**Atmospheric
Measurement
Techniques**

Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues

I. De Smedt[1], M. Van Roozendael[1], T. Stavrakou[1], J.-F. Müller[1], C. Lerot[1], N. Theys[1], P. Valks[2], N. Hao[2], and R. van der A[3]

[1]Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[2]Institut für Methodik der Fernerkundung (IMF), Deutsches Zentrum für Luft und Raumfahrt (DLR), Oberpfaffenhofen, Germany
[3]Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

Correspondence to: I. De Smedt (isabelle.desmedt@aeronomie.be)

Abstract. We present a new dataset of formaldehyde vertical columns retrieved from observations of GOME-2 on board the EUMETSAT MetOp-A platform between 2007 and 2011. The new retrieval scheme, which has been optimised for GOME-2, includes a two-step fitting procedure that strongly reduces the impact of spectral interferences between H_2CO and BrO, and a modified DOAS approach that better handles ozone absorption effects at moderately low sun elevations. Owing to these new features, the noise in the H_2CO slant columns is reduced by up to 40 % in comparison to baseline retrieval settings used operationally. Also, the previously reported underestimation of the H_2CO columns in tropical and mid-latitude regions has been largely eliminated, improving the agreement with coincident SCIAMACHY observations. To compensate for the drift of the GOME-2 slit function and to mitigate the instrumental degradation effects on H_2CO retrievals, an asymmetric Gaussian line-shape is fitted during the irradiance calibration. Additionally, external parameters used in the tropospheric air mass factor computation (surface reflectances, cloud parameters and a priori profile shapes of H_2CO) have been updated using most recent databases. Similar updates were also applied to the historical datasets of GOME and SCIAMACHY, leading to the generation of a consistent multi-mission H_2CO data record covering the time period from 1997 until 2011. Comparing the resulting time series of monthly averaged H_2CO vertical columns in 12 large regions worldwide, the correlation coefficient between SCIAMACHY and GOME-2 columns is generally higher than 0.8 in the overlap period, and linear regression slopes differ by less than 10 % from unity in most of the regions. In comparison to SCIAMACHY, the largely improved spatial sampling of GOME-2 allows for a better characterisation of formaldehyde distribution at the regional scale and/or at shorter timescales, leading to a better identification of the emission sources of non-methane volatile organic compounds.

1 Introduction

Atmospheric formaldehyde (H_2CO) is an intermediate product common to the degradation of many volatile organic compounds. While the global formaldehyde background is due to methane oxidation, emissions of non-methane volatile organic compounds (NMVOCs) from biogenic, biomass burning and anthropogenic continental sources result in important and localised enhancements of the H_2CO concentration. Monitoring the spatial and temporal variability of NMVOC emissions is essential for a better understanding of the processes that not only control the production and the evolution of tropospheric ozone, playing a key role in air quality and climate change, and also of the hydroxyl radical OH and secondary organic aerosols. For these reasons, H_2CO satellite observations have been increasingly used in

combination with tropospheric chemistry transport models to constrain NMVOC emissions in top-down inversion approaches (e.g. Palmer et al., 2006; Fu et al., 2007; Millet et al., 2008; Stavrakou et al., 2009b, c; Curci et al., 2010; Barkley et al., 2011; Fortems-Cheiney et al., 2012; Marais et al., 2012).

In emission regions, the bulk of formaldehyde lies in the boundary layer, which can only be sounded from space using UV-Vis nadir instruments. Formaldehyde tropospheric concentrations have been successfully retrieved from the successive mid-morning polar orbiting sensors operated by ESA on the ERS-2 and ENVISAT platforms, i.e. GOME (1996–2003) and SCIAMACHY (2002–2012) (Chance et al., 2000; Palmer et al., 2001; Wittrock et al., 2006; De Smedt et al., 2008). In addition, the OMI sensor launched on the NASA AURA platform in 2004 provides complementary H_2CO column measurements in the early afternoon (Kurosu, 2008; Millet et al., 2008). The present paper focuses on formaldehyde retrievals from the second Global Ozone Monitoring Experiment (GOME-2), which was launched in October 2006 on board the Meteorological Operational satellite-A (MetOp-A). Being part of the EUMETSAT Polar System (EPS) that represents the European contribution to the Initial Joint Polar-Orbiting Operational Satellite System (IJPS), the mission consists of a series of three MetOp satellites, planned to be successively launched at approximately five-year intervals (Callies et al., 2000). The measurements made by the three GOME-2 instruments therefore have the potential to extend the successful GOME and SCIAMACHY H_2CO mid-morning data record by more than a decade (Vrekoussis et al., 2010; De Smedt, 2011). Moreover, the higher spatial sampling of GOME-2 in comparison to GOME and SCIAMACHY is expected to allow for a better identification of the spatial structures of short-lived tropospheric species like H_2CO, and for a more precise determination of the variability of tropospheric emissions (Lerot et al., 2010; Richter et al., 2011). Satellite observations of H_2CO will be further extended by TROPOMI, the successor of OMI, to be launched in 2015 as part of the GMES Sentinel 5 precursor mission. Owing to its much improved spatial resolution, this instrument will offer an unprecedented view of the spatiotemporal variability of the tropospheric emission. It will be followed by the Sentinel-4 on the geostationary Meteosat Third Generation mission, and at the 2020 horizon by the Sentinel-5 to be operated on the METOP Second Generation platform.

H_2CO columns can be retrieved in the near ultraviolet (UV) using the differential optical absorption spectroscopy (DOAS) technique (Platt and Stutz, 2008), which involves two main steps. First, the effective slant column amount (corresponding to the integrated H_2CO concentration along the mean atmospheric optical path) is derived through a least-squares fit of the measured Earth reflectance spectrum to laboratory absorption cross-sections. Second, slant columns are converted into vertical columns by means of air mass factors (AMF) obtained from suitable radiative transfer calculations,

accounting for the presence of clouds and aerosols, surface properties and best-guess H_2CO vertical profiles. In the UV, the sensitivity to H_2CO concentration in the boundary layer is intrinsically limited from space due to the combined effect of Rayleigh and Mie scattering that hampers the penetration of the solar radiation in the lowest atmospheric layers. As a result, the retrieval of formaldehyde from space is noise sensitive and error prone. While the precision (or random uncertainty) is driven by the signal-to-noise ratio of the recorded spectra, the accuracy (or systematic uncertainty) is limited by the current knowledge of the external parameters needed in the different retrieval steps (De Smedt et al., 2008). To fully exploit the potential of satellite data, applications relying on tropospheric H_2CO observations require high-quality long-term time series, provided with well characterised errors and averaging kernels, and retrieved consistently from the different sensors.

In the present paper, we concentrate (1) on optimising the H_2CO retrieval settings for the GOME-2 instrument and (2) on bringing the resulting data product into harmonisation with the historical GOME and SCIAMACHY time series. Air mass factor parameters are updated by making use of new a priori databases of improved spatial and temporal resolutions. After a brief introduction to the GOME-2 instrument (Sect. 2), H_2CO columns retrieval issues are discussed and improved DOAS settings are introduced in Sect. 3. Section 4 discusses the impact of the updates brought to the air mass factor calculations. The resulting GOME-2 H_2CO vertical column data product is then presented in Sect. 5. Uncertainties are characterised and, to establish the link with historical datasets, the consistency of GOME-2 and SCIAMACHY H_2CO data is investigated for the period from January 2007 to December 2011 in Sect. 6.

2   GOME-2 measurements

The MetOp-A satellite was launched in October 2006 in a sun-synchronous polar orbit with an equator crossing time of 09:30 LT (descending node) and a repeat cycle of 29 days. GOME-2 is an improved version of the GOME instrument on the ERS-2 satellite (Callies et al., 2000; Munro et al., 2006). It is a nadir-scanning UV-VIS spectrometer with four main optical channels, covering the spectral range between 240 and 790 nm with a spectral resolution between 0.26 nm and 0.51 nm (FWHM). Additionally, two polarisation components are measured with polarisation measurement devices (PMDs) at 15 broadband channels covering the full spectral range.

GOME-2 measures the solar radiation backscattered by the atmosphere and reflected from the surface of the Earth in a nadir-viewing geometry. The direct sun spectrum is also measured via a diffuser plate once per day. An important improvement of the GOME-2 instrument compared to GOME/ERS-2 is the use of a quartz quasi-volume diffuser

for the direct sun measurements. The sun-angle dependent differential structures in the bi-directional scattering distribution function (BSDF) for this diffuser is strongly reduced compared the ground aluminium diffuser as used in GOME/ERS-2 (Valks et al., 2011).

The default swath width of the GOME-2 scan is 1920 km, twice as large compared to GOME and SCIAMACHY, allowing for global Earth coverage within 1.5–3 days at the Equator. The along-track dimension of the instantaneous field of view is ∼ 40 km, while the across-track dimension depends on the integration time used for each channel. For the nominal swath of 1920 km and integration time of 187.5 ms, the ground pixel size is 80×40 km^2. The GOME-2 spatial resolution is therefore finer than GOME (320×40 km^2) but coarser than SCIAMACHY (60×30 km^2) and OMI (13×24 km^2 at true nadir).

For this work, we have used the EUMETSAT GOME-2 level 1B data v4.11 from January 2007 to December 2010 and v5.12 from 2011 onwards.

3 Formaldehyde slant column retrieval

3.1 Initial baseline settings

H$_2$CO slant columns have been retrieved from the GOME and SCIAMACHY sensors using analysis settings developed in De Smedt et al. (2008, 2010). In a first step of the present work, these retrieval settings were applied to GOME-2 without any further modification. This constitutes our initial slant column retrieval baseline.

