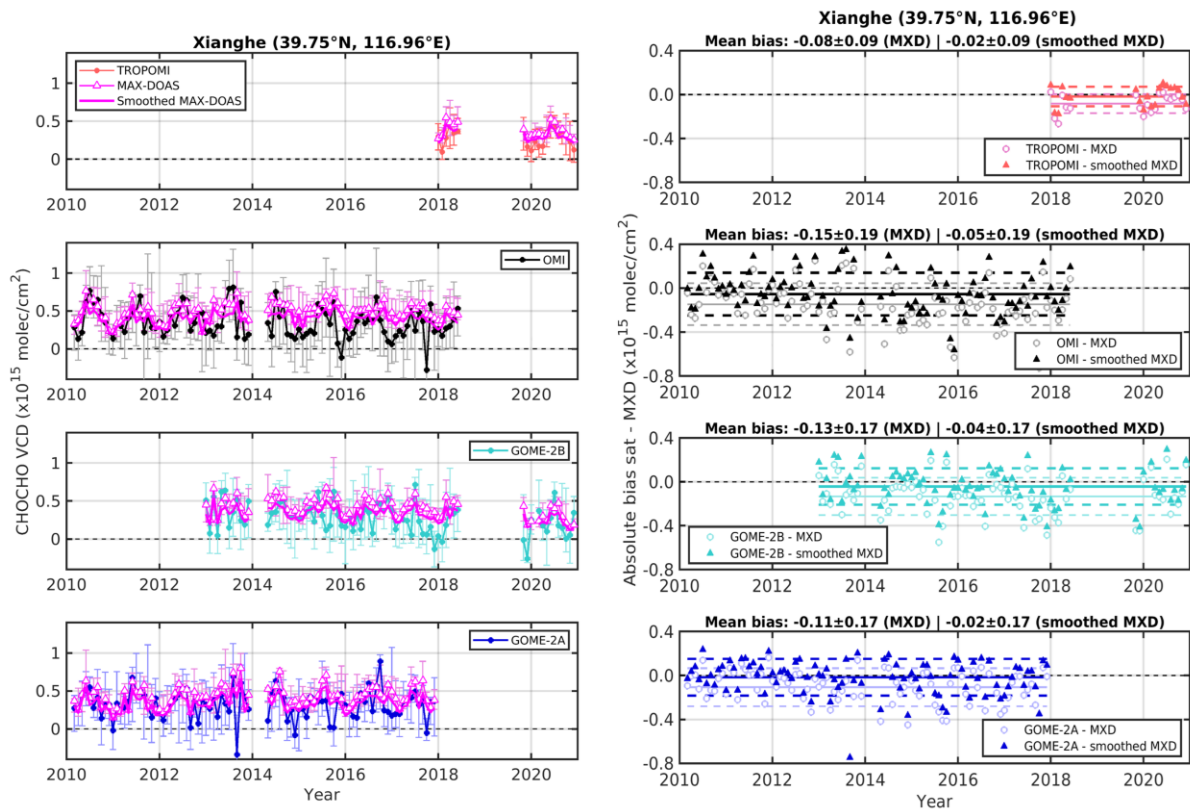



Figure 16 : Comparison of the monthly median glyoxal tropospheric vertical columns retrieved from satellite and MAX-DOAS (MXD) instruments in Xianghe (China). The four left panels compare the time series from TROPOMI, OMI and GOME-2A/B with the MXD time series. MXD columns are also shown when smoothed with the satellite averaging kernels. The error bars represent the 25 and 75% percentiles. The four right panels show the corresponding time series of the satellite-MD absolute differences. Both original and smoothed MXD data are shown. Mean bias and standard deviation of the differences are given (in 10^{15} molec/cm² units) in the panel titles and are also represented in the right panels with the full and dashed coloured lines.

In Figure 17, we compare the median satellite and MAX-DOAS seasonal cycles of the glyoxal tropospheric columns at three stations (Xianghe, Chiba and Phimai) where the time series present a good overlap with the OMI and GOME-2A and B records, in addition to TROPOMI. In Xianghe, the seasonal cycle of the smoothed MAX-DOAS columns is also shown, illustrating again the reduction of the satellite/MAX-DOAS bias when the a priori profile error component is removed. Note that the OMI and GOME-2B seasonal cycles are computed using data until end of 2013 and 2016 to limit the impact of the increasing number of outliers. In each comparison panel, the MAX-DOAS cycle is always computed using the same time range as the satellite instrument. Overall, the seasonal patterns are consistently captured by the satellite and MAX-DOAS instruments. In Xianghe, the GOME-2A and TROPOMI cycles follow closely the MAX-DOAS curves, although TROPOMI slightly underestimates the MAX-DOAS columns during winter months. OMI and GOME-2B also reproduce the general seasonal pattern but show a somewhat more scattered curve, likely due to their slightly less stable time series. In Chiba where the glyoxal signal is mostly driven by biogenic emissions, the agreement between the satellites and the MAX-DOAS measurements is excellent both in terms of variability and absolute values. Again, OMI shows a larger scatter (as also indicated by the larger error bars representing the inter-annual variability). In Phimai, where pyrogenic emissions are responsible for large glyoxal columns especially in the first few months of the year, the seasonal variability seen by the satellites and the MAX-DOAS is very consistent. A negative bias larger than for other stations is nevertheless observed. This can be related to other studies that identified larger biases in NO_2 or HCHO DOAS products for elevated column conditions (e.g. De Smedt et al., 2021; Verhoelst et al., 2021; Vigouroux et al., 2020). Possible causes for such biases are the different air masses probed by the satellite and ground-based instruments, their different vertical sensitivity as well as the a priori vertical profile information used in the retrieval algorithms.

