Atmos. Meas. Tech. Discuss., 7, 6065–6112, 2014
www.atmos-meas-tech-discuss.net/7/6065/2014/
doi:10.5194/amtd-7-6065-2014
© Author(s) 2014. CC Attribution 3.0 License.

Atmospheric Measurement Techniques Discussions
Open Access

This discussion paper is/has been under review for the journal Atmospheric Measurement Techniques (AMT). Please refer to the corresponding final paper in AMT if available.

Glyoxal retrieval from the Ozone Monitoring Instrument

C. C. Miller[1], G. G. Abad[2], H. Wang[2], X. Liu[2], T. Kurosu[2,*], D. J. Jacob[1,3], and K. Chance[2]

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA
[2]Harvard Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
[3]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA
[*]now at: Atmospheric Observations, Jet Propulsion Laboratory, Pasadena, California, USA

Received: 1 May 2014 – Accepted: 3 June 2014 – Published: 18 June 2014

Correspondence to: C. C. Miller (cmiller@fas.harvard.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We present an algorithm for the retrieval of glyoxal from backscattered solar radiation, and apply it to spectra measured by the Ozone Monitoring Instrument (OMI). The algorithm is based on direct spectrum fitting, and adopts a two-step fitting routine to account for liquid water absorption. Previous studies have shown that glyoxal retrieval algorithms are highly sensitive to the position of the spectral fit window. This dependence was systematically tested on real and simulated OMI spectra. We find that a combination of errors resulting from uncertainties in reference cross sections and spectral features associated with the Ring effect are consistent with the fit-window dependence observed in real spectra. This implies an optimal fitting window of 435–461 nm, consistent with previous satellite glyoxal retrievals. The results from the retrieval of simulated spectra also support previous findings that have suggested that glyoxal is sensitive to NO_2 cross section temperature. The retrieval window limits of the liquid water retrieval are also tested. A retrieval window 385–470 nm reduces interference with strong spectral features associated with sand. We show that cross track dependent offsets (stripes) present in OMI can be corrected using offsets derived from retrieved slant columns over the Sahara, and apply the correction to OMI data. Average glyoxal columns are on average lower than those of previous studies likely owing to the choice of reference sector for offset correction. OMI VCDs are lower compared to other satellites over the tropics and Asia during the monsoon season, suggesting that the new retrieval is less sensitive to water vapor abundance. Consequently we do not see significant glyoxal enhancements over tropical oceans. OMI derived glyoxal-to-formaldehyde ratios over biogenic and anthropogenic source regions are consistent with surface observations.

1 Introduction

The oxidation of non-methane volatile organic compounds (NMVOC)s is an important atmospheric process for both air quality and climate (Lippmann, 1989; Kanakidou et al.,

2005; Lelieveld and Dentener, 2000; Forster et al., 2007). NMVOC emissions inventories are subject to large uncertainties associated with extrapolating limited data on activity rates and emission factors. Emission estimates for biogenic isoprene, the largest NMVOC source on a global scale (Guenther et al., 2006) are uncertain by almost a factor of two (Pfister et al., 2008). Anthropogenic NMVOC emissions are also highly uncertain (Borbon et al., 2013; Parrish et al., 2012), especially in developing regions that rely on potentially unrepresentative foreign emissions data (Klimont et al., 2002; Streets et al., 2003; Wei et al., 2008).

Glyoxal (CHO-CHO) is a short lived product of NMVOC atmospheric oxidation observable from space (Wittrock et al., 2006; Vrekoussis et al., 2010; Lerot et al., 2010; Chance, 2006). Satellite observations of formaldehyde (HCHO), another product of NMVOC atmospheric oxidation have proven to be a valuable constraint on NMVOC emissions estimates (Abbot et al., 2003; Palmer et al., 2006; Millet et al., 2008; Marais et al., 2012; Fu et al., 2007; Curci et al., 2010). Glyoxal-HCHO ratios could provide information on NMVOC speciation (Vrekoussis et al., 2010; DiGangi et al., 2012). Satellite observations of glyoxal are made in a longer wavelength range (~ 435–460 nm) than HCHO (~ 330–360 nm) and are therefore less sensitive to molecular scattering that diminishes sensitivity to the lower troposphere (Palmer et al., 2001). Stavrakou et al. (2009) used both glyoxal and HCHO retrievals from SCIAMACHY to constrain global NMVOC emissions, and estimated that current models underestimate glyoxal by a factor of approximately two. However glyoxal-HCHO ratios differ between ground and satellite observations (DiGangi et al., 2012), a discrepancy that needs to be resolved to interpret glyoxal observations from space.