As detailed in Table 1, the H$_2$CO absorption features are fitted to the Meller and Moortgat (2000) laboratory measurements in the 328.5–346 nm wavelength range. In this interval, the absorption cross-sections of O$_3$ at 228 K and 243 K, NO$_2$ and BrO are also included. The weak interference by the oxygen collisional complex O$_4$ is not explicitly treated but subsumed in a polynomial of fifth order that also accounts for other broadband contributions to the atmospheric attenuation such as Rayleigh and Mie scattering. To correct for the Ring effect, two pseudo absorption cross-sections generated according to Vountas et al. (1998) are used. All reference datasets are convolved using the wavelength-dependent GOME-2 slit function determined from pre-flight measurements (Siddans et al., 2006). The daily solar irradiance is used as reference spectrum to calculate the atmospheric optical density, which in principle allows for the retrieval of absolute slant column densities. However, residual biases due to unresolved spectral interferences remain a limiting factor for the retrieval of weak absorbers such as H$_2$CO. These are compensated using a reference sector normalisation approach (De Smedt et al., 2008) so that, in practice, the end product of the slant column retrieval procedure is a differential slant column. The reference sector is chosen in the central Pacific Ocean (140°–160° W), where the only significant
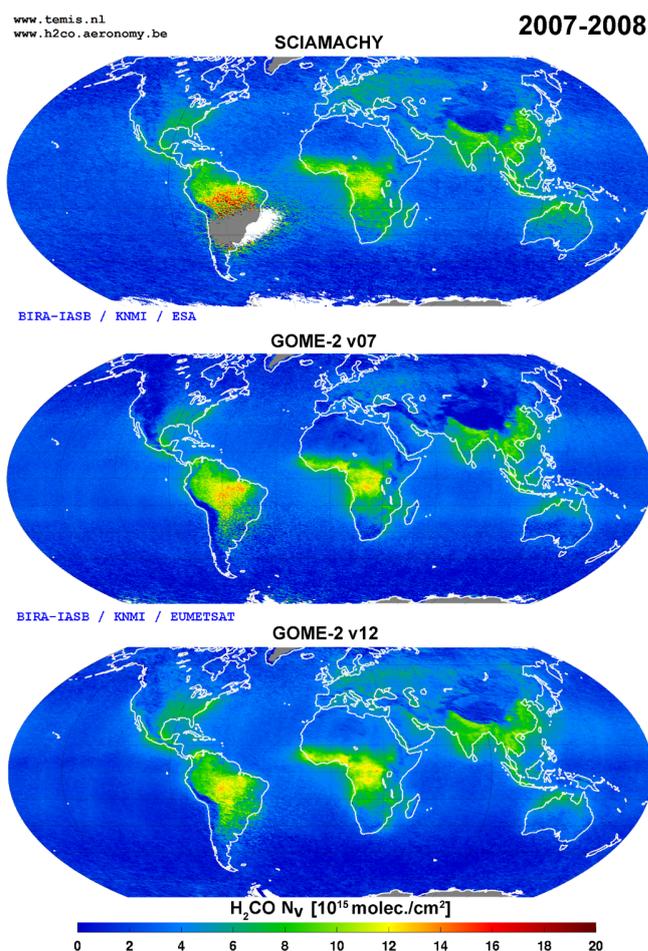

Fig. 1. H$_2$CO vertical column retrieved from SCIAMACHY and GOME-2 between 2007 and 2008. The initial retrieval settings used for GOME-2 v07 are based on the GOME and SCIAMACHY settings (De Smedt et al., 2008, 2010), while v12 includes the improved retrieval settings developed for GOME-2 (see text for details).

source of H$_2$CO is methane oxidation. On a daily basis, the latitudinal dependency of the H$_2$CO slant columns in the reference sector is modelled by a polynomial and subtracted from the slant columns ($\Delta N_s = N_s - N_{s0}$).

The final H$_2$CO vertical column (N_v) is obtained using the following equation:

$$N_v = \frac{\Delta N_s}{M} + N_{v0}^{\mathrm{CTM}}, \tag{1}$$

where ΔN_s is the differential slant column, M is the tropospheric air mass factor, and N_{v0}^{CTM} is the H$_2$CO background obtained from the tropospheric 3D-CTM IMAGES (Stavrakou et al., 2009b) in the reference sector. The tropospheric AMF calculation details are presented in Sect. 3.

Note that these formaldehyde retrieval settings have been implemented in the SCIAMACHY Level 2 operational processing environment at the German Aerospace center (DLR)

Table 1. Initial DOAS settings (v07) used to retrieve H_2CO slant columns from GOME-2 spectra.

Fitting interval	328.5–346 nm
Absorption cross-sections	
H_2CO	Meller and Moortgat (2000), 298 K
Ozone	Brion et al. (1998), 228 K and 243 K
BrO	Fleischmann et al. (2004), 223 K
NO_2	Vandaele et al. (2002), 220 K
Ring effect	2 vectors generated using SCIATRAN (Rozanov et al., 2001; Vountas et al., 1998)
Slit function	GOME-2 pre-flight slit function depending on the detector wavelength (Siddans et al., 2006)
Polynomial	5th order
Intensity offset correction	Linear offset
Reference spectrum (I_0)	Daily solar spectrum measured by GOME-2

(SGP OL version 6)[1], as well as in the GOME-2 operational data processor (GDP version 4.4)[2] developed for the EU-METSAT Satellite Application Facility on Ozone and Atmospheric Chemistry Monitoring (O3M SAF) at DLR. Hereafter, the H_2CO columns retrieved using the initial settings are referred to as GOME-2 H_2CO v07, while the label v12 is assigned to results obtained using improved settings described below.

The global distribution of the v07 H_2CO vertical columns derived over the 2007–2008 period is displayed in Fig. 1 (second panel) and compared with similarly sampled SCIA-MACHY results (first panel). The tropospheric AMFs have been consistently calculated using the same input parameters for both sensors, as further described in Sect. 4. The GOME-2 columns are slightly lower than those of SCIAMACHY, mainly over mid-latitude continental regions. As expected from the much better sampling of GOME-2 observations in comparison to SCIAMACHY, the 2-yr averaged GOME-2 map is significantly smoother. Moreover, over eastern South America, the GOME-2 H_2CO retrievals are clearly less affected by the South Atlantic anomaly. This feature of the GOME-2 H_2CO product is probably the result of a better instrumental shielding against extra-terrestrial high energy particles. Note that the Amazon region is characterised by large and particularly uncertain biogenic emissions of NMVOCs (Barkley et al., 2008, 2011; Müller et al., 2008; Stavrakou et al., 2009a, b, 2010), and therefore GOME-2 appears to be particularly well suited to address these issues.

Figure 2 compares time series of monthly averaged SCIA-MACHY and GOME-2 H_2CO differential slant columns ΔN_s derived over continents in the period from 2007 until 2011 in three latitude bands from $10°$ S to $50°$ N. For these comparisons, GOME-2 pixels having line-of-sight angles smaller than $32°$ (similar to SCIAMACHY) were selected. The GOME-2 H_2CO ΔN_s obtained with the initial settings (v07, maroon line) agree well with those of SCIAMACHY
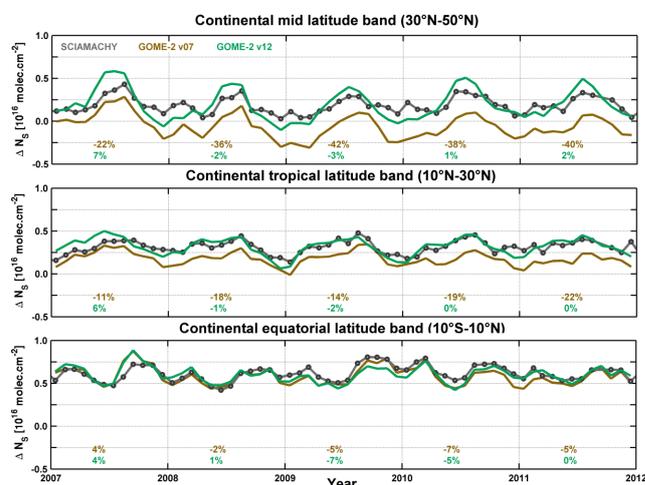
Fig. 2. Monthly H_2CO differential slant columns averaged in 3 latitude bands over continents, as retrieved from SCIAMACHY and GOME-2 measurements. The maroon line shows the GOME-2 results obtained with the initial retrieval settings (v07), while the green line shows the results obtained with the improved settings (v12, see text for details). Averaged relative differences between GOME-2 and SCIAMACHY observations are given at the bottom of each subplot, for each year and for the two versions of the GOME-2 results.

in the equatorial continental areas where the H_2CO concentrations are the largest. However, a small systematic negative bias of -5 % in average is observed. The amplitude of this negative bias increases with latitude and time, ranging from -11 % in 2007 in the tropical band, to -40 % in 2011 at mid-latitudes.

The scatter of the individual GOME-2 H_2CO measurements has also been compared to the scatter obtained with GOME and SCIAMACHY. As shown in De Smedt et al. (2008), the single-pixel H_2CO retrievals are shot-noise limited. Since GOME, SCIAMACHY and GOME-2 have similar instrumental designs, and since the H_2CO retrieval settings have been aligned, differences in the standard deviations in the retrieved slant columns should be proportional
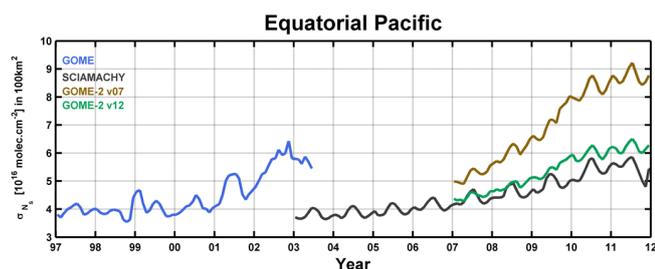
Fig. 3. H_2CO slant column standard deviation scaled to a pixel size of $10 \times 10\,km^2$, retrieved from GOME, SCIAMACHY and GOME-2 over the equatorial Pacific. Two versions of GOME-2 results are shown: the initial retrieval settings (v07) and the improved settings (v12, see text for details).

to the square root of the different instrument ground pixel areas. The spread of the differential slant columns in the clean equatorial Pacific has been analysed over the entire time period of measurements for each instrument. In this area, one can assume that the H_2CO production due to NMVOC oxidation can be neglected, and that the daily variations of the CH_4 oxidation are weak. The standard deviation of the retrieved H_2CO columns (σ_{N_s}) is therefore a measure of the random noise of the measurements. In order to compare the instrument performances, the standard deviations have been corrected for each pixel area and scaled to a common pixel area of $10 \times 10\,km^2$, more comparable to OMI or to the future TROPOMI spatial resolution (Veefkind et al., 2012). The results of this analysis are shown in Fig. 3. At the beginning of their measurement periods, the scaled standard deviations found for GOME, SCIAMACHY and GOME-2 v07 are respectively 3.8, 3.7 and 5.2×10^{15} molec. cm^{-2}. The standard deviations of the GOME-2 v07 results are therefore 35 % larger than expected. Note that a similar noise excess has been identified in GOME-2 NO_2 retrievals at visible wavelengths (Richter et al., 2011). Moreover, while the rate of increase of the noise of the GOME and SCIAMACHY H_2CO ΔN_s is less than 3 % yr^{-1} during their first five years of operation (6 % yr^{-1} for the next five years of SCIAMACHY), the rate of increase of the noise for GOME-2 v07 reaches 14 % yr^{-1} between 2007 and 2011 (from 5.2 to 9×10^{15} molec. cm^{-2}). This behaviour is related to the fast throughput degradation of GOME-2 documented in Lang et al. (2009).