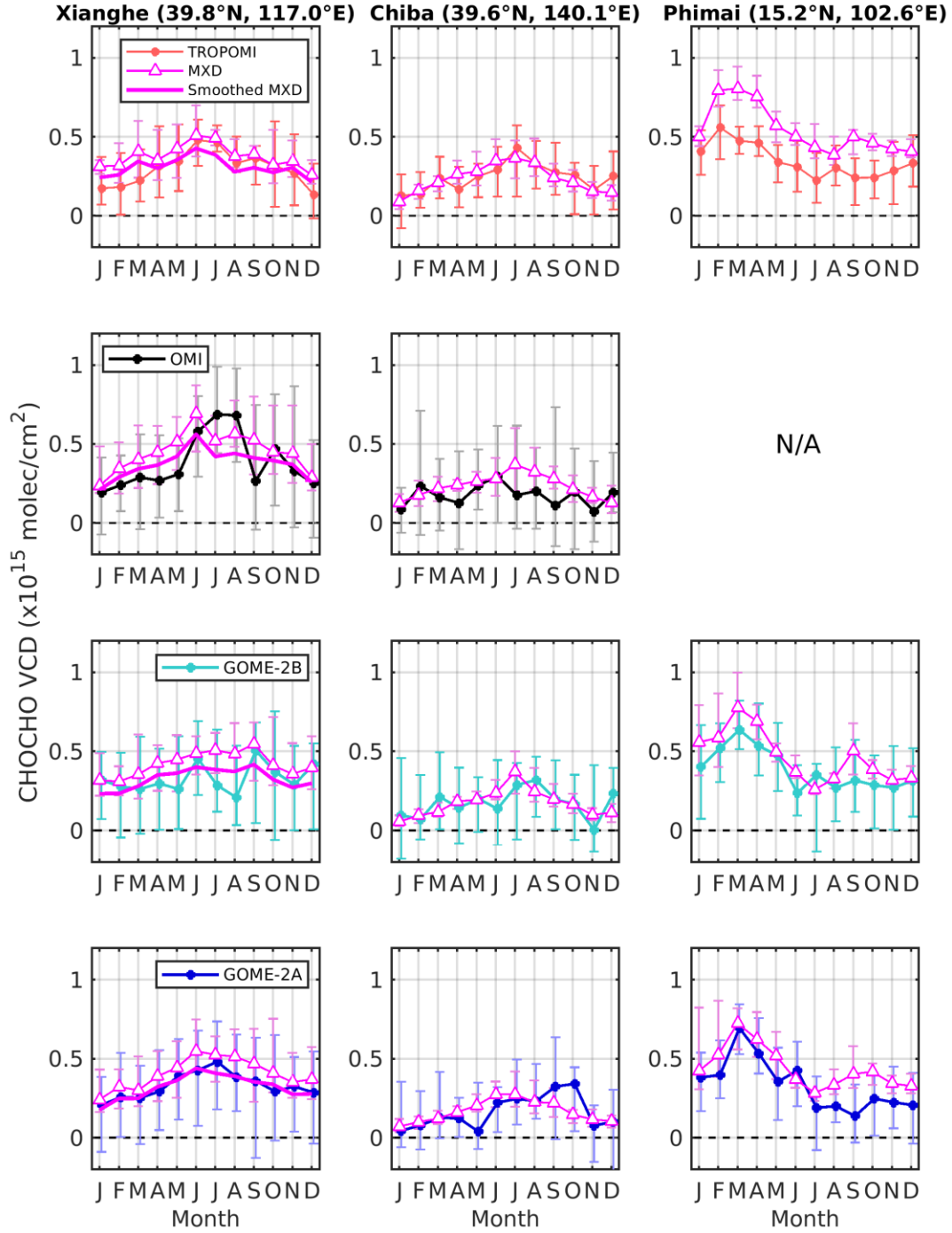



Figure 17 : Comparison of the monthly median glyoxal tropospheric vertical column seasonal cycle as retrieved from TROPOMI, OMI, GOME-2A/B and MXD in Xianghe (China), Chiba (Japan) and Phimai (Thailand). The columns correspond to the three stations and the rows to the different satellites. In Xianghe, MXD data smoothed with the satellite averaging kernels are also shown. The error bars represent the interannual variability (25% and 75% percentiles based on the full time series available). Note that the comparison of with the MAX-DOAS data in Phimai is not shown as the latter starts in 2014 when OMI is degraded.

809

810   In Figure 18, we compare again the seasonal cycle of glyoxal VCDs retrieved from TROPOMI with that from

811   more recent MAX-DOAS time series at six different stations. Four of them are located at mid-latitude in Europe

812   and show relatively low glyoxal columns, while larger average values are measured at the two other stations, in

813   Northern India (Mohali and Pantnagar). In Vienna/Austria and Athens/Greece, TROPOMI and MAX-DOAS

814   glyoxal columns agree very well and show consistent seasonal dependencies with maximum and minimum values

815   during summertime and wintertime, respectively.   On the other hand, at the higher latitude stations of

816   Bremen/Germany and Uccle/Belgium, the consistency of the seasonal variations seen from space and from the

817   ground is somewhat poorer. While the glyoxal columns agree well during summertime, the satellite columns tend

818   to underestimate MAX-DOAS values in winter, the latter showing almost no seasonal variation. Satellite glyoxal

819   retrievals at those latitudes are challenging in winter because of the low sun elevation causing a reduced sensitivity

820   to the lowermost atmospheric layers. As mentioned in section 3.2, observations with solar zenith angles larger

821   than 70° are filtered for this reason, which explains the gap between November and January at those two stations.

822   In Uccle, we have also tested the impact of smoothing the MAX-DOAS columns with the satellite averaging

823   kernels (similarly as for Xianghe), which turned out to be very small. The absence of any seasonal dependence in

824   the cities of Brussels (Uccle) and Bremen, in contrast to that observed (although limited) in Vienna and Athens,

825   is to some extent puzzling. One should keep in mind however that the glyoxal retrievals from MAX-DOAS

826   measurements are also challenging and it cannot be excluded that errors in ground-based data might also partly

827   contribute to the observed differences.

828   In Mohali and Pantnagar, glyoxal columns are much larger and the seasonal variability is driven by fire emissions

829   and meteorological factors such as the monsoon. At those two stations, the glyoxal seasonal variability is very

830   well reproduced by TROPOMI. In terms of absolute values, the TROPOMI columns agree reasonably well in

831   Mohali but, they significantly underestimate the (large) MAX-DOAS columns in Pantnagar. The reason why the

832   systematic satellite/ground-based bias is so different between those two stations is unclear. MAX-DOAS columns

833   are clearly higher in Pantnagar than in Mohali pointing either to possible local differences in air quality, not

834   reflected in the satellite data, or to inconsistencies in the ground-based data sets. Although the agreement of the

835   absolute columns is reasonable in Mohali, the typical behaviour is an underestimation of the columns by the

836   satellites, as discussed before. Note also that those sites are significantly contaminated by aerosols, which are

837   neglected in the satellite retrievals (apart from the stringent cloud filtering). MAX-DOAS data have also been

838   analysed using very different approaches, which may also cause differences. This calls for a more detailed

839   analysis, which would require an homogenization of the MAX-DOAS data treatment, a more sophisticated

840   approach for the computation of the satellite AMFs (e.g. with an explicit aerosol treatment) and possibly some

841   independent information on the glyoxal vertical distribution. This being said, the nice consistency in the glyoxal

842   column seasonal variability by the different systems is remarkable in itself. Table 4 provides an overview of the

843   correlation coefficient between the satellite and the MAX-DOAS glyoxal columns at all considered stations. For

844   stations where the analysis was possible for all satellite sensors, the correlation coefficient was found to be

845   significantly better for TROPOMI than for the other instruments. It is also clear that correlation coefficients are

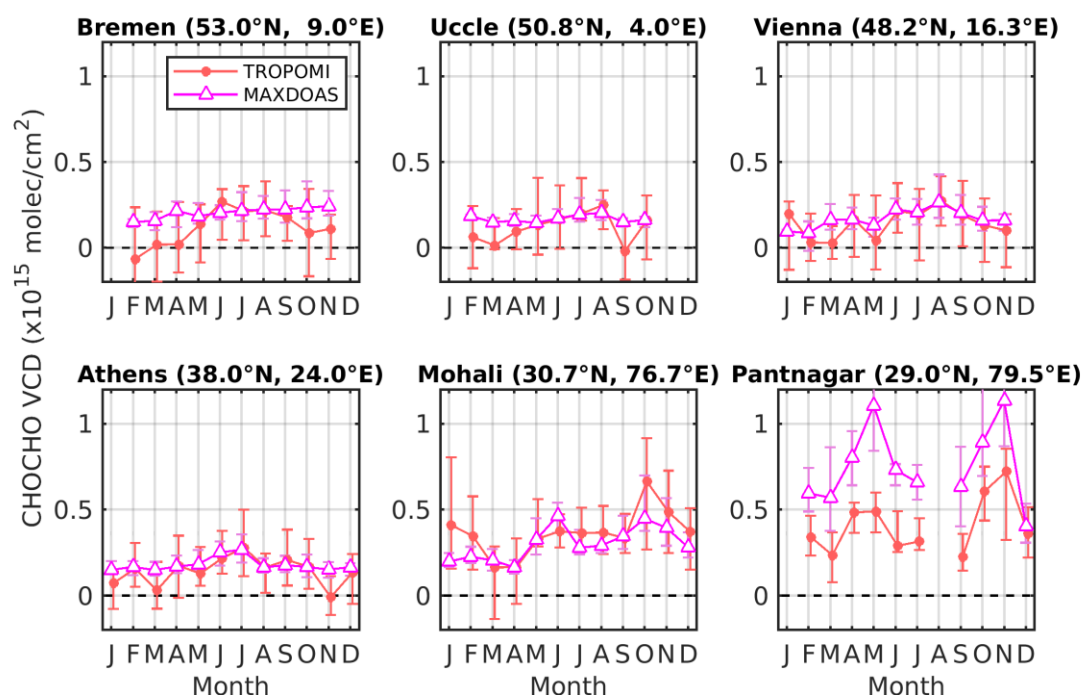
846   better for sites characterised by large and highly variable glyoxal columns (e.g. Asian stations). Apart from the

847   Bremen station where the negative bias during winter leads to a low correlation coefficient, all other values are

848   quite reasonable (between 0.61 and 0.87) for TROPOMI. Table 4 also gives the mean bias as derived from the

849 comparison of the satellite and MAX-DOAS glyoxal column seasonal cycle as well as the standard deviation of

850 the differences. As discussed above, the mean differences are generally lower than 1×10^{14} molec/cm², except for

851 high columns where differences are noticeably higher.