Glyoxal optical depths are very weak (O(10^{-3})). This makes the retrieval highly sensitive to fitting errors from stronger absorbers, as well as instrument calibration errors and the spectral structure of surface reflectance. Recent studies have shown that glyoxal concentrations retrieved from satellite (Alvarado et al., 2013) and from the surface (Sinreich et al., 2013) are highly sensitive to the settings of the retrieval, in particular, the choice of fit window.

Here we present a glyoxal retrieval for the Ozone Monitoring Instrument (OMI) with optimized retrieval settings. OMI offers superior spatial resolution and temporal coverage compared to existing satellite instruments (GOME-2 and SCIAMACHY) used to retrieve glyoxal. However retrieving glyoxal from OMI spectra is more challenging compared to the aforementioned instruments due to its lower spectral resolution and smaller signal-to-noise ratio. We use simulated OMI spectra to test retrieval accuracy and apply a systematic approach (Vogel et al., 2013) to optimize the glyoxal fit window. OMI uses a 2-D CCD array in contrast to the linear photodiode array detectors used in GOME-2 and SCIAMACHY. This makes retrieved slant columns subject to cross track biases. We present a simple method to correct for these glyoxal offsets.

2 Methods

OMI was launched on the NASA Aura satellite in sun-synchronous orbit in July 2004, with an equatorial crossing time of 13:42 LT. It is a CCD spectrometer measuring backscattered solar radiation with a 13 km × 24 km nadir resolution and daily global coverage. It's spectral range is 270–500 nm divided over three channels, allowing the retrieval of both HCHO and glyoxal. Glyoxal is retrieved in the visible channel (full spectral range 350–500 nm). The visible channel CCD is divided into 60 across track positions, with an average spectral sampling distance of 0.21 nm and average spectral resolution of 0.63 nm (full width at half maximum).

Glyoxal vertical column densities (VCDs, molecules cm^{-2}) are determined using a two-step approach widely employed for optically thin trace gas retrievals in the UV Visible spectral region. In the first step modelled spectra are directly fitted to observed OMI radiances to determine slant column densities (SCDs) that represent the integrated glyoxal number density through the mean photon path from the Sun to the instrument. In the second step, the SCDs are translated to VCDs using air mass factors (AMFs) computed using a radiative transfer model (Palmer et al., 2001).

2.1 Fitting glyoxal slant columns

Glyoxal SCDs are determined using the direct spectrum fitting approach described by Chance (1998). Here the state vector $\boldsymbol{x} \in \mathbb{R}^n$, consisting of a set of variables impacting the observed radiance (including the glyoxal SCD) is estimated from a set of observed radiance values at a number of discrete wavelengths (λ). Let $\boldsymbol{y} \in \mathbb{R}^m$ be the vector of these discrete radiance values. Assuming that the noise variance in the measured spectrum is the same for all wavelengths, the optimal estimate for the state $\hat{\boldsymbol{x}}$ is found by the least squares difference between the observed radiance and a model spectrum $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{b})$, a function of the state vector and a set of unoptimized parameters $\boldsymbol{b}$

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\arg\min} \sum_{i=1}^{m} (y_i - F_i(\boldsymbol{x}, \boldsymbol{b}))^2 \tag{1}$$

The modeled spectrum ($F(\lambda)$) consists of a solar source term $I_0(\lambda)$ that is then modified by trace gas absorption $\tau(\lambda)$, a common mode spectrum $R(\lambda)$ constructed by averaging a set of spectrum fit residuals, and scaling and baseline polynomials ($P_{\mathrm{sc}}(\lambda)$ and $P_{\mathrm{bl}}(\lambda)$ respectively), intended to account for broad-band spectral features;

$$F(\lambda) = \left[I_0(\lambda) \exp(-\tau(\lambda)) + R(\lambda)\right] P_{\mathrm{sc}}(\lambda) + P_{\mathrm{bl}}(\lambda) \tag{2}$$