The aforementioned discrepancies (bias and noise excess) prompted us to further investigate and tentatively improve our slant column retrievals, leading to the modified GOME-2 H_2CO retrieval settings presented below.

3.2 Improved DOAS retrieval

A common way to maximize the sensitivity to weak absorbers and reduce the noise in the retrieved slant columns is to increase the size of the fitting window, including more
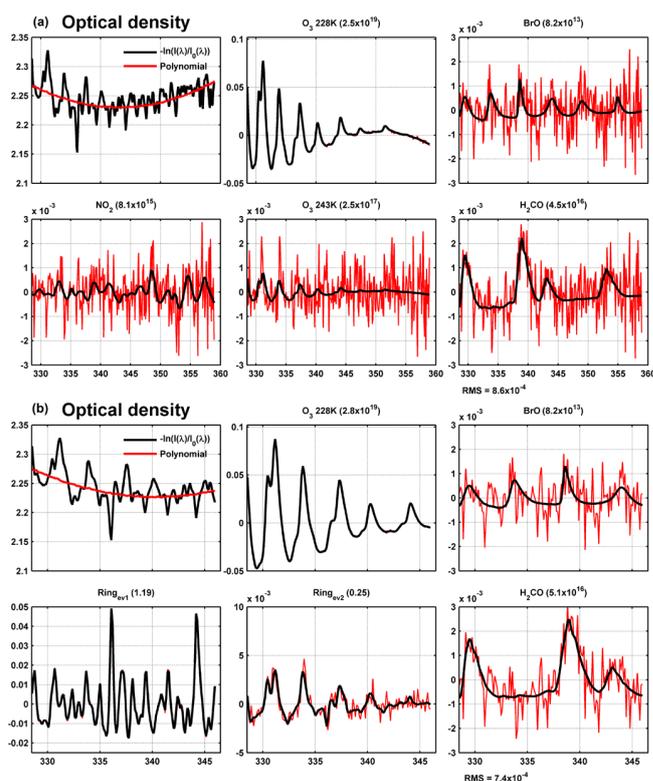

Fig. 4. Illustration of the 2-step DOAS fit of a GOME-2 spectrum overpassing a region of large biogenic emissions in southeastern US (9 August 2007, orbit 4176, SZA: 28°). **(a)** First step in the wavelength interval 328.5–359 nm, **(b)** second step in the wavelength interval 328.5–346 nm. The BrO slant column is not fitted in the second step. Each plot shows the optical density in black [no unit] and the residuals in red. The fitted slant columns are given above [molec. cm^{-2}]. The root mean square of the fit (RMS) in each wavelength interval is given at the bottom of the figure [no unit].

absorption bands and thereby increasing the information content in the DOAS analysis (see e.g. Richter et al., 2011). Although the flexibility to extend the H_2CO fitting interval has been shown to be limited at longer wavelengths by a known O_4-related artefact above arid regions and oceans (De Smedt et al., 2008), and at shorter wavelength by the strong ozone absorption, we decided to investigate the potential of using a larger fitting range to possibly reduce the known correlation between H_2CO and BrO absorption features (Theys et al., 2011). Since, in contrast to SCIAMACHY, GOME-2 does not suffer from strong polarization-related spectral features in channel 2, various fitting windows could be tested at both longer and shorter wavelengths beyond the limits of the baseline 328.5–346 nm window. Note that a similar approach has been adopted in Theys et al. (2011) with a focus on BrO retrieval. Here, our approach uses a two-step DOAS fit retrieval that effectively reduces the noise in H_2CO. Firstly, BrO slant columns are fitted in a wide wavelength interval (328.5–359 nm) that includes six BrO absorption bands and

Table 2. Improved DOAS settings (v12) used to retrieve H_2CO slant columns from GOME-2 spectra.

Fitting interval – 1	328.5–359 nm
Absorption cross-sections	
$\quad$ H_2CO	Meller and Moortgat (2000), 298 K
$\quad$ Ozone	Brion et al. (1998), 228 K and 243 K
$\quad$ BrO	Fleischmann et al. (2004), 223 K
$\quad$ NO_2	Vandaele et al. (2002), 220 K
Polarisation vectors	Eta and Zeta from the GOME-2 calibration key data (EUMETSAT, 2009)
Fitting interval – 2	328.5–346 nm
Absorption cross-sections	
$\quad$ H_2CO	Meller and Moortgat (2000), 298 K
$\quad$ Ozone	Brion et al. (1998), 228 K and 243 K
$\quad$ BrO (not fitted)	Fleischmann et al. (2004), 223 K
$\quad$ NO_2	Vandaele et al. (2002), 220 K
Ring effect	2 vectors generated using SCIATRAN (Rozanov et al., 2001; Vountas et al., 1998)
Slit function	Asymmetric Gaussian slit function fitted during the irradiance calibration (Cai et al., 2012)
Polynomial	5th order
Intensity offset correction	Linear offset
Reference spectrum (I_0)	Daily solar spectrum measured by GOME-2

minimises the correlation with H_2CO. H_2CO columns are then retrieved in the original 328.5–346 nm interval, but using the BrO slant column values determined in the first step (see Table 2 for details). This approach allows to efficiently decorrelate BrO from H_2CO retrievals while, at the same time, the O_4-related bias in arid regions is avoided.

Figure 4 illustrates our two-step procedure in the case of a GOME-2 spectrum measured on 9 August 2007 and overpassing a strong H_2CO signal due to isoprene oxidation in the southeastern United States. Note that in the 328.5–359 nm interval, two additional polarisation functions (Eta and Zeta, from the GOME-2 calibration key data; EUMETSAT, 2009) are included in the analysis to correct for residual calibration sensitivity issues at the edges of the scan. Note also that the BrO columns obtained in this interval have been systematically checked and found to be consistent with results published in Theys et al. (2011). More particularly, the BrO columns do not present any interference with elevated H_2CO columns in tropical regions.

Furthermore, in order to better cope with the strong ozone absorption at wavelengths shorter than 336 nm, the method of Puķīte et al. (2010) has been applied. It is a Taylor series approach that describes the O_3 slant column as function of wavelength and of vertical optical depth. At the first order, the method consists in adding two cross-sections to the fit: $\lambda\sigma_{O_3}$ and $\sigma_{O_3}^2$ (Eq. 11 in Puķīte et al., 2010). This correction, which is equivalent to the AMF-modified DOAS approach previously used in Theys et al. (2011) and De Smedt et al. (2011), allows reducing the fitting residuals and the H_2CO slant columns underestimation at large solar zenith angles, when the stratospheric ozone absorption becomes optically thick.
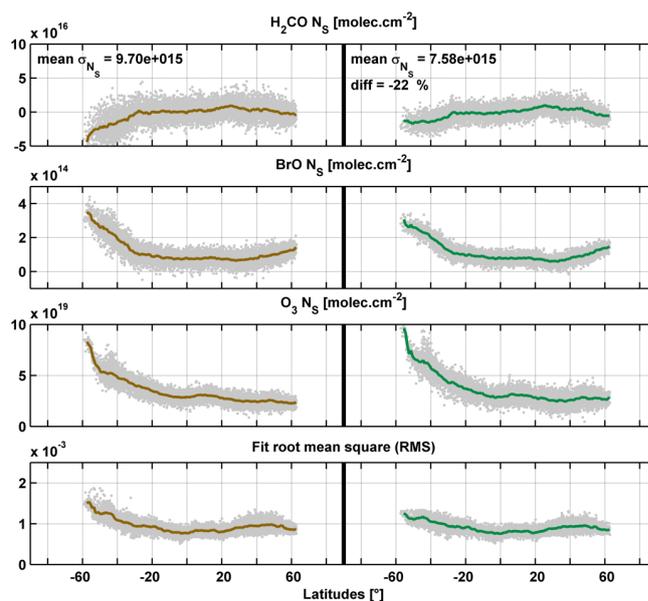

Fig. 5. H_2CO, BrO and O_3 slant columns, and root mean square of the DOAS fit for one GOME-2 orbit (13 August 2007, orbit number 4227). The left column corresponds to the initial retrieval settings (v07), while the right column shows the effect of the BrO and O_3 corrections (v12).

Figure 5 illustrates the effects of the two corrections (dual fit for BrO and H_2CO, and O_3 correction) for one particular GOME-2 orbit on 13 August 2007, which is typical of background level H_2CO concentrations. The upper panels of Fig. 5 show the H_2CO differential slant columns before and after the corrections (left and right column); the mean standard deviation σ_{N_s} are inset each plot. The second and third

panels show the BrO and O_3 slant columns, while the lower panels show the root mean square (RMS) of the DOAS fit. Owing to the first correction (dual fit for BrO and H_2CO), the scatter in the H_2CO slant columns is reduced by about 20 %, while the second correction (O_3 AMF correction) allows for a slight reduction of the RMS below 40° S, and a significant reduction of the H_2CO slant columns underestimation at southern mid-latitudes, directly related to the larger O_3 columns in the Southern Hemisphere in August.

3.3 Improved characterisation of the GOME-2 slit function

Although the GOME-2 slit function has been thoroughly characterised before the launch of the instrument (Siddans et al., 2006), recent investigations (Lacan and Lang, 2011; Dikty and Richter, 2011) have shown that the width of the GOME-2 slit function has been narrowing with time. To study the impact of these changes on our H_2CO retrievals, effective slit functions have been derived from measured solar irradiance spectra by adjustment to the high resolution solar reference of Chance and Kurucz (2010) and assuming an asymmetric Gaussian line-shape described by the following equation (Cai et al., 2012):

$$
\begin{aligned}
f(x) = & (1 - \delta(x))e^{-4\ln(2)\frac{x^2}{(w(1-a))^2}} \\
& + \delta(x)e^{-4\ln(2)\frac{x^2}{(w(1+a))^2}},
\end{aligned} \tag{2}
$$

where x is the difference in wavelength and $\delta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$ is a step function. The asymmetric Gaussian line-shape is characterised by w, the full width at half maximum (FWHM) and by a, the asymmetry factor (AF), where a is defined such as the left and right widths at half maximum are $w_L = w(1-a)$ and $w_R = w(1+a)$.