852

853 **Figure 18 : Comparison of the monthly median glyoxal tropospheric vertical column seasonal cycle as retrieved from**

854 **TROPOMI and MXD at four European stations (Bremen, Uccle, Vienna, Athens) and at two Indian stations (Mohali,**

855 **Pantnagar). The error bars represent the interannual variability (25% and 75% percentiles based on the full time**

856 **series available).**

857

858 **Table 4 : Correlation coefficients between the satellite and MAX-DOAS monthly median glyoxal tropospheric**

859 **vertical columns as well as mean absolute difference and associated standard deviation at nine stations.**

	Correlation coefficient $Mean\ bias \pm standard\ deviation\ (x10^{14}\ molec/cm^2)$								
	Xianghe	**Chiba**	**Phimai**	**Bremen**	**Uccle**	**Vienna**	**Athens**	**Mohali**	**Pantnagar**
TROPOMI	0.87 *-0.8±0.6*	0.80 *0.1±0.6*	0.85 *-2.0±0.8*	0.13 *-0.9±0.9*	0.67 *-0.5±0.7*	0.73 *-0.3±0.6*	0.61 *-0.4±0.6*	0.70 *0.6±0.9*	0.78 *-3.5±1.5*
OMI (until 2013)	0.70 *-0.7±1.3*	0.32 *-0.6±0.8*	N/A						

GOME-2B (until 2016)	0.37 *-0.9±0.9*	0.66 *0.0±0.7*	0.88 *-0.8±0.8*						
GOME-2A	0.92 *-0.8±0.4*	0.58 *-0.1±0.9*	0.86 *-1.1±0.8*						

860 **6. Conclusions**

861 We presented the first global TROPOMI glyoxal tropospheric column product derived from three years (2018-
862 2020) of visible radiance measurements. The DOAS-based algorithm, which relies largely on previous
863 developments for heritage satellite nadir-viewing instruments, has been further improved in different aspects. In
864 particular, the use of additional pseudo cross-sections in the DOAS spectral fit allows mitigating the effect of the
865 instrumental spectral response function perturbations in case of scene brightness inhomogeneity, which otherwise
866 would lead to systematic biases in the retrieved glyoxal columns. This helps removing artefacts along the coasts
867 and reducing pseudo-noise in regions covered by persistent broken clouds. The glyoxal slant columns are also
868 empirically corrected for biases caused by the NO_2 misfit in case of strong absorption. Finally, the background
869 correction procedure has been optimized for the TROPOMI characteristics and the a priori glyoxal vertical
870 distribution, essential to the AMF computation, is now provided by the CTM MAGRITTE, an updated version of
871 the IMAGES model, running at the higher spatial resolution of 1°x1°. The glyoxal column retrievals have been
872 fully characterized with an error budget considering the different error components introduced in each of the
873 algorithm modules. This allows extending the glyoxal column data product with total random and systematic error
874 estimates provided for every observation, with corresponding averaging kernels and a priori profiles.

875 Glyoxal tropospheric columns have also been derived from data of the OMI, GOME-2A and GOME-2B satellite
876 instruments using retrieval baselines similar to the TROPOMI algorithm. An extensive inter-comparison of those
877 four data sets emphasised their excellent consistency with absolute mean glyoxal column differences found to be
878 generally lower than $0.5x10^{14}$ molec/cm². This demonstrates that glyoxal retrievals respond in the same manner
879 to our selection of settings for all nadir-viewing satellite instruments. Because of this sensitivity, the retrievals
880 may be easily impacted by spectral features caused by instrumental degradation. We have shown that the stability
881 of the OMI and GOME-2 data records is somewhat degraded after a few years of operations. Glyoxal retrievals
882 are characterized by a high level-of-noise, requiring significant spatio-temporal averaging to extract meaningful
883 signals. With both a much larger number of observations and a finer spatial resolution, TROPOMI outperforms
884 by far the previous instruments in its ability to provide high quality and detailed glyoxal fields. Although
885 consistently identified in our four satellite data sets, the origin of the glyoxal oceanic signal remains unclear. There
886 appears to be an inconsistency between what is measured from space and most glyoxal concentration
887 measurements conducted in marine boundary layer campaigns. Non-negligible glyoxal concentrations in the free
888 troposphere as measured during one campaign (Volkamer et al., 2015) might reconcile the satellite and field data.
889 On the other hand, part of this signal may also be partly caused by remaining spectral interferences (e.g. with
890 water vapour).

891 Satellite observations have also been compared with a few independent MAX-DOAS data sets from stations
892 located in Asia and Europe. Owing to the scarcity of MAX-DOAS glyoxal data sets, especially covering several

seasons, this validation exercise is therefore unprecedented. Based on a thorough analysis at the Xianghe station (China), where a 10-year time series of MAX-DOAS data is available, and on the comparison of seasonal cycles at other stations, we conclude that satellite and MAX-DOAS instruments observe consistent glyoxal signals and have similar intra-annual variations. This is reflected by the strong correlation coefficients, ranging between 0.61 and 0.87 for TROPOMI, with the exception of one mid-latitude station where the correlation is poorer. In general, the satellite and MAX-DOAS columns also agree in absolute values with differences less than 1×10^{14} molec/cm², at least for stations with moderate columns. In Xianghe, we showed that the application of the satellite averaging kernels to the MAX-DOAS data further reduces the mean differences. There are however two stations (Phimai/Thailand and Pantnagar/India) where the satellite/MAX-DOAS bias is more significant, despite a reasonable agreement of the measured seasonal variations. Although the origin of this bias is not fully understood, the MAX-DOAS columns at those stations are very high and it is not uncommon to have such biases in UV-Visible satellite retrievals for strongly polluted sites. It cannot be excluded that part of the bias originates from the MAX-DOAS retrieval strategy at those sites. We have also indications that the satellite observations are low-biased during wintertime at mid-high latitudes where both the glyoxal signal is weak and the sensitivity to the boundary layer is reduced. The comparisons of OMI, GOME-2 and MAX-DOAS glyoxal columns also show reasonable agreement and similar intra-annual variability. Both the correlation coefficients and the scatter of the satellite/ground differences were however less good than those of TROPOMI. This points again to the better performance of TROPOMI for the detection of glyoxal from space and to its enhanced capability at providing information on VOC emissions. For future work, it would be beneficial to dedicate more efforts in the homogenization of the MAX-DOAS glyoxal retrievals in terms of both spectral analysis and slant-to-vertical column conversion in order to strengthen their potential for the validation of satellite data sets such as the one presented in this work.

Data availability

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

Author contributions

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

Competing interests

The authors have the following competing interests: Thomas Wagner is chief-executive editor of AMT. Andreas Richter is executive editor of AMT. Diego Loyola, Andreas Richter, Michel Van Roozendael and Thomas

931   Wagner act as associate editors for AMT.

932

933   **Acknowledgments**

934   This work contains modified Copernicus Sentinel-5 Precursor satellite data (2018-2020). It has been supported

935   by the European Space Agency via the preparation of the Level-2 Prototype Processor of the future Copernicus

936   Sentinel-5 satellite (contract #4000118463/16/NL/AI) and via the GLYRETRO project, part of the Sentinel-

937   5p+Innovation programme (contract #4000127610/19/I-NS). EUMETSAT, the Belgian Federal Science Policy

938   Office (BELSPO) and the German Aerospace Center (DLR) are acknowledged for their respective financial

939   support of the GOME-2 algorithmic developments through the AC SAF Continuous Development and Operations

940   Phase (CDOP-3) and the ProDEx B-ACSAF contribution to the ACSAF. The Vienna MAX-DOAS instrument is

941   part of the VINDOBONA project, which is funded by the Austrian Science Fund (FWF): I 2296-N29, the German

942   Science Foundation (DFG): Ri1800/6-1, and A1 Telekom Austria. IISER Mohali Atmospheric Chemistry Facility

943   is gratefully acknowledged for supporting the MAX-DOAS operations in Mohali, India. We thank Caroline Fayt

944   and Christian Herman from BIRA-IASB for maintaining the Uccle and Xianghe MAX-DOAS instruments.

945

946   **References**

947   Abbot, D. S., Palmer, P. I., Martin, R. V., Chance, K. V., Jacob, D. J. and Guenther, A.: Seasonal and

948   interannual variability of North American isoprene emissions as determined by formaldehyde column

949   measurements from space, Geophys. Res. Lett., 30(17), n/a-n/a, doi:10.1029/2003GL017336, 2003.

950   Aliwell, S. R., Van Roozendael, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P.,

951   Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J. C., Pfeilsticker, K. and Pundt, I.: Analysis for BrO in

952   zenith-sky spectra: An intercomparison exercise for analysis improvement, J. Geophys. Res. Atmos., 107(14),

953   ACH 10-1, doi:10.1029/2001JD000329, 2002.

954   Alvarado, L. M. A., Richter, A., Vrekoussis, M., Wittrock, F., Hilboll, A., Schreier, S. F. and Burrows, J. P.: An

955   improved glyoxal retrieval from OMI measurements, Atmos. Meas. Tech., 7(12), 4133–4150, doi:10.5194/amt-

956   7-4133-2014, 2014.