The source spectrum ($I_0(\lambda)$) is derived from the monthly running mean of a set of daily solar irradiance spectra measured by OMI during the end of its orbit ($b_{\mathrm{sol}}(\lambda)$). Due to the satellite's orbital motion relative to the sun, the solar irradiance spectra are Doppler shifted relative to the earth observations. To account for this, both the solar and earthshine grids are calibrated using a high resolution solar reference spectrum (Chance and Kurucz, 2010). Since the measured spectra are not fully Nyquist sampled, direct interpolation of the measured solar spectrum to the earthshine grid introduces aliasing. To account for this, an additional term $b_u(\lambda)$ derived from the difference in a fully sampled and under sampled solar reference spectrum is included in the fit

(Chance et al., 2005). Finally an inelastic Raman scattering source term ($b_r(\lambda)$) to account for "filling in" of the solar lines from O_2 and N_2 rotational transitions is included as described in Chance and Spurr (1997).

$$I_0(\lambda) = b_{\mathrm{sol}}(\lambda) + x_u b_u(\lambda) + x_r b_r(\lambda) \tag{3}$$

The source spectrum is attenuated by trace gas absorption. The total optical depth ($\tau(\lambda)$) is the sum of the contributions from each absorber.

$$\tau(\lambda) = \sum_j x_j b_j(\lambda) \tag{4}$$

Here, x_j and $b_j(\lambda)$ are the SCD and reference cross section (RCS) of species j respectively.

Table 1 summarises the RCSs included in the fitting procedure. The two strongest glyoxal absorption bands lie in the 430–460 nm spectral region, as shown in Fig. 1. In addition to glyoxal, absorption due to ozone (O_3), nitrogen dioxide (NO_2), water vapor (H_2O) and the oxygen collision complex (O_2–O_2) contribute significantly to the total optical depth, and are therefore included in the fitting process. Previous work has shown that surface extinction from liquid water is significant over clear surface waters, where the mean photon path through the ocean is significant (Vrekoussis et al., 2009). Lerot et al. (2010) found that the cross correlation between the glyoxal and liquid water RCSs within the glyoxal fit region is too high for simultaneous fitting. Our OMI retrieval adopts the Lerot et al. (2010) approach of including pre-fitted optical depths from a separate liquid water retrieval that takes advantage of the broad spectral features of liquid water outside the glyoxal fit window. All RCSs are degraded to the OMI instrument resolution through convolution (denoted $\otimes$) with the measured instrument transfer function $\Gamma(\lambda)$ (Dirksen et al., 2006), and then splined to the instrument wavelength grid. As a measured solar spectrum is used in the fitting process, the convolution to the source

(I_0) and absorption (τ) expressions is done separately. Thus if $I_0^{hr}(\lambda)$ and $\tau^{hr}(\lambda)$ denote the solar spectrum and total optical depth at infinitely high resolution, the first term of Eq. (3) is given by;

$$I_0(\lambda)\exp(-\tau(\lambda)) = I_0^{hr}(\lambda) \otimes \Gamma(\lambda) \exp\left(-\tau^{hr}(\lambda) \otimes \Gamma(\lambda)\right) \quad (5)$$

However in reality, the instrument distorts the spectrum after trace gas absorption. Thus the convolution must be applied last in the true expression.

$$I_0(\lambda)\exp(-\tau(\lambda))_{true} = \left[I_0^{hr}(\lambda)\exp(-\tau^{hr}(\lambda))\right] \otimes \Gamma(\lambda) \quad (6)$$

The difference between equations 5 and 6 is referred to as the solar I_0 effect (Aliwell et al., 2002). For glyoxal, correcting for the I_0 effect is important, as the I_0 effects of strongly absorbing interfering species are comparable in magnitude to observed glyoxal optical depths. For small optical depths the RCS of species j can be corrected using a high resolution solar reference spectrum ($I_{sol}(\lambda)$). Assuming a small reference column density x_j^{ref};

$$b_j(\lambda) = \frac{1}{x_j^{ref}} \ln\left(\frac{I_{sol}(\lambda) \otimes \Gamma(\lambda)}{I_{sol}(\lambda) \otimes \exp\left(-x_j^{ref} b_j^{hr}(\lambda) \otimes \Gamma(\lambda)\right)}\right) \quad (7)$$

The above correction is insensitive to reference column densities for all interfering species considered over ranges typically observed in the atmosphere. Here, reference column densities for each species were chosen so that the optical depth used in Eq. (7) is approximately 10^{-3}. This magnitude is small enough to be in the range where the exponential in Eq. (7) is approximately linear, and thus should be a good approximation for the I_0 effects for all shallow optical depths.