The result of this effective calibration is presented in Fig. 6. Retrieved slit width and asymmetry factor parameters are displayed for the entire period from 2007 to 2011, and for the wavelengths 330, 340, 350 and 360 nm. In agreement with Dikty and Richter (2011), we find that in 2011, after 5 yr of operations, the GOME-2 slit function has narrowed by about 8 % compared to its value at launch. Note that transient changes in the slit function visible in Fig. 6 coincide with peaks of the instrument throughput, related to operations performed on the GOME-2 instrument[34]. In contrast to the slit width, the asymmetry factor is more stable, showing only localised instabilities, particularly after the major throughput test performed in September 2009 (EUMETSAT
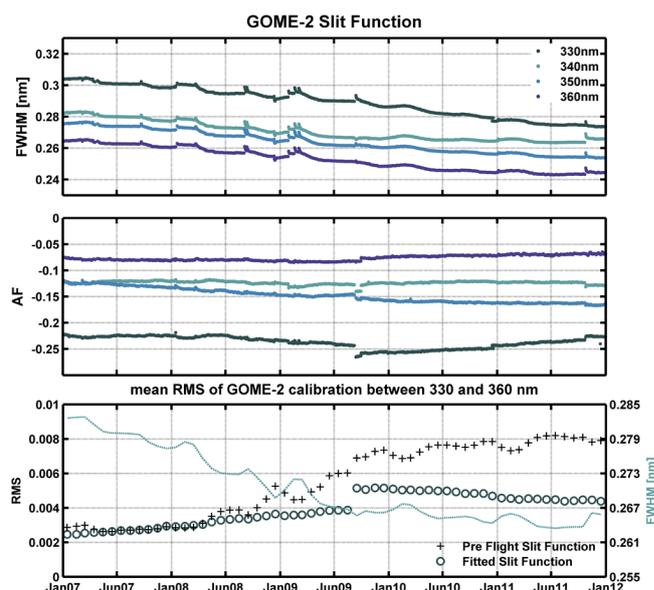

Fig. 6. (a) GOME-2 slit function width (FWHM) and asymmetry factor (AF) fitted during the calibration procedure of the DOAS analysis. **(b)** Mean residuals of the GOME-2 solar spectrum calibration between 330 and 360 nm, using the pre-flight slit function or fitting a Gaussian asymmetric slit function. The FWHM of the fitted Gaussian asymmetric slit function is shown as second y-axis.

Product and Service News, 2009)[5]. The lower panel of Fig. 6 compares the residuals of the calibration procedure obtained using (1) the pre-flight slit function and (2) fitting an asymmetric Gaussian line-shape. One can see that these residuals correlate well with the width of the slit function plotted as second y-axis in the same figure. It also appears that the calibration results are markedly worse just after the throughput test of September 2009, suggesting that this test also had an impact on the shape of the slit function.

3.4 Summary of the slant column improvements

The new retrieval settings (v12) are detailed in Table 2. Figure 7 sums up the global effect of the different retrieval improvements in comparison to the initial settings (v07). The first panel shows a comparison of the monthly averaged H_2CO differential slant columns on the global scale. The final H_2CO ΔN_s (v12) are higher by 16 % in 2007 and by 50 % in 2011. As a result, the general underestimation of the GOME-2 columns compared to SCIAMACHY is significantly reduced. At the same time, the impact of the instrumental degradation is minimised. This is due in part to the BrO and O_3 corrections, and for the rest to the fit of the slit function. The improvement of v12 over v07 is also illustrated in Figs. 1 and 2, which show a good agreement between SCIAMACHY and GOME-2 ΔN_s at all latitudes.

---

[3]GOME-2 Long Term Monitoring Report (G2RP-R1 PPF4.0).

[4]GOME-2/Metop-A instrument, PPF and auxiliary-data change history, Ref.: EUM/OPS-EPS/TEN/08/0716, v6 Draft, Date: 11 January 2012.

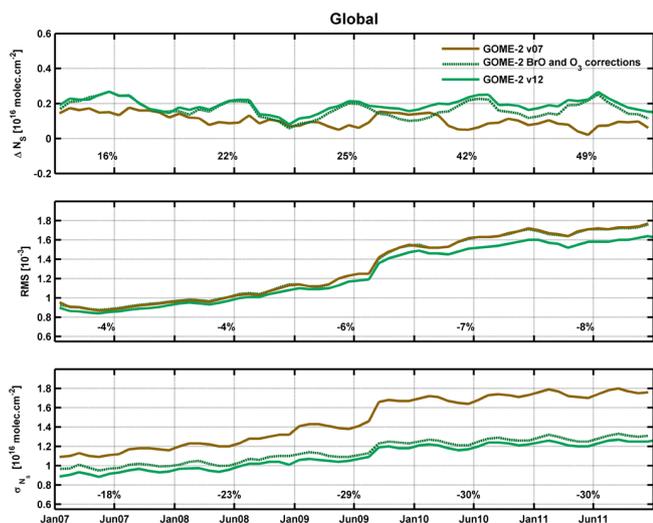[5]GOME-2 Products Guide, Ref.: EUM/OPS-EPS/MAN/07/0445, Issue: v2D, Date: 6 March 2009.

Fig. 7. Globally averaged time evolution of the normalized H_2CO slant colums (first panel); the root mean square of the fit (second panel); and the H_2CO slant columns standard deviation (third panel) for the intitial settings (v07), the BrO and O3 corrections, and the final improved GOME-2 retrieval settings, including the fit of the GOME-2 slit function (v12). Yearly averaged relative differences between GOME-2 v07 and v12 results are given at the bottom of each subplot.

Also, the effect of the South Atlantic anomaly has been reduced in v12 compared to v07.

The second panel of Fig. 7 compares the root mean squares (RMS) of the fits. While the BrO and O3 corrections have little impact on the RMS, the fit of the slit function decreases the RMS by 4 % in 2007. This effect becomes more important in time (8 % in 2011) and is expected to continue to grow.

The third panel of Fig. 7 clearly shows that the reduction of the noise in the H_2CO slant columns is largely driven by the pre-fit of BrO although it is reinforced by the fit of the slit function. The noise reduction ranges from 18 % in 2007 to 30 % in 2011. This effect is also shown in Fig. 3, where the random noise of the GOME-2 v12 slant columns in 2007 is only 17 % larger than the nominal values of GOME and SCIAMACHY (instead of 35 % larger with v07). Importantly, the impact of the instrumental degradation has been reduced in v12 since the degradation rate of the random noise has been reduced by a factor 2 (from 14 % yr^{-1} to 7 % yr^{-1}).

4 Air mass factors

The method used to calculate the tropospheric H_2CO air mass factors has been described in detail in De Smedt et al. (2008). Briefly, AMFs are computed using vertically resolved scattering weighting functions and a priori profile shapes following the formulation of Palmer et al. (2001). In the case of cloudy scenes, a correction is applied using a Lambertian reflecting cloud model and the independent pixel

approximation (Martin et al., 2002). No explicit correction is applied for aerosols but the cloud correction scheme accounts for a large part of the aerosol scattering (Boersma et al., 2004, 2011; De Smedt et al., 2008). In all cases, observations with effective cloud fractions exceeding 40 % are filtered out. For larger cloud fractions, the error in the vertical column increases rapidly. Furthermore, the information content below the cloud altitude is weak because the column is dominated by a priori profile information, as can be deduced from the averaging kernels that are provided for each observation (De Smedt, 2011). The AMF variation with wavelength is found to be generally less than 5 % in the 328.5–346 nm wavelength range and a single air mass factor representative for the entire retrieval interval can be used at 340 nm. However, for solar zenith angles larger than 60°, the increase of stratospheric ozone absorption leads to air mass factor variations of up to 10 % within the wavelength interval and, again, the sensitivity of the measurement in the lower troposphere decreases rapidly. For these reasons, the data are filtered out for solar zenith angles larger than 60°.

In this work, the scattering weighting functions are evaluated with the LIDORT v3.0 radiative transfer model (Spurr, 2001, 2008) for a number of representative viewing geometries, surface albedos and altitudes, and stored in a look-up table. To reduce the errors associated to topography, scattering weighting functions are calculated with a fine altitude grid near the ground (200 m sampling up to 2 km), and the a priori profile shapes are rescaled to the ground altitude following Zhou et al. (2009). To take advantage of the better GOME-2 sampling rate, the precision of the external parameters used to calculate the air mass factors has been improved as much as possible. The same changes have been applied to the GOME and SCIAMACHY air mass factors in order to ensure full consistency between the time series. Compared to the previous version of the product (De Smedt et al., 2010), the albedo climatology, the cloud data version and the a priori profile shapes have been updated.

The monthly albedo climatology of Kleipool et al. (2008) is now used (at 342 nm). It provides the mode (most frequently observed value) of the Lambertian equivalent reflectances derived from OMI observations at the spatial resolution of 0.5°. This replaces the monthly climatology of Koelemeijer et al. (2003), which provided the minimum of the Lambertian equivalent reflectances derived from GOME measurements at the resolution of 1° (at 335 nm). The impact of this change on the vertical columns of H_2CO is illustrated in the two first panels of Fig. 8, which show the vertical columns differences averaged over two seasons in 2007 (DJF and JJA). These differences have to be considered relative to the corresponding seasonal vertical columns shown in Fig. 9. For the DJF period, the new albedo climatology leads to H_2CO columns systematically lower over eastern China and North equatorial Africa by about -2.5 and -1×10^{15} molec. cm^{-2}, respectively (-40 % and -10 % of the total columns). During the JJA period, systematically
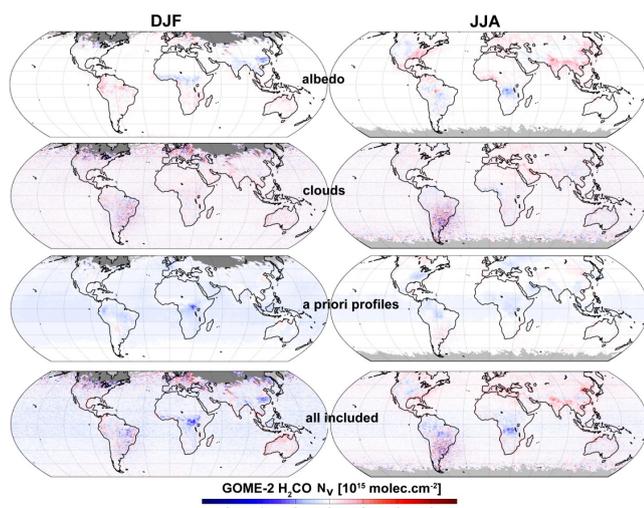
Fig. 8. Absolute differences in H$_2$CO vertical columns (VC) from air mass factor parameter updates (new version-previous version). The left and right columns show the seasonally averaged VC differences, respectively in DJF and in JJA 2007. First row: update of the surface albedo climatology (previous version: Koelemeijer et al., 2003, new version: Kleipool et al., 2008). Second row: update of the cloud product (previous version: Fresco v5, new version: Fresco v6). Third row: update of the a priori profiles from IMAGESv2 (previous version: monthly profiles, $5 \times 5°$, new version: daily profiles, $2 \times 2.5°$). The two lower panels show the combined effects of the three AMF parameter updates.

higher columns are derived over eastern US, India and Eastern Asia (2 to 3×10^{15} molec. cm^{-2}, 15 to 30 %), while lower columns are observed in Africa, south of the Equator (-2×10^{15} molec. cm^{-2} or -15 %). Over the Mediterranean Sea, vertical columns are 10 % higher during summertime due to lower reflectances over waters in the Kleipool climatology compared to Koelemeijer.