957   Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Kalisz Hedegaard, A. B., Schneising, O. and

958   Burrows, J. P.: Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus

959   Sentinel-5 Precursor satellite during the 2018 Canadian wildfires, Atmos. Chem. Phys., 20(4), 2057–2072,

960   doi:10.5194/acp-20-2057-2020, 2020.

961   Azam, F. and Richter, A.: GOME2 on MetOp Follow-on analysis of GOME2 in orbit degradation Final Report,

962   2015.

963   Barkley, M. P., Smedt, I. De, Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D.,

964   Guenther, A., Paulot, F., Marais, E. and Mao, J.: Top-down isoprene emissions over tropical South America

965   inferred from SCIAMACHY and OMI formaldehyde columns, J. Geophys. Res. Atmos., 118(12), 6849–6868,

966   doi:10.1002/jgrd.50552, 2013.

967   Barkley, M. P., González Abad, G., Kurosu, T. P., Spurr, R., Torbatian, S. and Lerot, C.: OMI air-quality

968   monitoring over the Middle East, Atmos. Chem. Phys, 17, 4687–4709, doi:10.5194/acp-17-4687-2017, 2017.

969   Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R.,

970   Wiedinmyer, C., Kaiser, J. W., Sindelarova, K. and Guenther, A.: Nine years of global hydrocarbon emissions

971   based on source inversion of OMI formaldehyde observations, Atmos. Chem. Phys., 16(15), 10133–10158,

972   doi:10.5194/acp-16-10133-2016, 2016.

973   Beekmann, M. and Vautard, R.: A modelling study of photochemical regimes over Europe: robustness and

974   variability, Atmos. Chem. Phys., 10(20), 10067–10084, doi:10.5194/acp-10-10067-2010, 2010.

975   Behrens, L. K., Hilboll, A., Richter, A., Peters, E., Alvarado, L. M. A., Kalisz Hedegaard, A. B., Wittrock, F.,

976   Burrows, J. P. and Vrekoussis, M.: Detection of outflow of formaldehyde and glyoxal from the African

977   continent to the Atlantic Ocean with a MAX-DOAS instrument, Atmos. Chem. Phys., 19(15), 10257–10278,

978   doi:10.5194/acp-19-10257-2019, 2019a.

979   Behrens, L. K., Hilboll, A., Richter, A., Peters, E., Alvarado, L. M. A., Hedegaard, A. B. K., Wittrock, F.,

980   Burrows, J. P. and Vrekoussis, M.: Detection of outflow of formaldehyde and glyoxal from the African

981   continent to the Atlantic Ocean with a MAX-DOAS instrument, Atmos. Chem. Phys., 19(15), 10257–10278,

982   doi:10.5194/acp-19-10257-2019, 2019b.

983   Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y. and Wagner, T.: The Mainz profile algorithm

984   (MAPA), Atmos. Meas. Tech., 12(3), 1785–1806, doi:10.5194/amt-12-1785-2019, 2019.

985   Benavent, N., Garcia-Nieto, D., Wang, S. and Saiz-Lopez, A.: MAX-DOAS measurements and vertical profiles

986   of glyoxal and formaldehyde in Madrid, Spain, Atmos. Environ., 199, 357–367,

987   doi:10.1016/j.atmosenv.2018.11.047, 2019.

988   Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., Van Der A, R. J., Sneep, M., Van Den Oord, G.

989   H. J., Levelt, P. F., Stammes, P., Gleason, J. F. and Bucsela, E. J.: Near-real time retrieval of tropospheric NO2

990   from OMI, Atmos. Chem. Phys., 7(8), 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

991   Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V.,

992   Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y. and Brunner, D: An improved

993   tropospheric NO$_2$ column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4(9),

994   1905–1928, doi:10.5194/amt-4-1905-2011, 2011.

995   Brinksma, E. J., Pinardi, G., Volten, H., Braak, R., Richter, A., Schönhardt, A., van Roozendael, M., Fayt, C.,

996   Hermans, C., Dirksen, R. J., Vlemmix, T., Berkhout, A. J. C., Swart, D. P. J., Oetjen, H., Wittrock, F., Wagner,

997   T., Ibrahim, O. W., de Leeuw, G., Moerman, M., Curier, R. L., Celarier, E. A., Cede, A., Knap, W. H.,

998   Veefkind, J. P., Eskes, H. J., Allaart, M., Rothe, R., Piters, A. J. M. and Levelt, P. F.: The 2005 and 2006

999   DANDELIONS NO$_2$ and aerosol intercomparison campaigns, J. Geophys. Res., 113(D16), D16S46,

1000   doi:10.1029/2007JD008808, 2008.

1001   Cao, H., Fu, T. M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., De Smedt, I., Zhang, Q., Van

1002   Roozendael, M., Hendrick, F., Chance, K., Li, J., Zheng, J. and Zhao, Y.: Adjoint inversion of Chinese non-

methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, Atmos. Chem. Phys., 18(20), 15017–15046, doi:10.5194/acp-18-15017-2018, 2018.

Chan, A. W. H., Chan, M. N., Surratt, J. D., Chhabra, P. S., Loza, C. L., Crounse, J. D., Yee, L. D., Flagan, R. C., Wennberg, P. O. and Seinfeld, J. H.: Role of aldehyde chemistry and NOx concentrations in secondary organic aerosol formation, Atmos. Chem. Phys. Discuss., 10(4), 10219–10269 [online] Available from: http://www.atmos-chem-phys-discuss.net/10/10219/2010/, 2010.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J. and Chance, K.: Glyoxal retrieval from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 7(11), 3891–3907, doi:10.5194/amt-7-3891-2014, 2014.

Chan Miller, C., Jacob, D. J., Marais, E. A., Yu, K., Travis, K. R., Kim, P. S., Fisher, J. A., Zhu, L., Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Kaiser, J., Min, K.-E., Brown, S. S., Washenfelder, R. A., González Abad, G. and Chance, K.: Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data, Atmos. Chem. Phys., 17(14), 8725–8738, doi:10.5194/acp-17-8725-2017, 2017.

Chance, K. and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, J. Quant. Spectrosc. Radiat. Transf., 111(9), 1289–1295, doi:10.1016/j.jqsrt.2010.01.036, 2010.

Chance, K. V. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, Appl. Opt., 36(21), 5224, doi:10.1364/ao.36.005224, 1997.

Clémer, K., Van Roozendael, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Spurr, R., Wang, P. and De Mazière, M.: Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, Atmos. Meas. Tech., 3(4), 863–878, doi:10.5194/amt-3-863-2010, 2010.

Coburn, S., Ortega, I., Thalman, R., Blomquist, B., Fairall, C. W. and Volkamer, R.: Measurements of diurnal variations and eddy covariance (EC) fluxes of glyoxal in the tropical marine boundary layer: description of the Fast LED-CE-DOAS instrument, Atmos. Meas. Tech. Discuss., 7(10), 6245–6285, doi:10.5194/amt-7-3579-2014, 2014.

Curci, G., Palmer, P. I., Kurosu, T. P., Chance, K. and Visconti, G.: Estimating European volatile organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10(23), 11501–11517, doi:10.5194/acp-10-11501-2010, 2010.

Danckaert, T., Fayt, C., van Roozendael, M., De Smedt, I., Letocard, V., Merlaud, A. and Pinardi, G.: QDOAS software user manual. [online] Available from: http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf, 2017.

Danielson, J. J. and Gesch, D. B.: Global multi-resolution terrain elevation data 2010 (GMTED2010), 2011.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H. and Van Roozendael, M.: Twelve years of

1038   global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos.

1039   Chem. Phys., 8(16), 4947–4963 [online] Available from: http://www.atmos-chem-phys.net/8/4947/2008/, 2008.

1040   De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen,

1041   C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F. and Van Roozendael, M.: Diurnal, seasonal

1042   and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2

1043   observations, Atmos. Chem. Phys., 15(21), 12519–12545, doi:10.5194/acp-15-12519-2015, 2015.

1044   De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A.,

1045   Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma,

1046   K. F. and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and

1047   from the QA4ECV project, Atmos. Meas. Tech., 11(4), 2395–2426, doi:10.5194/amt-11-2395-2018, 2018.