A common mode spectrum $R(\lambda)$ constructed by averaging the fit residuals of spectra between 30° N and 30° S is included in the final spectrum fit. $R(\lambda)$ is intended to account for systematic residuals uncorrelated with the RCSs, for instance those due to errors in the specification of the instrument transfer function. The scaling and baseline polynomials account for broad band spectral effects, including Rayleigh and Mie scattering, wavelength dependent surface reflection and instrument offsets.

$$P_{sc}(\lambda) = \sum_{k=0}^{n_{sc}} x_k^{sc}(\lambda - \bar{\lambda})^k \quad (8)$$

$$P_{bl}(\lambda) = \sum_{k=0}^{n_{bl}} x_k^{bl}(\lambda - \bar{\lambda})^k \quad (9)$$

Here $\bar{\lambda}$ is the mean wavelength over the fitting window. The choice of the appropriate polynomial orders (n_{sc} and n_{bl}) impacts retrieval accuracy. Lower order polynomials may not fully account for the broadband spectral features not physically modelled, whereas a polynomial of too high order may increase error through overfitting. Here we set $n_{sc} = 3$ and $n_{bl} = 1$. This choice was made by performing a set of sensitivity tests systematically varying n_{sc} and n_{bl} over a subset of OMI orbits. Polynomial degrees lower than this order induced latitudinal dependent biases, and larger orders resulted in similar SCDs to the orders selected.

2.2 Determination of glyoxal vertical column densities

The spectrum fitting algorithm described in the previous section returns a slant column measurement of glyoxal ($x_{glyoxal} \equiv \Omega_s$). A more geophysically relevant quantity is the vertical column density (Ω_v), defined as the number density per unit area integrated through the height of the atmosphere. The ratio of these quantities is called the Air Mass Factor (A).

$$A = \frac{\Omega_s}{\Omega_v} \tag{10}$$

For optically thin absorbers including glyoxal, radiative transfer simulations required to determine A can be decoupled from the profile of the trace gas being measured (Palmer et al., 2001).

$$A = \int_0^\infty W(z)S(z)\mathrm{d}z \tag{11}$$

$W(z)$ are called scattering weights, and represent the number of times the radiation reaching the satellite has traversed the layer $[z, d+\mathrm{d}z]$. Here, $W(z)$ for each OMI observation is interpolated from a lookup table calculated with the VLIDORT v2.4RT radiative transfer model, and taking into account instrument viewing geometry, cloud fraction and height, surface height, and reflectance (González Abad et al., 2014). We use data from the OMI O_2–O_2 cloud retrieval algorithm and seasonally dependent OMI lambertian equivalent surface reflectances database from Kleipool et al. (2008) as inputs for the lookup table. Although aerosols are not explicitly accounted for, their impact on the scattering weights is partially accounted for through the cloud retrieval algorithm.

$S(z)$ in Eq. (11) is the vertical shape factor, representing the normalized glyoxal profile ($n(z)$):

$$S(z) = \frac{n(z)}{\Omega_v} \tag{12}$$

Shape factors are computed from a monthly averaged glyoxal climatology for 2007 from the GEOS-Chem chemical transport model (v9-01-03). The model is driven by

GEOS-5 assimilated meteorology with a resolution of 2°×2.5° and 47 vertical levels. We use a modified version of the simulation described in Fu et al. (2008), with significant updates to NMVOC chemistry (Miller et al., 2012).

2.3 Error estimation

Errors in the retrieved glyoxal slant column density are estimated following the methods of Rogers (2000). The difference between the retrieved ($\hat{\boldsymbol{x}}$) and true state ($\boldsymbol{x}$) of the atmosphere arises from a combination of parameter errors ($\boldsymbol{b} - \hat{\boldsymbol{b}}$), noise in the measured spectrum ($\boldsymbol{\epsilon}$) and forward model errors ($\boldsymbol{\Delta f}$).

$$\boldsymbol{x} - \hat{\boldsymbol{x}} = \mathbf{G}_y\mathbf{K}_b(\boldsymbol{b} - \hat{\boldsymbol{b}}) + \mathbf{G}_y\boldsymbol{\epsilon} + \mathbf{G}_y\boldsymbol{\Delta f}(\boldsymbol{x}, \boldsymbol{b}) \tag{13}$$