Regarding the cloud parameters, version 6 of the FRESCO cloud product now replaces version 5 (Koelemeijer et al., 2001; Wang et al., 2008). Both versions provide an effective cloud fraction and cloud top height, assuming a Lambertian-reflecting cloud with an albedo of 0.8. The main difference lies in the surface albedo climatology used for the cloud retrieval, which was the Koelemeijer climatology in version 5 and has been replaced by the MERIS climatology (Popp et al., 2011) over continents in the new version. As can be seen in the second row panels of Fig. 8, the mean impact of this new cloud version is relatively small. A stronger effect can be observed locally over Iraq and Syria, Northern Australia and generally along the coasts, reflecting the better spatial resolution of the MERIS database. Note that the MERIS albedo database does not provide measurement in the UV and cannot be used as albedo climatology for the H$_2$CO retrieval.

Finally, the monthly a priori H$_2$CO vertical profiles simulated by the IMAGESv2 tropospheric CTM at the resolution of $4° \times 5°$ (Stavrakou et al., 2009a) have been replaced

by daily profiles provided by an updated version of IMAGESv2 with a spatial resolution of $2 \times 2.5°$ (Stavrakou et al., 2009b, 2011). This new version of IMAGES provides the global distribution of 80 chemical components for 40 vertical levels between the Earth's surface and 22.5 km altitude. Advection is driven by monthly mean operational ECMWF fields, and daily fields are used for temperature, water vapour, boundary layer mixing, and cloud optical depths. The model time step is 6 h, but diurnal variations in the photorates and in the concentrations are accounted for through correction factors computed via a diurnal cycle simulation with a 20-min time step. The diurnal variations are used to estimate the formaldehyde H$_2$CO model profiles at the overpass time of GOME-2. The chemical mechanism comprises 20 explicit NMVOCs. The degradation mechanism for the majority of the NMVOCs is mostly based on the Master Chemical Mechanism (MCM) (Saunders et al., 2003), whereas for isoprene, the Mainz Isoprene Mechanism (MIM version 2, Taraborelli et al., 2009) has been used. The a priori biogenic emissions of isoprene are obtained from the MEGAN-ECMWF inventory (Müller et al., 2008, and update in 2010, http://tropo.aeronomie.be/models/isoprene.htm). The a priori pyrogenic emissions are provided by the Global Fire Emission Database (GFEDv3, van der Werf et al., 2010) by applying updated emission factors (2007) from Andreae and Merlet (2001). The a priori anthropogenic emissions are obtained from the RETRO database for the year 2000 (Schultz et al., 2007), which is overwritten by EMEP over Europe, and REAS over Asia (Ohara et al., 2007). The impact of using this new version of IMAGES to estimate the a priori profile shapes of H$_2$CO, and also the reference sector correction (N_{v0}^{CTM}), is illustrated in the third line panels of Fig. 8. The global background is lower by 0.5×10^{15} molec. cm^{-2} during the December-January-February period. Over central Africa and western South America, the H$_2$CO vertical columns are lower in average by about 4×10^{15} molec. cm^{-2} (less than 15 % of the column in these period/regions), reflecting a priori profile shapes larger in the free troposphere than in the previous version. During the June-July-August period, the new H$_2$CO columns are slightly higher around Beijing (by about 2×10^{15} molec. cm^{-2}), reflecting profile shapes peaking more near the surface in this urban area, owing to the better spatial resolution of the new simulated profiles.

The lower line panels of Fig. 8 sum up the impacts of these different changes in the air mass factor calculations. The largest effects can be attributed to the new albedo climatology, with contribution of the new a priori H$_2$CO profiles over Africa and around Beijing. In large regions, the averaged differences are relatively moderate compared to the total H$_2$CO columns (lower than 20 %), but more locally, the differences can reach ± 40 %. Overall, the spatial resolution of the product is improved by using better resolved albedo and profile shape climatology, and the temporal resolution is improved by using daily profile shapes. In order to provide
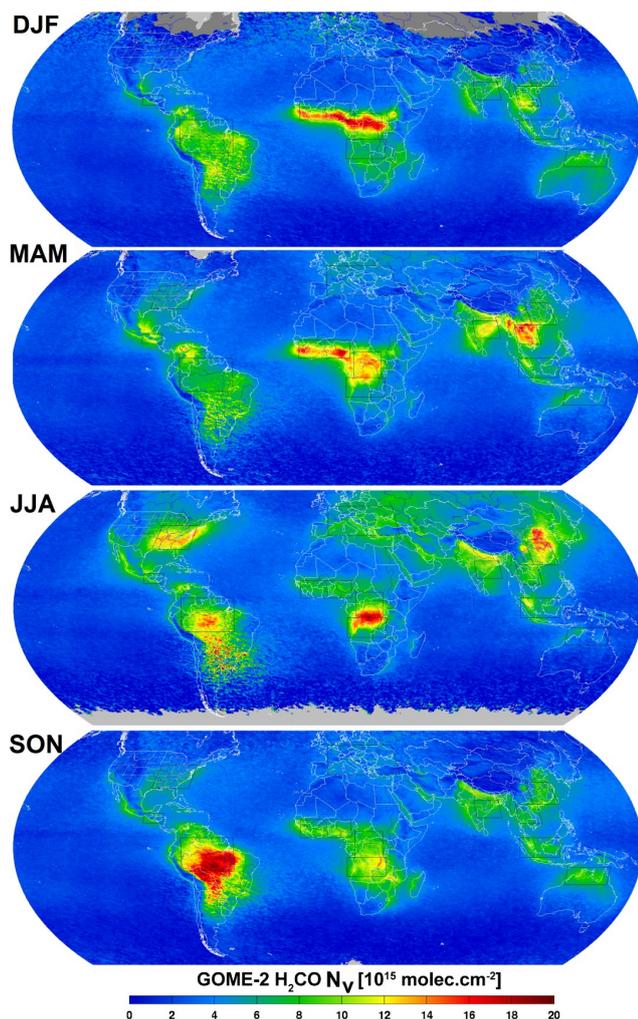
Fig. 9. GOME-2 (v12) seasonal H_2CO vertical columns averaged between 2007 and 2011.

a consistent multi-mission dataset of H_2CO vertical columns between 1997 and 2011, the GOME and SCIAMACHY air mass factors have been recalculated applying the same updates of the external parameters.

5　Error analysis and data product characterisation

Figure 9 presents the GOME-2 H_2CO columns averaged seasonally between 2007 and 2011 on a spatial grid of $0.25°$, while Fig. 10 presents the monthly variations of the H_2CO vertical columns (upper panels) and their relative errors (lower panels) averaged over the period 2007–2011 in the 12 regions defined by the black boxes in Fig. 9. For these figures, pixels have been selected with solar zenith angles below $60°$, cloud fractions lower than 40 % and fitting residuals lower than 1.5×10^{-3}. The method to calculate the total error in the vertical columns of formaldehyde has been described in detail in De Smedt et al. (2008), and is based on

the following equation (assuming that the errors are independent, which is known not to be strictly valid; Boersma et al., 2004):

$$\sigma_{N_v}^2 = \frac{1}{M^2} \times \frac{\sigma_{N_{s,rand}}^2}{N} + \frac{1}{M^2} \times \sigma_{N_{s,syst}}^2$$
$$+ \sigma_{N_{v0}}^{CTM} + \left(\frac{\Delta N_s}{M^2}\right)^2 \times \sigma_M^2, \tag{3}$$

where $\sigma_{N_{s,rand}}^2$ and $\sigma_{N_{s,syst}}^2$ are, respectively, the random and systematic parts of the error in the slant column. The random error in the slant columns is directly related to the residuals of the DOAS fits, while the systematic error takes into account uncertainties of the absorption cross-sections datasets, the correlation between the different absorption cross-sections in the considered wavelength interval, and the absorber concentrations for every observation. N is the number of satellite observations considered in the vertical column average. $\sigma_{N_{v0}}^{CTM}$ is the estimated error in the reference sector correction. Finally, σ_M is the air mass factor error that takes into account uncertainties in the albedo, the cloud parameters and the profile shapes, weighted by the AMF sensitivity to these parameters, that changes for each observation condition. It is understood that AMF errors have systematic and also random components that may average out in space or in time. However, these components can hardly be separated in practice and we consider these uncertainties as systematic. The final error estimated with Eq. (3) is therefore an upper limit of the real error in the vertical columns. The different contributions to the total error are shown separately in Fig. 10. The total error on monthly averaged columns is generally comprised between 30 and 40 %, except in wintertime in mid-latitudes where the error exceeds 60 %. This limitation is explained by the low H_2CO emissions in these regions during wintertime coupled by the fact that the sensitivity of the satellite measurements decreases in the boundary layer when the solar zenith angle increases (increasing the sensitivity of the AMF to any error in the external parameters). In these large regions, the random error in the monthly averaged columns is almost negligible compared to the systematic errors. The systematic uncertainties in slant columns and in air mass factors are generally the largest sources of error, and most of the time of similar magnitudes. However, when the H_2CO columns are lower, the error related to the reference sector correction becomes more significant (up to 30 %). Another important error source of the AMF calculation is the lack of correction in the case of absorbing aerosols, principally in biomass burning conditions. Indeed, it has been shown in several studies that the impact of aerosols on air mass factors is mainly significant when the aerosol layer is above the bulk of formaldehyde, and when the aerosol optical thickness is high, typical of biomass burning conditions (Leitão et al., 2010; Gonzi et al., 2011; Barkley et al., 2012). A full treatment of clouds and aerosols in radiative transfer will only be possible if clouds are aerosols are represented separately as
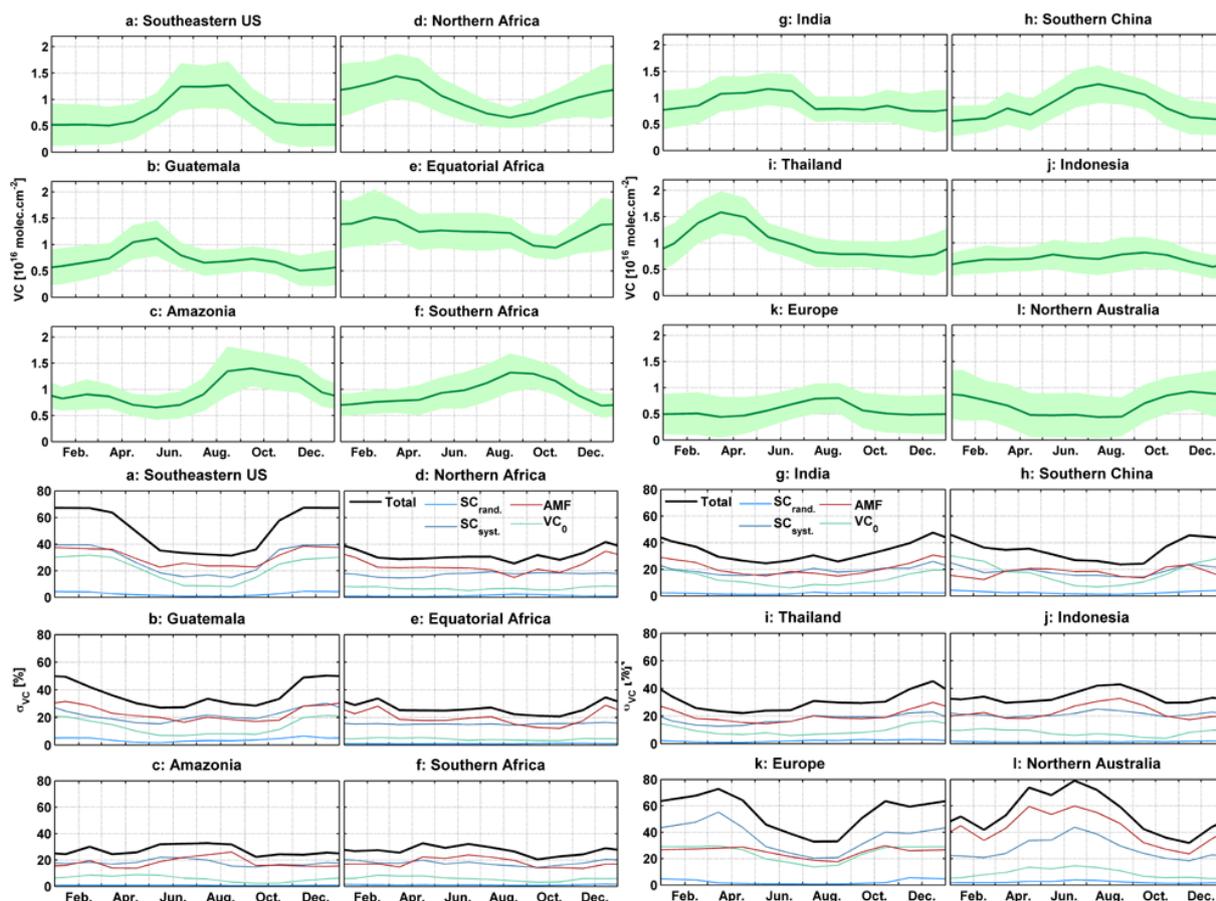
Fig. 10. GOME-2 H_2CO vertical columns averaged regionally between 2007 and 2011 (upper panels), and their estimated relative errors (lower panels) with the contributions of each error source. The limits of the regions are shown in Fig. 9.