1048   De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L.,

1049   Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot,

1050   C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys,

1051   N., Vlietinck, J., Wagner, T., Wang, T., Yu, H. and Van Roozendael, M.: Comparative assessment of

1052   TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column

1053   measurements, Atmos. Chem. Phys., 21(16), 12561–12593, doi:10.5194/ACP-21-12561-2021, 2021.

1054   DiGangi, J. P., Henry, S. B., Kammrath, A., Boyle, E. S., Kaser, L., Schnitzhofer, R., Graus, M., Turnipseed,

1055   A., Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin, R. L., Kim, S., Nakashima, Y., Wolfe, G. M.,

1056   Kajii, Y., Apel, E. C. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A., Keutsch, F. N., Park, J.-H. J.-H.,

1057   Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin III, R. L., Kim, S., Nakashima, Y., Wolfe, G. M., Kajii,

1058   Y., Apel, E. C. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A. and Keutsch, F. N.: Observations of

1059   glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, Atmos. Chem.

1060   Phys., 12(20), 9529–9543, doi:10.5194/acp-12-9529-2012, 2012.

1061   Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem.

1062   Phys., 3(5), 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.

1063   EUMETSAT: GOME-2 Product Guide (EUM/OPS-EPS/MAN/07/0445), (March) [online] Available from:

1064   http://www.eumetsat.int, 2011.

1065   Fu, T.-M. T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M. and Henze, D. K.: Global budgets of

1066   atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, J.

1067   Geophys. Res., 113(D15), D15303, doi:10.1029/2007JD009505, 2008.

1068   van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., Maasakkers, J. D. and Veefkind, J. P.: TROPOMI ATBD

1069   of the total and tropospheric NO2 data products, S5p/TROPOMI, (1.4.0), 1–76 [online] Available from:

1070   https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-NO2-data-products, 2019.

1071   Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V.,

1072   Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J. M., Gamache, R. R., Hodges, J. T., Jacquemart, D.,

1073   Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M. A. H., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G.,

1074   Barbe, A., Császár, A. G., Devi, V. M., Furtenbacher, T., Harrison, J. J., Hartmann, J. M., Jolly, A., Johnson, T.

1075   J., Karman, T., Kleiner, I., Kyuberis, A. A., Loos, J., Lyulin, O. M., Massie, S. T., Mikhailenko, S. N.,

1076   Moazzen-Ahmadi, N., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Polyansky, O. L., Rey, M., Rotger,

1077   M., Sharpe, S. W., Sung, K., Starikova, E., Tashkun, S. A., Auwera, J. Vander, Wagner, G., Wilzewski, J.,

1078   Wcisło, P., Yu, S. and Zak, E. J.: The HITRAN2016 molecular spectroscopic database, J. Quant. Spectrosc.

1079   Radiat. Transf., 203, 3–69, doi:10.1016/j.jqsrt.2017.06.038, 2017.

1080   Gratsea, M., Vrekoussis, M., Richter, A., Wittrock, F., Schönhardt, A., Burrows, J., Kazadzis, S., Mihalopoulos,

1081   N. and Gerasopoulos, E.: Slant column MAX-DOAS measurements of nitrogen dioxide, formaldehyde, glyoxal

1082   and oxygen dimer in the urban environment of Athens, Atmos. Environ., 135, 118–131,

1083   doi:10.1016/J.ATMOSENV.2016.03.048, 2016.

1084   Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K. and Wang, X.: The

1085   Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated

1086   framework for modeling biogenic emissions, Geosci. Model Dev., 5(6), 1471–1492, doi:10.5194/gmd-5-1471-

1087   2012, 2012.

1088   Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N.

1089   M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M.

1090   E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S.

1091   H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R. and Wildt, J.: The formation, properties and impact of

1092   secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9(14), 5155–5236 [online]

1093   Available from: http://www.atmos-chem-phys.net/9/5155/2009/, 2009.

1094   Hendrick, F., Clémer, K., Wang, P., De Mazière, M., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G.,

1095   Stavrakou, T., Vlemmix, T. and Van Roozendael, M.: Four years of ground-based MAX-DOAS observations of

1096   HONO and NO2 in the Beijing area, Atmos. Chem. Phys., 14(2), 765–781, doi:10.5194/acp-14-765-2014, 2014.

1097   Hönninger, G., von Friedeburg, C. and Platt, U.: Multi axis differential optical absorption spectroscopy (MAX-

1098   DOAS), Atmos. Chem. Phys., 4(1), 231–254, doi:10.5194/acp-4-231-2004, 2004.

1099   Hoque, H. M. S., Irie, H. and Damiani, A.: First MAX-DOAS Observations of Formaldehyde and Glyoxal in

1100   Phimai, Thailand, J. Geophys. Res. Atmos., 123(17), 9957–9975, doi:10.1029/2018JD028480, 2018.

1101   Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean,

1102   M., Schaaf, E. and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic

1103   compounds: A global gridded data set for 1970-2012, Atmos. Chem. Phys., 17(12), 7683–7701,

1104   doi:10.5194/acp-17-7683-2017, 2017.

1105   Irie, H., Takashima, H., Kanaya, Y., Boersma, K. F., Gast, L., Wittrock, F., Brunner, D., Zhou, Y. and Van

1106   Roozendael, M.: Eight-component retrievals from ground-based MAX-DOAS observations, Atmos. Meas.

1107   Tech., 4(6), 1027–1044, doi:10.5194/amt-4-1027-2011, 2011.

1108   Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press. [online] Available from:

1109   https://press.princeton.edu/books/hardcover/9780691001852/introduction-to-atmospheric-chemistry (Accessed

1110   19 March 2021), 2000.

1111   Javed, Z., Liu, C., Khokhar, M., Tan, W., Liu, H., Xing, C., Ji, X., Tanvir, A., Hong, Q., Sandhu, O., Rehman,

1112   A., Javed, Z., Liu, C., Khokhar, M. F., Tan, W., Liu, H., Xing, C., Ji, X., Tanvir, A., Hong, Q., Sandhu, O. and

1113   Rehman, A.: Ground-Based MAX-DOAS Observations of CHOCHO and HCHO in Beijing and Baoding,

1114   China, Remote Sens., 11(13), 1524, doi:10.3390/rs11131524, 2019.

1115   Jin, X., Jin, X., Fiore, A., Fiore, A., Boersma, K. F., Boersma, K. F., Smedt, I. De and Valin, L.: Inferring

1116   Changes in Summertime Surface Ozone-NOx-VOC Chemistry over U.S. Urban Areas from Two Decades of

1117   Satellite and Ground-Based Observations, Environ. Sci. Technol., 54(11), 6518–6529,

1118   doi:10.1021/acs.est.9b07785, 2020.

1119   Kaiser, J., Wolfe, G. M., Min, K. E., Brown, S. S., Miller, C. C., Jacob, D. J., deGouw, J. A., Graus, M.,

1120   Hanisco, T. F., Holloway, J., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Washenfelder, R. A. and

1121   Keutsch, F. N.: Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor

1122   speciation, Atmos. Chem. Phys., 15(13), 7571–7583, doi:10.5194/acp-15-7571-2015, 2015.

1123   Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat,

1124   R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J.,

1125   Vacanti, G., Vonk, F. and Veefkind, P.: Pre-launch calibration results of the TROPOMI payload on-board the

1126   Sentinel-5 Precursor satellite, Atmos. Meas. Tech., 11(12), 6439–6479, doi:10.5194/amt-11-6439-2018, 2018.

1127   Kleipool, Q. L., Dobber, M. R., de Haan, J. F. and Levelt, P. F.: Earth surface reflectance climatology from 3

1128   years of OMI data, J. Geophys. Res., 113(D18), D18308, doi:10.1029/2008JD010290, 2008.

1129   Kluge, F., Hüneke, T., Knecht, M., Lichtenstern, M., Rotermund, M., Schlager, H., Schreiner, B. and

1130   Pfeilsticker, K.: Profiling of formaldehyde, glyoxal, methylglyoxal, and CO over the Amazon: Normalized

1131   excess mixing ratios and related emission factors in biomass burning plumes, Atmos. Chem. Phys., 20(20),

1132   12363–12389, doi:10.5194/acp-20-12363-2020, 2020.

1133   Knote, C., Hodzic, A., Jimenez, J. L., Volkamer, R., Orlando, J. J., Baidar, S., Brioude, J., Fast, J., Gentner, D.