$\mathbf{K}_b$ is the forward model Jacobian with respect to the model parameters and $\mathbf{G}_y$ is the sensitivity of the retrieved state to changes in the observed radiance. OMI spectra are fitted with the Gauss–Newton based ELSUNC least squares algorithm (Lindström and Wedin, 1988), which additionally uses a truncated QR method far from the solution. Near fit convergence, $\mathbf{G}_y$ is approximated by the sensitivity derived from a Gauss Newton iteration, which is the pseudoinverse of the forward model Jacobian ($\mathbf{K}_x = \partial \boldsymbol{F}/\partial \boldsymbol{x}$)

$$\mathbf{G}_y = \frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{y}} = \left(\mathbf{K}_x^T\mathbf{K}_x\right)^{-1}\mathbf{K}_x^T \tag{14}$$

Parameter errors arise from variables that are not optimized in the fitting process. The main source of parameter error in the glyoxal retrieval is due to uncertainty in the RCSs. Let $\boldsymbol{b}_i \in \mathbb{R}^m$ be the vector of RCS values for species i, assembled on the same wavelength grid as the radiance values $\boldsymbol{y}$. If there are p RCSs included in the retrieval then the full vector of parameters is

$$\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_p \end{pmatrix} \tag{15}$$

The Jacobian for $\boldsymbol{b}_i$ is defined as $\mathbf{K}_b^i = \partial F / \partial \boldsymbol{b}_i$. The full parameter Jacobian matrix corresponding to Eq. (15) is given by;

$$\mathbf{K}_b = \begin{pmatrix} \mathbf{K}_b^1 & \mathbf{K}_b^2 & \dots & \mathbf{K}_b^p \end{pmatrix} \tag{16}$$

As each RCS used in the retrieval derives from independent laboratory measurements, it is reasonable to assume that the RCS error of one species is uncorrelated with the RCSs of the others. Thus the expression for the full error covariance matrix of the parameters has a block diagonal structure.

$$\mathbf{S}_b = \begin{pmatrix} \mathbf{S}_b^1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_b^2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{S}_b^p \end{pmatrix} \tag{17}$$

Here, $\mathbf{S}_b^i$ is the error covariance matrix corresponding to the RCS of the ith species. $\mathbf{S}_b^i$ are approximated as diagonal, and are constructed using relative errors reported in Table 1. The covariance matrix describing how parameter errors impact errors in the fitted variables ($\mathbf{S}_x^b$) can now be found by propagating the parameter errors in Eq. (13) using Eqs. (14), (16) and (17).

$$\begin{aligned} \mathbf{S}_x^b &= \mathbf{G}_y \mathbf{K}_b \mathbf{S}_b \left(\mathbf{G}_y \mathbf{K}_b\right)^{\mathrm{T}} \\ &= \sum_{i=1}^{p} \mathbf{G}_y \mathbf{K}_b^i \mathbf{S}_b^i \left(\mathbf{G}_y \mathbf{K}_b^i\right)^{\mathrm{T}} \end{aligned} \tag{18}$$

The covariance matrix for the measurement error is estimated from the root mean square of the fit residuals (ϵ_{rms}) adjusted by the number of statistical degrees of freedom. No correlation in noise signal between measurement pixels is assumed leading to the following expression for the noise covariance matrix.

$$\mathbf{S}_\epsilon = \epsilon_{\mathrm{rms}}^2 \left(\frac{m}{m-n}\right) \mathbf{I}_{n \times n} \tag{19}$$

Here m are the number of points in the spectrum and n the number of fit variables. This leads to the following estimate of the random error component of the fit variable covariance matrix

$$\begin{aligned} \mathbf{S}_x^\epsilon &= \mathbf{G}_y \mathbf{S}_\epsilon \mathbf{G}_y^{\mathrm{T}} \\ &= \epsilon_{\mathrm{rms}}^2 \left(\frac{m}{m-n}\right) (\mathbf{K}_x^{\mathrm{T}} \mathbf{K}_x)^{-1} \end{aligned} \tag{20}$$

Inference of the forward model error term in Eq. (13) is complicated by an incomplete knowledge of the true atmospheric state, and by how the state maps to the observed spectra. The polynomials included in the fitting process are only approximations for the true physics (scattering, reflectance and instrument effects that cause radiometric offsets). Although no forward model error estimate is included in the retrieval, the sensitivity to forward model error can be assed by testing the retrieval algorithm against model spectra, where the true atmospheric state is known. This will be done in the next section.