scattering layers and if detailed information on aerosol optical properties is available at the global scale (Valks et al., 2011). Besides the cloud and aerosol impacts, it is recognised that the errors related to uncertainties in the a priori profile shapes can be locally as large as 40 % (Barkley et al., 2012).

As observed with previous satellite measurements, the highest annual columns are found over tropical regions in Africa and Amazon, where biogenic and biomass burning sources dominate the emissions (Barkley et al., 2008; Müller et al., 2008; Stavrakou et al., 2009a), and in Southeastern Asia, where biogenic, biomass burning and anthropogenic sources all contribute to the observed H_2CO column (Fu et al., 2007; Stavrakou et al., 2009a). The formaldehyde seasonal variations are primarily related to the increase of biogenic emissions during local summer months at mid-latitudes over deciduous forests (in America, in Europe, in Northern Asia and in Australia), and during the dry season over evergreen tropical forests (in Amazon and Africa). In the Tropics, biomass burning also contributes significantly to the H_2CO columns (Gonzi et al., 2011; Stavrakou et al., 2009b). Over the Amazon, burning is more widespread between September and November (SON) (Barkley et al., 2011), while over

Africa, a dipole pattern exists owing to different seasonal burning either side of the Equator (Stavrakou et al., 2009a; Marais et al., 2012). Localised emissions along the Nile Valley appear in the JJA map, such as along the plains in Iraq up to the Persian Gulf. Over Asia, biogenic emissions dominate the H_2CO signal, coupled with seasonal agricultural burning and forest fires (Fu et al., 2007). In Southeastern Asia, important biomass burning events take place in India and in Indochina from March to May. In highly populated regions of Asia, anthropogenic emissions of NMVOCs have been reported over major cities and over industrial Eastern China, the Indo-Gangetic plains and central India where biofuel is extensively used (Fu et al., 2007). However, even there, the biogenic contribution to the total H_2CO column largely dominates. For this reason, the anthropogenic signal should be easier to isolate during wintertime. This is not the most favourable situation since the quality of the H_2CO satellite retrieval decreases in the case of low Sun observations. In the GOME-2 map for the SON period, anthropogenic H_2CO signals are visible in urban areas of central and southern China including Beijing, Chongqing, Wuhan and Changsha, Xi'an, Hong Kong and the Pearl River Delta,
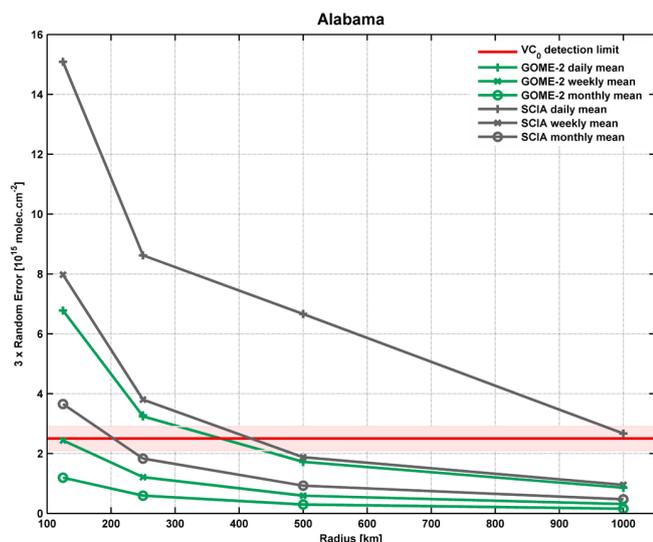
Fig. 11. Detection limit of the SCIAMACHY and GOME-2 daily, weekly and monthly averaged H_2CO vertical columns, in function of the radius considered around Alabama (southeastern US) in August 2007. The red line shows the IMAGES correction in the remote Pacific, considered as the background detection limit.

and in major Indian cities, including New Delhi and Kolkata. In North America, biogenic NMVOCs dominate the H_2CO variability. The spatial distribution of the biogenic emissions seen by GOME-2 is better defined than previously observed with SCIAMACHY or GOME. For example, the summertime formaldehyde concentrations follows closely the spatial distribution of the eastern deciduous forests along the spine of the Appalachian mountains, but also along a narrower band along the western coast from California to Mexico (Millet et al., 2008; Boeke et al., 2011). Similar patterns are observed with the OMI instrument, which currently offers the best spatial resolution and temporal coverage (Millet et al., 2008). During fall, weaker H_2CO signals are detectable in the vicinity of some urban areas, like for example around Los Angeles and Phoenix, or between Houston and Dallas in eastern Texas, where local petrochemical industry is important (Millet et al., 2008). Maximum summertime concentrations in Europe present lower levels than in America and are observed over broadleaf forests in Central and Eastern Europe (over Germany, Poland, Hungary, Ukraine and Western Russia), and more surprisingly also over closed seas (Mediterranean, Aegean, Black and Caspian Seas), as already observed with OMI measurements (Curci et al., 2010; Sabolis et al., 2011) and with aircraft field campaigns (Klippel et al., 2011).

As already mentioned, the identification of regional emissions patterns in the GOME-2 observations are much better defined than with the GOME and SCIAMACHY observations, owing to the better sampling rate of GOME-2 and despite the fact that its ground pixels are larger than those

of SCIAMACHY. The precision of the current GOME and SCIAMACHY H_2CO datasets does not allow for a better time/space resolution than monthly averaged columns over rather large regions in order to sufficiently reduce the noise in the observations (see e.g. Dufour et al., 2009). The better sampling rate of GOME-2 allows for a significant reduction of the noise in the averaged columns. To illustrate this, we compared the GOME-2 and SCIAMACHY detection limits, defined as three times the random error (first term of Eq. 3). In Fig. 11, the detection limits are shown for daily, weekly and monthly averaged H_2CO columns in August 2007 in southeastern US, as a function of the circle radius considered around the state of Alabama (centre: 35° N, 87° W). The value of the reference sector correction (N_{v0}^{CTM}, shown in red in the figure) represents the background level of H_2CO at this latitude and is considered here as a threshold for the desired minimum detection limit of the observations. It can be seen that the SCIAMACHY observations need to be averaged monthly within a radius of at least 200 km (or weekly in a radii of 500 km) to provide sufficiently low random errors. Using GOME-2 observations, it is possible to work with weekly averaged values within 150 km, or even with daily averaged values within radius larger than 500 km. This opens new perspectives for the exploitation of satellite H_2CO observations, together with the OMI early afternoon H_2CO observations.