1134   R., Goldstein, A. H., Hayes, P. L., Knighton, W. B., Oetjen, H., Setyan, A., Stark, H., Thalman, R., Tyndall, G.,

1135   Washenfelder, R., Waxman, E. and Zhang, Q.: Simulation of semi-explicit mechanisms of SOA formation from

1136   glyoxal in aerosol in a 3-D model, Atmos. Chem. Phys, 14, 6213–6239, doi:10.5194/acp-14-6213-2014, 2014.

1137   Kosarev, A. N., Kostianoy, A. G. and Zonn, I. S.: Kara-Bogaz-Gol bay: Physical and chemical evolution, Aquat.

1138   Geochemistry, 15(1–2), 223–236, doi:10.1007/s10498-008-9054-z, 2009.

1139   Kumar, V., Sarkar, C. and Sinha, V.: Influence of post-harvest crop residue fires on surface ozone mixing ratios

1140   in the N.W. IGP analyzed using 2years of continuous in situ trace gas measurements, J. Geophys. Res., 121(7),

1141   3619–3633, doi:10.1002/2015JD024308, 2016.

1142   Kumar, V., Beirle, S., Dörner, S., Mishra, A. K., Donner, S., Wang, Y., Sinha, V. and Wagner, T.: Long term

1143   MAX-DOAS measurements of NO2, HCHO and aerosols and evaluation of corresponding satellite data

1144   products over Mohali in the Indo-Gangetic plain, Atmos. Chem. Phys., (2), 1–62, doi:10.5194/acp-2020-404,

1145   2020.

1146 Lawson, S. J., Selleck, P. W., Galbally, I. E., Keywood, M. D., Harvey, M. J., Lerot, C., Helmig, D. and

1147 Ristovski, Z.: Seasonal in situ observations of glyoxal and methylglyoxal over the temperate oceans of the

1148 Southern Hemisphere, Atmos. Chem. Phys., 15(1), 223–240, doi:10.5194/acp-15-223-2015, 2015.

1149 Lerot, C., Stavrakou, T., De Smedt, I., Müller, J. F., Van Roozendael, M., M̈uller, J. F. and Van Roozendael,

1150 M.: Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global

1151 model, Atmos. Chem. Phys., 10(24), 12059–12072, doi:10.5194/acp-10-12059-2010, 2010.

1152 Levelt, P. F., Van Den Oord, G. H. J., Dobber, M. R., Mälkki, A., Visser, H., De Vries, J., Stammes, P.,

1153 Lundell, J. O. V and Saari, H.: The Ozone Monitoring Instrument, IEEE Trans. Geosci. Remote Sens., 44(5),

1154 1093, doi:10.1109/TGRS.2006.872333, 2006.

1155 Li, J., Mao, J., Min, K.-E., Washenfelder, R. A., Brown, S. S., Kaiser, J., Keutsch, F. N., Volkamer, R., Wolfe,

1156 G. M., Hanisco, T. F., Pollack, I. B., Ryerson, T. B., Graus, M., Gilman, J. B., Lerner, B. M., Warneke, C., de

1157 Gouw, J. A., Middlebrook, A. M., Liao, J., Welti, A., Henderson, B. H., McNeill, V. F., Hall, S. R., Ullmann,

1158 K., Donner, L. J., Paulot, F. and Horowitz, L. W.: Observational constraints on glyoxal production from

1159 isoprene oxidation and its contribution to organic aerosol over the Southeast United States, J. Geophys. Res.

1160 Atmos., 121(16), 9849–9861, doi:10.1002/2016JD025331, 2016.

1161 Liu, Z., Wang, Y., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., Shao, M., Chang, C.-C., Liu, S.-C.,

1162 Wang, H. and Chen, C.: Exploring the missing source of glyoxal (CHOCHO) over China, Geophys. Res. Lett.,

1163 39(10), L10812, doi:10.1029/2012GL051645, 2012.

1164 Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley,

1165 M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J. T., Krotkov, N., Stammes, P.,

1166 Wang, P., Eskes, H. J. and Krol, M.: Structural uncertainty in air mass factor calculation for NO2 and HCHO

1167 satellite retrievals, Atmos. Meas. Tech., 10(3), 759–782, doi:10.5194/amt-10-759-2017, 2017.

1168 Lorente, A., Boersma, K. F., Stammes, P., Tilstra, L. G., Richter, A., Yu, H., Kharbouche, S. and Muller, J.-P.:

1169 The importance of surface reflectance anisotropy for cloud and NO2 retrievals from GOME-2 and OMI, Atmos.

1170 Meas. Tech., 11(7), 4509–4529, doi:10.5194/amt-11-4509-2018, 2018.

1171 Loyola, D. G., Xu, J., Heue, K. P. and Zimmer, W.: Applying FP-ILM to the retrieval of geometry-dependent

1172 effective Lambertian equivalent reflectivity (GE-LER) daily maps from UVN satellite measurements, Atmos.

1173 Meas. Tech., 13(2), 985–999, doi:10.5194/amt-13-985-2020, 2020.

1174 Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E.,

1175 Rozemeijer, N., Vonk, F. and Veefkind, P.: In-flight calibration results of the TROPOMI payload on board the

1176 Sentinel-5 Precursor satellite, Atmos. Meas. Tech., 13(7), 3561–3580, doi:10.5194/amt-13-3561-2020, 2020.

1177 Lutz, R., Loyola, D., García, S. G. and Romahn, F.: OCRA radiometric cloud fractions for GOME-2 on MetOp-

1178 A/B, Atmos. Meas. Tech., 9(5), 2357–2379, doi:10.5194/amt-9-2357-2016, 2016.

1179 Mahajan, A. S., Prados-Roman, C., Hay, T. D., Lampel, J., Pöhler, D., Groβmann, K., Tschritter, J., Frieß, U.,

1180 Platt, U., Johnston, P., Kreher, K., Wittrock, F., Burrows, J. P., Plane, J. M. C. and Saiz-Lopez, A.: Glyoxal

1181 observations in the global marine boundary layer, J. Geophys. Res. Atmos., n/a–n/a,

1182 doi:10.1002/2013JD021388, 2014.

1183 Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet,

1184 D. B., Barkley, M. P., Paulot, F. and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of

1185 formaldehyde columns, Atmos. Chem. Phys., 12(14), 6219–6235, doi:10.5194/acp-12-6219-2012, 2012.

1186 Mason, J. D., Cone, M. T. and Fry, E. S.: Ultraviolet (250–550 nm) absorption spectrum of pure water, Appl.

1187 Opt., 55(25), 7163, doi:10.1364/AO.55.007163, 2016.

1188 Michaela Friedrich, M., Rivera, C., Stremme, W., Ojeda, Z., Arellano, J., Bezanilla, A., García-Reynoso, J. A.

1189 and Grutter, M.: NO2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City,

1190 Atmos. Meas. Tech., 12(4), 2545–2565, doi:10.5194/amt-12-2545-2019, 2019.

1191 Müller, J.-F. and Brasseur, G.: IMAGES: A three-dimensional chemical transport model of the global

1192 troposphere, J. Geophys. Res., 100(D8), 16445, doi:10.1029/94JD03254, 1995.

1193 Müller, J.-F., Stavrakou, T., Bauwens, M., Compernolle, S. and Peeters, J.: Chemistry and deposition in the

1194 Model of Atmospheric composition at Global and Regional scales using Inversion Techniques for Trace gas

1195 Emissions (MAGRITTE v1.0). Part B. Dry deposition, Geosci. Model Dev. Discuss., 1–49, doi:10.5194/gmd-

1196 2018-317, 2018.

1197 Müller, J.-F., Stavrakou, T. and Peeters, J.: Chemistry and deposition in the Model of Atmospheric composition

1198 at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE v1.1) – Part 1:

1199 Chemical mechanism, Geosci. Model Dev., 12(6), 2307–2356, doi:10.5194/gmd-12-2307-2019, 2019.

1200 Müller, J. F., Stavrakou, T., Wallens, S., De Smedt, I., Van Roozendael, M., Potosnak, M. J., Rinne, J., Munger,

1201 B., Goldstein, A. and Guenther, A. B.: Global isoprene emissions estimated using MEGAN, ECMWF analyses

1202 and a detailed canopy environment model, Atmos. Chem. Phys., 8(5), 1329–1341, doi:10.5194/acp-8-1329-

1203 2008, 2008.

1204 Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M.,

1205 Holdak, A., Kokhanovsky, A., Livschitz, J. and Eisinger, M.: The GOME-2 instrument on the Metop series of

1206 satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9(3),

1207 1279–1301, doi:10.5194/amt-9-1279-2016, 2016.