3 Retrieval optimisation

The choice of the retrieval spectral window is an important determinant of retrieval accuracy. Figure 2 shows the mean slant optical depths ($\overline{\tau}$) averaged between

435–460 nm for orbit number 10430 on 1 July 2006 (o10430) simulated with VLIDORT using GEOS-Chem species profiles. There is a glyoxal hot spot over central Africa from biogenic and pyrogenic NMVOC emissions with optical depths reaching a maximum of $\overline{\tau} \approx 1.8 \times 10^{-4}$ (slant column of 1.5×10^{15} molecules cm^{-2}). Coincident with the glyoxal hotspot are much higher absorptions by H_2O and NO_2, reaching optical depths as high as $\overline{\tau} \approx 1.2 \times 10^{-3}$ and $\overline{\tau} \approx 6 \times 10^{-3}$ respectively. The simulated slant optical depths of NO_2 and O_3 exhibit a strong dependence on solar/instrument viewing geometry, that can be attributed to strong stratospheric absorption. In the following sections, we will use this case study orbit to evaluate the sensitivity of our retrieval to the settings of the forward model, and the position of the fit window. We start by using simulated spectra to guide the initial design of the retrieval. We additionally test the sensitivity of the retrieval algorithms for the pre-fitted liquid water absorption and glyoxal using real OMI spectra.

3.1 Observing system simulation experiments

The model described in Eq. (2) is only a semi-physical approximation of the true spectrum. We therefore test its performance relative to a model closer to the true physics through an Observing System Simulation Experiment (OSSE). The approach is summarised in Fig. 3. For each OMI track, GEOS-Chem chemical and meteorological profiles were sampled for the instrument viewing geometry and the results averaged onto a 2° latitude grid for computational expediency. The version of GEOS-Chem used here does not simulate stratospheric chemistry, so O_3 and NO_2 zonal climatologies derived from the OMI total column ozone retrieval (Liu et al., 2010) and a stratospheric model (McLinden et al., 2000) were included above the model tropopause. Clear sky synthetic spectra were modeled with VLIDORT on a 0.01 nm grid, using the viewing geometry of OMI. The wavelength dependence of surface reflectance was accounted for by interpolating the OMI derived coarse resolution Kleipool et al. (2008) reflectance database at 400–480 nm using a 3rd order polynomial. Absorption from glyoxal, O_2–O_2, H_2O vapor, O_3 and NO_2 are included in the simulations. We account for the temperature

dependence of O_3 and NO_2 RCSs using the parameterisations of Liu et al. (2007) and Vandaele et al. (2003) respectively. Simulated spectra are convolved with a 0.65 nm FWHM Gaussian approximating the OMI instrument transfer function, and then sampled onto the OMI radiance wavelength grid. The observed solar spectrum is simulated by convolving the high resolution solar reference with the same Gaussian, followed by sampling to the OMI solar irradiance grid. Finally, the retrieval algorithm is applied and the results compared to the "true" state (i.e. GEOS-Chem).

The RCS of O_3 and NO_2 exhibit temperature dependencies that could induce errors in the retrieval if not properly accounted for. Alvarado et al. (2013) found significant improvements in their spectrum fit residuals over heavily polluted regions by incorporating two independent NO_2 RCSs at different temperatures into their glyoxal retrievals. We first tested the impact of four different RCS temperature choices on the retrieval using a 435–460 nm window. Figure 4 shows the difference between the retrieved and true slant column densities. The preliminary version of the OMI glyoxal retrieval used a NO_2 RCS temperature of 220 K (Chance, 2006). Figure 4 shows that using this RCS temperature induces a significant positive global bias in the retrieval, as well as a local bias between 0–15° S over the region with strong pyrogenic emissions. We also tested a 240 K RCS temperature which is closer to the average temperature of the environment of photons absorbed by NO_2. Although this reduces the global bias, the pyrogenic hotspot remains. Including two independent NO_2 RCSs at different temperatures (230 K and 290 K) removes the 0–15° S bias whilst slightly increasing the overall bias, likely due to the added cross correlation caused by fitting the second RCS. This is consistent with the reductions observed by Alvarado et al. (2013). We therefore include two NO_2 RCSs at different temperatures in the operational retrieval to avoid interferences from boundary layer NO_2. Including an additional O_3 RCS (243 K) does not improve the retrieval owing to the small temperature dependence of O_3 between 400–500 nm.