6 Comparison of the GOME, SCIAMACHY and GOME-2 H_2CO datasets

Figure 12 compares the time series of the GOME, SCIAMACHY and GOME-2 monthly averaged columns in the target regions whose boundaries are displayed in Fig. 9. The value of the reference sector correction is also shown in the plots as an indication of the H_2CO detection limit. This allows discriminating between regions (or time periods) where the total column is constrained by the satellite observations themselves rather than by the model background. In the Tropics, the observed H_2CO columns are well above the detection limit throughout the year. In mid-latitude regions, the summer H_2CO observations are also elevated, but in winter they are generally close to, or just above, the detection limit. For each region, the correlation of GOME-2 with co-incident SCIAMACHY observations is given within the plot, as well as the slope of the regression line of GOME-2 versus SCIAMACHY. The correlation coefficients are higher than 0.8 in almost all regions. They are slightly lower (0.7) in Europe, Indonesia and northern Australia. The difference between the SCIAMACHY and GOME-2 H_2CO columns is found to be lower than 5 % in the regions of Guatemala, Amazonia, South Africa, India, southern China, Thailand and Indonesia. Note that in India and China, the positive trends previously detected in the GOME and SCIAMACHY H_2CO columns (De Smedt et al., 2010) are supported by
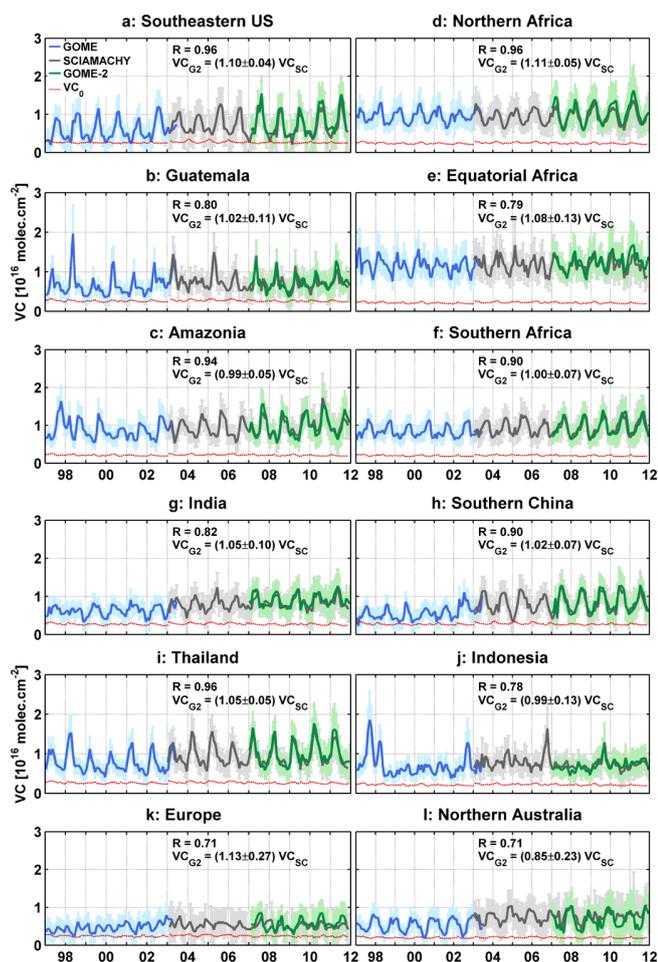
Fig. 12. Time series of monthly and regionally averaged H_2CO vertical columns retrieved from GOME (in blue), SCIAMACHY (in grey) and GOME-2 (in green). The value of the reference sector correction is shown in red as an indication of the detection limit of the observations. The limits of the regions are shown in Fig. 9. The correlation values and regression lines between SCIAMACHY and GOME-2 observations are given in the inset of each plot for the period 2007–2011.

the GOME-2 observations. The GOME-2 columns are about 10 % larger than SCIAMACHY in northern and equatorial Africa, such as in southeastern US and Europe. Larger differences are found in Australia, where the GOME-2 columns are 15 % lower than SCIAMACHY, but appear more consistent with the GOME time series. The discrepancy between SCIAMACHY and GOME-2 in this region is mostly important during the winters of 2009 and 2010, when the amplitude of the seasonal variations in the SCIAMACHY data is significantly reduced. An instrumental degradation effect in SCIAMACHY time series cannot be excluded. As already mentioned, instrumental degradation has noticeable effects on the H_2CO random error (see Fig. 3), but no significant impact on the H_2CO columns was detected so far (except over Australia for SCIAMACHY). It must be noted that the

daily normalisation procedure applied to eliminate systematic zonal artefacts also helps in maintaining the long-term stability of the H_2CO vertical columns by minimising the sensitivity of the retrieved slant columns to long-term instrumental degradation effects. The general level of agreement obtained between the three sensors in all regions is very satisfactory with regard to the estimated uncertainties in the total columns, which range between 30 % and 40 % for these monthly and spatially averaged data (see Fig. 10).

7 Conclusions and perspectives

Global distributions of formaldehyde tropospheric columns have been retrieved from earthshine backscatter radiance spectra recorded by GOME-2 on METOP-A between 2007 and 2011. Improved DOAS retrieval settings have been developed for this instrument, with the ultimate goal to provide a consistent global long-term multi-sensor dataset of H_2CO mid-morning columns based on GOME, SCIAMACHY and GOME-2 measurements. The updated GOME-2 settings include a two-step fitting procedure to minimise interferences between H_2CO and BrO spectral structures, and a modified DOAS approach to better cope with the strong O_3 absorption effects at moderate and large solar zenith angles. To handle the impact of time-dependent GOME-2 slit function changes, an asymmetric Gaussian line-shape is derived as part of a calibration procedure applied on daily solar spectra. These corrections reduce the noise in the H_2CO slant columns, soften the instrumental degradation effects and improve the detection of H_2CO at mid-latitudes. Consistent air mass factors are applied to the three instruments based on up-to-date surface albedo and cloud parameters complemented by a new database of a priori H_2CO profile shapes of improved spatiotemporal representativeness. The resulting SCIAMACHY and GOME-2 formaldehyde columns, monthly averaged over the main NMVOC emission region of between 2007 and 2011, show excellent agreement, with correlation coefficients higher than 0.8 and mean column differences lower than 10 % almost everywhere.

The whole dataset is available on the formaldehyde pages of the TEMIS website (www.temis.nl) and is delivered together with a detailed error estimate. For individual measurements, the random error in the slant column is the largest source of uncertainty, while systematic errors dominate on monthly averages. These are mostly related to imperfect cloud and aerosols corrections, and by uncertainties in the a priori H_2CO vertical profile shape. As regards the latter error source, vertical columns might be improved for particular locations by using more accurate a priori profiles, for example based on input from regional models, ground-based or aircraft measurements. To this aim, the averaging kernels and the a priori profiles are provided for each individual vertical column.

More generally, the validation of this and similar datasets of short-lived tropospheric trace gas measurements currently remains a challenge due to the scarcity of suitable reference ground datasets and more generally to the difficulty of handling the impact of the different spatial and temporal sampling of satellite and ground-based measurements. Therefore, more measurements and validation studies making use of in-situ and FTIR or MAXDOAS ground-based remote sensing systems are definitely needed over a variety of regions, in particular in the Tropics and at the suburban level in mid-latitudes.

Finally, the harmonised and fully documented long-term global formaldehyde dataset presented in this work is ideally suited to support global air quality and chemistry-climate related studies (e.g. Sartelet et al., 2012). In order to fully exploit multi-platform H_2CO observations, and to use them synergistically with other trace gases observations in multi-compounds inversion schemes, it is essential to homogenise as far as possible the retrieval settings as well as the external databases, and to properly characterise the information contents, in order to reduce as much as possible systematic bias between the observations.

References

Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochem. Cy., 15, 955–966, 2001.

Barkley, M. P., Palmer, P. I., Kuhn, U., Kesselmeier, J., Chance, K. V., Kurosu, T. P., Martin, R. V., Helmig, D., and Guenther, A.: Net ecosystem fluxes of isoprene over tropical South America inferred from Global Ozone Monitoring Experiment (GOME) observations of HCHO columns, J. Geophys. Res., 113, D20304, doi:10.1029/2008JD009863, 2008.

Barkley, M. P., Palmer, P. I., Ganzeveld, L., Arneth, A., Hagberg, D., Karl, T., Guenther, A., Paulot, F., Wennberg, P. O., Mao,

J., Kurosu, T. P., Chance, K., Müller, J.-F., De Smedt, I., Van Roozendael, M., Chen, D., Wang, Y., and Yantosca, R.: Can a "state of the art" chemistry transport model simulate Amazonian tropospheric chemistry?, J. Geophys. Res., 116, D16302, doi:10.1029/2011JD015893, 2011.

Barkley, M. P., Kurosu, T. P., Chance, K, De Smedt, I., Van Roozendael, M., Arneth, A., Hagberg, D., and Guenther, A.: Assessing sources of uncertainty in formaldehyde air mass factors over tropical South America: Implications for top-down isoprene emission estimates, J. Geophys. Res., 117, D13304, doi:10.1029/2011JD016827, 2012.

Boeke, N. L., Marshall, J. D., Alvarez, S., Chance, K. V., Fried, A., Kurosu, T. P., Rappengluck, B., Richter, D., Walega, J., Weibring, P., and Millet, D. B.: Formaldehyde columns from the Ozone Monitoring Instrument: Urban versus background levels and evaluation using aircraft data and a global model, J. Geophys. Res., 116, 5303DO, doi:10.1029/2010JD014870, 2011.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO 2 retrieval from space, J. Geophys. Res., 109, D04311m doi:10.1029/2003JD003962, 2004.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905–1928, doi:10.5194/amt-4-1905-2011, 2011.

Brion, J., Chakir, A., Charbonnier, J., Daumont, D., Parisse, C., and Malicet, J.: Absorption spectra measurements for the ozone molecule in the 350–830 nm region, J. Atmos. Chem., 30, 291–299, 1998.

Cai, Z., Liu, Y., Liu, X., Chance, K., Nowlan, C. R., Lang, R., Munro, R., and Suleiman, R.: Characterization and correction of Global Ozone Monitoring Experiment 2 ultraviolet measurements and application to ozone profile retrievals, J. Geophys. Res., 117, D07305, doi:10.1029/2011JD017096, 2012.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2- Metop's second-generation sensor for operational ozone monitoring, ESA Bull., 102, 28–36, 2000.

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, 1289–1295, 2010.

Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., and Jacob D. J.: Satellite observations of formaldehyde over North America from GOME, Geophys. Res. Lett., 27, 3461–3464, 2000.

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, 11501–11517, doi:10.5194/acp-10-11501-2010, 2010.

De Smedt, I.: Long-Term Global Observations of Tropospheric Formaldehyde Retrieved from Spaceborne Nadir UV Sensors, Ph.D. thesis, Faculty of Applied Sciences, University of Brussels, Belgium, 2011.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 4947–4963,

doi:10.5194/acp-8-4947-2008, 2008.

De Smedt, I., Stavrakou, T., Müller, J. F., van Der A, R. J., and Van Roozendael, M.: Trend detection in satellite observations of formaldehyde tropospheric columns, Geophys. Res. Lett., 37, L18808, doi:10.1029/2010GL044245, 2010.

Dikty, S. and Richter, A.: GOME-2 on MetOp-A Support for Analysis of GOME-2 In-Orbit Degradation and Impacts on Level 2 Data Products, Final Report, Version 1.2, 14 October 2011.

Dufour, G., Wittrock, F., Camredon, M., Beekmann, M., Richter, A., Aumont, B., and Burrows, J. P.: SCIAMACHY formaldehyde observations: constraint for isoprene emission estimates over Europe?, Atmos. Chem. Phys., 9, 1647–1664, doi:10.5194/acp-9-1647-2009, 2009.

EUMETSAT, GOME-2 Products Guide, Ref.: EUM/OPS-EPS/MAN/07/0445, Issue: v2D, Date: 6 March 2009.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. Photochem. Photobiol. A, 168, 117–132, 2004.