1208 Myriokefalitakis, S., Vrekoussis, M., Tsigaridis, K., Wittrock, F., Richter, A., Brühl, C., Volkamer, R., Burrows,

1209 J. P. and Kanakidou, M.: The influence of natural and anthropogenic secondary sources on the glyoxal global

1210 distribution, Atmos. Chem. Phys., 8(16), 4965—4981 [online] Available from: http://www.atmos-chem-

1211 phys.net/8/4965/2008/, 2008.

1212 Noël, S., Bramstedt, K., Bovensmann, H., Gerilowski, K., Burrows, J. P., Standfuss, C., Dufour, E. and

1213 Veihelmann, B.: Quantification and mitigation of the impact of scene inhomogeneity on Sentinel-4 UVN UV-

1214 VIS retrievals, Atmos. Meas. Tech., 5(6), 1319–1331, doi:10.5194/amt-5-1319-2012, 2012.

1215 Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore,

1216 A. and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to

1217 formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res. Atmos., 106(D13),
1218 14539–14550, doi:10.1029/2000JD900772, 2001.

1219 Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C.,
1220 Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B. and Sumner, A. L.: Quantifying the seasonal and
1221 interannual variability of North American isoprene emissions using satellite observations of the formaldehyde
1222 column, J. Geophys. Res., 111(D12), D12315, doi:10.1029/2005JD006689, 2006.

1223 Peters, E., Wittrock, F., Richter, A., Alvarado, L. M. A., Rozanov, V. V. and Burrows, J. P.: Liquid water
1224 absorption and scattering effects in DOAS retrievals over oceans, Atmos. Meas. Tech., 7(12), 4203–4221,
1225 doi:10.5194/amt-7-4203-2014, 2014.

1226 Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy: Principles and Applications, Springer-
1227 Verlag., 2008.

1228 Puķīte, J., Kühl, S., Deutschmann, T., Platt, U. and Wagner, T.: Extending differential optical absorption
1229 spectroscopy for limb measurements in the UV, Atmos. Meas. Tech., 3(3), 631–653, doi:10.5194/amt-3-631-
1230 2010, 2010.

1231 Richter, A.: qa4ecv_no2_inhomogeneous_scenes_171221, in QA4ECV technical note, Treatment of
1232 inhomogeneous scenes., 2018.

1233 Richter, A. and Burrows, J. P.: Tropospheric NO2 from GOME measurements, Adv. Sp. Res., 29(11), 1673–
1234 1683, doi:10.1016/S0273-1177(02)00100-X, 2002.

1235 Richter, A., Begoin, M., Hilboll, A. and Burrows, J. P.: An improved NO$_2$ retrieval for
1236 the GOME-2 satellite instrument, Atmos. Meas. Tech., 4(6), 1147–1159, doi:10.5194/amt-4-1147-2011, 2011.

1237 Richter, A., Hilboll, A., Sanders, A., Peters, E. and Burrows, J. P.: Inhomogeneous scene effects in OMI NO 2
1238 observations, EGU General Assembly 2018. [online] Available from: http://www.doas-
1239 bremen.de/posters/egu_2018_richter.pdf (Accessed 19 May 2021), 2018.

1240 Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., Birk, M., Bizzocchi, L.,
1241 Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coudert, L. H., Devi, V. M., Drouin, B.
1242 J., Fayt, A., Flaud, J. M., Gamache, R. R., Harrison, J. J., Hartmann, J. M., Hill, C., Hodges, J. T., Jacquemart,
1243 D., Jolly, A., Lamouroux, J., Le Roy, R. J., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T.,
1244 Mikhailenko, S., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V., Perrin, A.,
1245 Polovtseva, E. R., Richard, C., Smith, M. A. H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G. C.,
1246 Tyuterev, V. G. and Wagner, G.: The HITRAN2012 molecular spectroscopic database, J. Quant. Spectrosc.
1247 Radiat. Transf., 130, 4–50, doi:10.1016/j.jqsrt.2013.07.002, 2013.

1248 Rozanov, A., Rozanov, V., Buchwitz, M., Kokhanovsky, A. and Burrows, J. P.: SCIATRAN 2.0 - A new
1249 radiative transfer model for geophysical applications in the 175-2400 nm spectral region, in Advances in Space
1250 Research, vol. 36, pp. 1015–1019, Elsevier Ltd., 2005.

1251 Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Pepijn

1252   Veefkind, J. and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech,
1253   10, doi:10.5194/amt-10-1957-2017, 2017.

1254   Schreier, S. F., Richter, A., Peters, E., Ostendorf, M., Schmalwieser, A. W., Weihs, P. and Burrows, J. P.: Dual
1255   ground-based MAX-DOAS observations in Vienna, Austria: Evaluation of horizontal and temporal NO2,
1256   HCHO, and CHOCHO distributions and comparison with independent data sets, Atmos. Environ. X, 5, 100059,
1257   doi:10.1016/J.AEAOA.2019.100059, 2020.

1258   Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W. and Burrows, J. P.: High spectral resolution ozone
1259   absorption cross-sections – Part 2: Temperature dependence, Atmos. Meas. Tech., 7(2), 625–636,
1260   doi:10.5194/amt-7-625-2014, 2014.

1261   Sinreich, R., Coburn, S., Dix, B. and Volkamer, R.: Ship-based detection of glyoxal over the remote tropical
1262   Pacific Ocean, Atmos. Chem. Phys., 10(23), 11359–11371, doi:10.5194/acp-10-11359-2010, 2010.

1263   Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Van Der Werf, G. R., Giglio, L. and Guenther,
1264   A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-
1265   based formaldehyde columns. [online] Available from: www.atmos-chem-phys.net/9/1037/2009/ (Accessed 19
1266   March 2021a), 2009.

1267   Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F.,
1268   Richter, A. and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne
1269   measurements and inverse modelling, Atmos. Chem. Phys., 9(21), 8431–8446 [online] Available from:
1270   http://www.atmos-chem-phys.net/9/8431/2009/, 2009b.

1271   Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T. and Zhang, Q.: Key
1272   chemical NO$_x$ sink uncertainties and how they influence top-down emissions of nitrogen oxides, Atmos. Chem.
1273   Phys., 13(17), 9057–9082, doi:10.5194/acp-13-9057-2013, 2013.

1274   Stavrakou, T., Müller, J.-F., Bauwens, M., Smedt, I. De, Lerot, C., Roozendael, M. Van, Coheur, P.-F.,
1275   Clerbaux, C., Boersma, K. F., A, R. van der, Song, Y., Jeong, S.-J., Huang, X., Song, Y., Li, M., Li, J., Zhu, T.,
1276   Yamaji, K., Werf, G. R. van der, Huang, X., Li, M., Li, J., Song, Y., Fu, T.-M., Levelt, P. F., Smedt, I. De,
1277   Smedt, I. De, Boersma, K. F., Lerot, C., Müller, J.-F., Stavrakou, T., Stavrakou, T., Müller, J.-F., Smedt, I. De,
1278   Roozendael, M. Van, Werf, G. van der, Giglio, L., Guenther, A., Stavrakou, T., Guenther, A., Karl, T., Harley,
1279   P., Wiedinmyer, C., Palmer, P. I., Geron, C., Li, M., Andreae, M. O., Merlet, P., Akagi, S., Kurokawa, J., Sun,
1280   L., Jin, X., Holloway, T., Safieddine, S., Chaudhry, Z., Shi, Z., Tao, J. Z., Wang, Z., Han, D., Li, S., Su, L.,
1281   Chen, L., Deng, X., Dee, D. P., Randerson, J., Chen, Y., Werf, G., Rogers, B., Morton, D., Kaiser, J.,
1282   Wiedinmyer, C., Kudo, S., Inomata, S., Warneke, C., Yokelson, R. J., Korontzi, S., McCarty, J., Loboda, T.,
1283   Kumar, S., Justice, C., Fu, T.-M., Stavrakou, T., Castellanos, P., Boersma, K. F., Werf, G. R. van der, Razavi,
1284   A., Stavrakou, T., Lin, J., Castellanos, P., Boersma, K. F., Torres, O., Haan, J. F. de, Barkley, M. P., Roberts,
1285   G., Wooster, M. J., Lagoudakis, E., Stavrakou, T., Miller, C. C., Jacob, D. J., et al.: Substantial Underestimation
1286   of Post-Harvest Burning Emissions in the North China Plain Revealed by Multi-Species Space Observations,
1287   Sci. Reports, Publ. online 31 August 2016; | doi10.1038/srep32307, 6, 615–619, doi:10.1038/SREP32307, 2016.