The sensitivity of the retrieval to window position was tested following Vogel et al. (2013), by systematically quantifying OSSE retrieval error as a function of lower and upper wavelength limits. Figure 5 shows the mean bias between the retrieved and true

glyoxal slant columns, as well as the slope of the linear regression of the retrieved vs. "true" slant columns. Retrievals for most window choices have a mean bias of less than 5×10^{13}, except when the window truncates the strongest glyoxal band. The window region centered around 445–463 nm performs optimally, as shown by the lowest mean bias and regression slope closest to 1. This corresponds to the strongest glyoxal absorption band. Extending the window down to 435 nm to include the second strongest glyoxal band slightly increases the mean bias and slope. Given the relatively low retrieval bias for most windows, the results of the OSSE indicate that the spectrum model ($F(\lambda)$) is capable of accounting for the physical effects simulated by the OSSE. These include:

- the I_0 effect
- the RCS temperature dependencies of O_3 and NO_2
- the broad-band corrections for Rayleigh scattering and surface reflectance
- the wavelength dependence of the slant column density of interfering gases
- the undersampling correction of the solar spectrum.

3.2 Stripe correction

Trace gas retrievals from 2-D CCD instruments such as OMI suffer from systematic cross-track biases, which appear as stripes when viewed in the along track direction. This has been attributed to the cross-track variability of the measured solar irradiances (Veihelmann and Kleipool, 2006). For OMI, this variability arises due to a combination of noise in the measured solar spectra, transient dark current signals and the angular and wavelength dependence of the diffuser used for irradiance measurements. The impact of these variations is significant for glyoxal due to its relatively weak absorption.

We investigated how solar spectrum variation impacts the OMI retrieval by adding noise to the solar spectra used in the retrieval OSSE (Fig. 6). A different Gaussian

noise realisation was added to each across-track solar spectrum. We chose a signal to noise ratio of 3000, to be roughly consistent with the noise level expected for averaging a month of OMI solar spectra. Striping is apparent in the slant columns retrieved from the synthetic spectra (Fig. 6b) with cross-track biases reaching as high as 1.5×10^{15} molecules cm^{-2}. Figure 6b also shows that the magnitudes of the stripes are constant with latitude. Thus determining the stripe offsets at one location should be sufficient for correcting the stripes at all locations.

The Sahara is a convenient region to determine the cross-track stripe offsets. Glyoxal concentrations in this region are expected to be negligible, with VCDs simulated by GEOS-Chem below 1.5×10^{13} molecules cm^{-2} all year. In addition, spectra over the Sahara have a high signal-to-noise ratio due to high surface reflectivity. Figure 6c shows the mean glyoxal SCD retrieved in the Saharan region defined by the limits 20–30° N, 10° W–30° E. Since there is essentially no glyoxal these represent the stripe offsets due to noise in the solar spectrum employed in the retrieval. Subtracting these offsets from Fig. 6a produces the stripe-corrected results in Fig. 6d. For reference, the synthetic data retrieved with noise-free solar spectrum is shown in Fig. 6a. The stripe-corrected retrieval is virtually indistinguishable from the noise-free case. We thus conclude that we can correct for stripes using this simple background subtraction approach, provided that the origin of the stripes is due purely to solar irradiance spectrum noise. This should generally be true, except for radiance/irradiance spectra impacted by so called random telegraph signals (RTS) caused by particle hits on the CCD. These lead to prolonged changes in dark current, which manifest as spikes in the observed spectra. To reduce the impact of RTS, we remove pixels that have been flagged as RTS in the level 1-B product (Kleipool, 2005). We identify additional spikes by comparing the residual difference between the modeled and measured spectra. Pixels in spectra whose residuals are 3 standard deviations from the mean are flagged as RTS. Spectra with these additionally flagged pixels are then refitted upon removal of the flagged pixels.