Fu, T.-M., Jacob, D. J., Palmer, P. I., Chance, K. V., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C., and Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone, J. Geophys. Res., 112, D06312, doi:10.1029/2006JD007853, 2007.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, Atmos. Chem. Phys., 12, 6699–6721, doi:10.5194/acp-12-6699-2012, 2012.

Gonzi, S., Palmer, P. I., Barkley, M. P., De Smedt, I., and Van Roozendael, M.: Biomass burning emission estimates inferred from satellite column measurements of HCHO: Sensitivity to co-emitted aerosol and injection height, Geophys. Res. Lett., 38, L14807, doi:10.1029/2011GL047890, 2011.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F.m and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res., 113, D18308, doi:10.1029/2008JD010290, 2008.

Klippel, T., Fischer, H., Bozem, H., Lawrence, M. G., Butler, T., Jöckel, P., Tost, H., Martinez, M., Harder, H., Regelin, E., Sander, R., Schiller, C. L., Stickler, A., and Lelieveld, J.: Distribution of hydrogen peroxide and formaldehyde over Central Europe during the HOOVER project, Atmos. Chem. Phys., 11, 4391–4410, doi:10.5194/acp-11-4391-2011, 2011.

Koelemeijer, R. B. A., Stammes, P., Hovenier, J. W., and de Haan, J. F.: A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106, 3475–3490, doi:10.1029/2000JD900657, 2001.

Koelemeijer, R. B. A., de Haan, J. F., and Stammes, P.: A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations, J. Geophys. Res., 108, 4070, doi:10.1029/2002JD002429, 2003.

Kurosu, T. P.: OMHCHO README FILE, available at: http://www.cfa.harvard.edu/atmosphere/Instruments/OMI/PGEReleases/READMEs/OMHCHO_README.pdf (last access: 14 August 2012), 2008.

Lacan, A. and Lang, R.: Investigation on GOME-2 throughput degradation, Final report, EUM/LEO/REP/09/0732 Issue 1.1, 16 July 2011.

Lang, R., Munro, R., Livschitz, Y., Dyer, R., and Lacan, A.: GOME-2 FM3 Long-Term In-Orbit Degradation – Basic Signatures After 2nd Throughput Test, EUMETSAT Technical report, EUM.OPS-EPS.DOC.09.0464, 2009.

Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q. J., Beekmann, M., and Burrows, J. P.: On the improvement of NO$_2$ satellite retrievals – aerosol impact on the airmass factors, Atmos. Meas. Tech., 3, 475–493, doi:10.5194/amt-3-475-2010, 2010.

Lerot, C., Stavrakou, T., De Smedt, I., Müller, J.-F., and Van Roozendael, M.: Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model, Atmos. Chem. Phys., 10, 12059–12072, doi:10.5194/acp-10-12059-2010, 2010.

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns, Atmos. Chem. Phys., 12, 6219–6235, doi:10.5194/acp-12-6219-2012, 2012.

Martin, R. V., Chance, K. V., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M., and Koelemeijer, R. B. A.: An improved retrieval of tropospheric nitrogen dioxide from GOME, J. Geophys. Res., 107, 4437, doi:10.1029/2001JD001027, 2002.

Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross section of HCHO between 223 and 323K in the wavelength range 225–375 nm, J. Geophys. Res., 105, 7089–7102, doi:10.1029/1999JD901074, 2000.

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K. V., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, J. Geophys. Res., 113, 1–18, doi:10.1029/2007JD008950, 2008.

Müller, J.-F., Stavrakou, T., Wallens, S., De Smedt, I., Van Roozendael, M., Potosnak, M. J., Rinne, J., Munger, B., Goldstein, A., and Guenther, A. B.: Global isoprene emissions estimated using MEGAN, ECMWF analyses and a detailed canopy environment model, Atmos. Chem. Phys., 8, 1329–1341, doi:10.5194/acp-8-1329-2008, 2008.

Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A. P.: GOME-2 on MetOp, Proc. of The 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, 2006.

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, Atmos. Chem. Phys., 7, 4419–4444, doi:10.5194/acp-7-4419-2007, 2007.

Palmer, P. I., Jacob, D. J., Chance, K. V., Martin, R. V., D, R. J., Kurosu, T. P., Bey, I., Yantosca, R., and Fiore, A.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106, 14539–14550, doi:10.1029/2000JD900772, 2001.

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K. V., Kurosu, T. P., Guenther, A., Wiedinmyer, C., Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B., and Sumner, A. L.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, J. Geophys. Res., 111, 1–14, doi:10.1029/2005JD006689, 2006.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy: Principles and Applications (Physics of Earth and Space Environments), Springer-Verlag, Berlin, Heidelberg, ISBN 978-3540211938, 2008.

Popp, C., Wang, P., Brunner, D., Stammes, P., Zhou, Y., and Grzegorski, M.: MERIS albedo climatology for FRESCO+ O2 A-band cloud retrieval, Atmos. Meas. Tech., 4, 463–483, doi:10.5194/amt-4-463-2011, 2011.

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

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

Rozanov, A., Rozanov, V., and Burrows, J. P.: A numerical radiative transfer model for a spherical planetary atmosphere: Combined differential integral approach involving the Piccard iterative approximation, J. Quant. Spectrosc. Radiat. Transfer, 69, 491–512, 2001.

Sabolis, A., Meskhidze, N., Curci, G., Palmer, P. I., and Gantt, B.: Interpreting elevated space-borne HCHO columns over the Mediterranean Sea using the OMI sensor, Atmos. Chem. Phys., 11, 12787–12798, doi:10.5194/acp-11-12787-2011, 2011.

Sartelet, K. N., Couvidat, F., Seigneur, C., and Roustan, Y.: Impact of biogenic emissions on air quality over Europe and North America, Atm. Environ., 53, 131–141, 2012.

Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, Atmos. Chem. Phys., 3, 161–180, doi:10.5194/acp-3-161-2003, 2003.

Schultz, M. G., Backman, L., Balkanski, Y., Bjoerndalsaeter, S., Brand, R., Burrows, J. P., Dalsoeren, S., de Vasconcelos, M., Grodtmann, B., Hauglustaine, D. A., Heil, A., Hoelzemann, J. J., Isaksen, I. S. A., Kaurola, J., Knorr, W., Ladstaetter-Weißenmayer, A., Mota, B., Oom, D., Pacyna, J., Panasiuk, D., Pereira, J. M. C., Pulles, T., Pyle, J., Rast, S., Richter, A., Savage, N., Schnadt, C., Schulz, M., Spessa, A., Staehelin, J., Sundet, J. K., Szopa, S., Thonicke, K., van het Bolscher, M., van Noije, T., van Velthoven, P., Vik, A. F., and Wittrock, F.: REanalysis of the TROpospheric chemical composition over the past 40 years (RETRO): A long-term global modeling study of tropospheric Julich/Hamburg, Germany, 48/2007 report on Earth chemistry, Julich/Hamburg, System Science of the Max Planck Institute for Meteorology, Hamburg, available at: http://retro.enes.org, ISSN 1614-1199, 2007.

Siddans, R., Latter, B. G., Kerridge, B. J., Smeets, J., Otter, G., and Slijkhuis, S.: Analysis of GOME-2 Slit function Measurements: Final Report Eumetsat Contract No. EUM/CO/04/1298/RM, 2006.

Spurr, R. J. D.: Linearized radiative transfer theory, A general discrete ordinate approach to the calculation of radiances and analytic weighting functions, with applications to atmospheric remote sensing, Thesis Manuscript, Technische Universiteit Eindhoven, 2001.

Spurr, R. J. D.: LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems, in Light Scattering Reviews, edited by: Kokhanovsky, A., Berlin, 229–271, 2008.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, Atmos. Chem. Phys., 9, 1037–1060, doi:10.5194/acp-9-1037-2009, 2009a.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006, Atmos. Chem. Phys., 9, 3663–3679, doi:10.5194/acp-9-3663-2009, 2009b.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, Atmos. Chem. Phys., 9, 8431–8446, doi:10.5194/acp-9-8431-2009, 2009c.

Stavrakou, T., Peeters, J., and Müller, J.-F.: Improved global modelling of HO$_x$ recycling in isoprene oxidation: evaluation against the GABRIEL and INTEX-A aircraft campaign measurements, Atmos. Chem. Phys., 10, 9863–9878, doi:10.5194/acp-10-9863-2010, 2010.

Stavrakou, T., Guenther, A., Razavi, A., Clarisse, L., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Karagulian, F., De Mazière, M., Vigouroux, C., Amelynck, C., Schoon, N., Laffineur, Q., Heinesch, B., Aubinet, M., Rinsland, C., and Müller, J.-F.: First space-based derivation of the global atmospheric methanol emission fluxes, Atmos. Chem. Phys., 11, 4873–4898, doi:10.5194/acp-11-4873-2011, 2011.

Taraborrelli, D., Lawrence, M. G., Butler, T. M., Sander, R., and Lelieveld, J.: Mainz Isoprene Mechanism 2 (MIM2): an isoprene oxidation mechanism for regional and global atmospheric modelling, Atmos. Chem. Phys., 9, 2751–2777, doi:10.5194/acp-9-2751-2009, 2009.

Theys, N., Van Roozendael, M., Hendrick, F., Yang, X., De Smedt, I., Richter, A., Begoin, M., Errera, Q., Johnston, P. V., Kreher, K., and De Mazière, M.: Global observations of tropospheric BrO columns using GOME-2 satellite data, Atmos. Chem. Phys., 11, 1791–1811, doi:10.5194/acp-11-1791-2011, 2011.

Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO$_2$ column retrieval for GOME-2, Atmos. Meas. Tech., 4, 1491–1514, doi:10.5194/amt-4-1491-2011, 2011.

Vandaele, A. C., Hermans, C., Fally, S., Carleer, M., Colin, R., Mérienne, M.-F., Jenouvrier, A., and Coquart, B.: High-resolution Fourier transform measurement of the NO$_2$ visible and near-infrared absorption cross sections: Temperature and pressure effects, J. Geophys. Res., 107, 4348, doi:10.1029/2001JD000971, 2002.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.

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

Vountas, M., Rozanov, V. V., and Burrows, J. P.: Ring effect: impact of rotational Raman scattering on radiative transfer in earth's atmosphere, J. of Quant. Spec. Rad. Trans., 60, 943–961, 1998.

Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem. Phys., 10, 10145–10160, doi:10.5194/acp-10-10145-2010, 2010.

Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.: FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8, 6565–6576, doi:10.5194/acp-8-6565-2008, 2008.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, Geophys. Res. Lett., 33, 1–5, doi:10.1029/2006GL026310, 2006.

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