1288   Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O2–O2 collision pairs

1289 between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15(37), 15371,
1290 doi:10.1039/c3cp50968k, 2013.

1291 Theys, N., Volkamer, R., Müller, J. F., Zarzana, K. J., Kille, N., Clarisse, L., De Smedt, I., Lerot, C.,
1292 Finkenzeller, H., Hendrick, F., Koenig, T. K., Lee, C. F., Knote, C., Yu, H. and Van Roozendael, M.: Global
1293 nitrous acid emissions and levels of regional oxidants enhanced by wildfires, Nat. Geosci., 13(10), 681–686,
1294 doi:10.1038/s41561-020-0637-7, 2020.

1295 Tilstra, L., Tuinder, O., Wang, P. and Stammes, P.: Directionally dependent Lambertian-equivalent reflectivity
1296 (DLER) of the Earth's surface measured by the GOME-2 satellite instruments, Atmos. Meas. Tech. Discuss., 1–
1297 29, doi:10.5194/amt-2020-502, 2021.

1298 Tilstra, L. G., Tuinder, O. N. E., Wang, P. and Stammes, P.: Surface reflectivity climatologies from UV to NIR
1299 determined from Earth observations by GOME-2 and SCIAMACHY, J. Geophys. Res. Atmos.,
1300 doi:10.1002/2016JD025940, 2017.

1301 Torres, O., Bhartia, P. K., Jethva, H. and Ahn, C.: Impact of the ozone monitoring instrument row anomaly on
1302 the long-term record of aerosol products, Atmos. Meas. Tech, 11, 2701–2715, doi:10.5194/amt-11-2701-2018,
1303 2018.

1304 Valks, P., Hao, N. and Lerot, C.: Algorithm Theoretical Basis Document for GOME-2 glyoxal column data
1305 records, SAF/AC/DLR/ATBD/GLY/01; Iss. 1/B., 2020.

1306 Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A.
1307 and Coquart, B.: Measurements of the NO2 absorption cross-section from 42 000 cm−1 to 10 000 cm−1 (238–
1308 1000 nm) at 220 K and 294 K, J. Quant. Spectrosc. Radiat. Transf., 59(3–5), 171–184, doi:10.1016/S0022-
1309 4073(97)00168-4, 1998.

1310 Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F.,
1311 Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors,
1312 R., Kruizinga, B., Vink, R., Visser, H. and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES
1313 mission for global observations of the atmospheric composition for climate, air quality and ozone layer
1314 applications, Remote Sens. Environ., 120, 70–83, doi:10.1016/j.rse.2011.09.027, 2012.

1315 Veefkind, J. P., Haan, J. F. de, Sneep, M. and Levelt, P. F.: Improvements to the OMI O_2–O_2 operational cloud
1316 algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9(12), 6035–6049,
1317 doi:10.5194/AMT-9-6035-2016, 2016.

1318 Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J. C., Eskes, H. J., Eichmann, K. U., Fjæraa, A. M.,
1319 Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A.,
1320 Folkert Boersma, K., Bognar, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter De
1321 La Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagkiozidis, D., Kivi,
1322 R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., Piters, A. J. M., Pommereau, J. P.,
1323 Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Cárdenas, C.
1324 R., De Miguel, L. S., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendael, M., Pepijn Veefkind, J.,

1325 Wagner, T., Wittrock, F., Yela González, M. and Zehner, C.: Ground-based validation of the Copernicus

1326 Sentinel-5P TROPOMI NO2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global

1327 networks, Atmos. Meas. Tech., 14(1), 481–510, doi:10.5194/amt-14-481-2021, 2021.

1328 Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I.,

1329 Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger,

1330 J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale,

1331 D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P. and Winkler, H.:

1332 TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-

1333 transform infrared stations, Atmos. Meas. Tech., 13(7), 3751–3767, doi:10.5194/amt-13-3751-2020, 2020.

1334 Vohra, K., Vodonos, A., Schwartz, J., Marais, E. A., Sulprizio, M. P. and Mickley, L. J.: Global mortality from

1335 outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem, Environ. Res.,

1336 195, 110754, doi:10.1016/j.envres.2021.110754, 2021.

1337 Volkamer, R., Spietz, P., Burrows, J. and Platt, U.: High-resolution absorption cross-section of glyoxal in the

1338 UV—vis and IR spectral ranges, J. Photochem. Photobiol. A Chem., 172(1), 35—46 [online] Available from:

1339 http://www.colorado.edu/chemistry/volkamer/publications/articles/Volkamer etal (2005) HR cross section

1340 glyoxal.pdf, 2005.

1341 Volkamer, R., San Martini, F., Molina, L. T., Salcedo, D., Jimenez, J. L. and Molina, M. J.: A missing sink for

1342 gas-phase glyoxal in Mexico City: Formation of secondary organic aerosol, Geophys. Res. Lett., 34(19),

1343 L19807, doi:10.1029/2007GL030752, 2007.

1344 Volkamer, R., Baidar, S., Campos, T. L., Coburn, S., DiGangi, J. P., Dix, B., Eloranta, E. W., Koenig, T. K.,

1345 Morley, B., Ortega, I., Pierce, B. R., Reeves, M., Sinreich, R., Wang, S., Zondlo, M. A. and Romashkin, P. A.:

1346 Aircraft measurements of BrO, IO, glyoxal, NO2, H2O, O2–O2 and aerosol extinction profiles in the tropics:

1347 comparison with aircraft-/ship-based in situ and lidar measurements, Atmos. Meas. Tech., 8(5), 2121–2148,

1348 doi:10.5194/amt-8-2121-2015, 2015.

1349 Voors, R., Dobber, M., Dirksen, R. and Levelt, P.: Method of calibration to correct for cloud-induced

1350 wavelength shifts in the Aura satellite's Ozone Monitoring Instrument, Appl. Opt., 45(15), 3652–3658,

1351 doi:10.1364/AO.45.003652, 2006.

1352 Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P.: Temporal and spatial variability of glyoxal as

1353 observed from space, Atmos. Chem. Phys., 9(13), 4485–4504, doi:10.5194/acp-9-4485-2009, 2009.

1354 Vrekoussis, M., Wittrock, F., Richter, A. and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what

1355 can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem. Phys., 10(21), 10145–

1356 10160, doi:10.5194/acp-10-10145-2010, 2010.

1357 Wagner, T., Beirle, S. and Deutschmann, T.: Three-dimensional simulation of the Ring effect in observations of

1358 scattered sun light using Monte Carlo radiative transfer models, Atmos. Meas. Tech., 2(1), 113–124,

1359 doi:10.5194/amt-2-113-2009, 2009.

1360 Wells, K. C., Millet, D. B., Payne, V. H., Deventer, M. J., Bates, K. H., de Gouw, J. A., Graus, M., Warneke, C.,

1361 Wisthaler, A. and Fuentes, J. D.: Satellite isoprene retrievals constrain emissions and atmospheric oxidation,
1362 Nature, 585(7824), 225–233, doi:10.1038/s41586-020-2664-3, 2020.

1363 Van Der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van
1364 Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J. and Kasibhatla, P. S.: Global fire emissions
1365 estimates during 1997-2016, Earth Syst. Sci. Data, 9(2), 697–720, doi:10.5194/essd-9-697-2017, 2017.

1366 Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle,
1367 S., Platt, U. and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space,
1368 Geophys. Res. Lett., 33(16), L16804, doi:10.1029/2006GL026310, 2006.

1369 World Health Organization: Ambient air pollution: A global assessment of exposure and burden of disease.
1370 [online] Available from: https://apps.who.int/iris/handle/10665/250141 (Accessed 19 March 2021), 2016.

1371 Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., Van Geffen, J. H. G. M., Beirle, S., Wagner, T.,
1372 Van Roozendael, M., Marchenko, S., Lamsal, L. N. and Eskes, H. J.: Improved slant column density retrieval of
1373 nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty
1374 characterization, and trends, Atmos. Meas. Tech. Discuss., 11(7), 1–47, doi:10.5194/amt-11-4033-2018, 2018.

1375 Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R. and Wang, P.: An improved tropospheric NO_2 retrieval for
1376 OMI observations in the vicinity of mountainous terrain, Atmos. Meas. Tech., 2(2), 401–416, doi:10.5194/amt-
1377 2-401-2009, 2009.

1378

1379

1380

1381

1382

1383

1384

1385

1386

1387

1388

1389

1390

1391