A particular stripe offset correction should apply for all spectra retrieved using the same OMI solar spectrum. Since the operational retrieval uses a 30 day running mean,

the stripe patterns should vary smoothly in time. For the real spectra, we must also consider how random noise in the radiance spectra propagates to random error in the stripe offsets. In principle this can be reduced by averaging retrievals over the normalisation region. We therefore create a time dependent offset for each track by taking a 5 day running mean of all retrieved slant columns for each track in the Saharan normalisation region. The 5 day window was chosen because this was the minimum window width required to reduce the uncertainty in the stripe offsets below 1×10^{14} molecules cm^{-2}. The associated stripe patterns for the month of July 2006 and their uncertainties are shown in Fig. 7. The magnitude of stripes determined from the real spectra are comparable in magnitude to those in the OSSE. We also see that the stripe pattern over the month time frame is fairly constant. Thus the 5 day averaging window appears small enough to capture the temporal variability of the stripe offsets.

To correct the stripes for a particular orbit, the derived stripe offset nearest in time is subtracted. Figure 8 shows the SCDs retrieved with real spectra from o10430 and an orbit taken on the same day over India (o10427) before and after the stripe correction is applied. Since the random uncertainty of the fits for individual spectra are large (O(10^{15}) molecules cm^{-2}) a 30 point running mean is applied to each track to aid visualisation of the stripe patterns. Figure 8 shows that the magnitudes of the stripes are significantly reduced upon applying the stripe offset correction. The correction performs similarly for both orbits, further evidence that the stripe patterns arise due to the common solar spectra employed in the retrieval. Thus the stripe correction offsets derived over the Sahara should apply globally.

3.3 Liquid water pre-fit

Retrieved glyoxal slant columns over clear oceanic waters are systematically negative when absorption from liquid water is not considered, due to anti-correlation between glyoxal and liquid water in the glyoxal fit window. Lerot et al. (2010) designed a two step retrieval procedure to correct for the impact of liquid water absorption, whereby liquid water is first derived in a larger fit window, and then held constant in the smaller

glyoxal fit window. We adopt the same approach for the OMI retrieval. In the retrieval of liquid water absorption, we additionally fit O_3, NO_2, and the O_2–O_2 collision complex. The liquid water retrieval uses a 1st order baseline and 5th order scaling polynomial. The higher order polynomial choice was needed to account for the impacts of surface reflectance over the broader fit window.

The sensitivity of the liquid water retrieval to window position was tested over two regions using real OMI spectra. The region over the Sahara used for the stripe correction for o10430 was selected for sensitivity tests to test a potential interference between liquid water absorption and surface reflectance from sand (Richter et al., 2011), which could induce errors in deriving the glyoxal background from this region. Sensitivity tests were also performed on an orbit taken on the same day (o10423), over the Pacific ocean between 0–30° N. This region contained significant liquid water absorption. Figure 9 shows the mean retrieved liquid water optical depth, total parameter error due to RCS uncertainty, and the spectrum fit residuals (adjusted for statistical degrees of freedom) for the two regions. Above 480 nm, retrievals over the Sahara become strongly negative. This is likely an artifact of the strong spectral dependence of reflectance from sandy surfaces, which contains a pronounced feature at approximately 480 nm (Richter et al., 2011). Optical depths over the clear ocean region are a maximum for a window setting of 397.5–470 nm, whilst the retrieval that minimises the spectrum fit RMS is for 410–467.5 nm. However in both these regions, retrievals over land are strongly negative. These biases may be explained by the large parameter uncertainty, mostly due the high uncertainty in the liquid water cross section. Extension of the lower window limit below 400 nm leads to a sharp reduction in parameter error, which occurs due to the strong increase of liquid water absorption below this wavelength (see Fig. 1). Retrieved liquid water optical depths over the Sahara in this spectral region are close to zero, suggesting that incorporation of the strong shoulder of the liquid water RCS below 400 nm acts to significantly reduce the cross correlation with the surface reflectance signal.

For the liquid water retrieval, we set the window interval at 385–470 nm. The retrieval window choice for liquid water was guided by not wanting to fit any unwanted surface

reflectance signals that could be corrected by the polynomials in the smaller glyoxal window. Since sand has a strong spectral feature at 480 nm, we do not consider upper window limits above this wavelength. In addition, the incorporation of the water absorption shoulder below 400 nm reduced the negative retrieval bias over sandy surfaces. Figure 10 shows the spatial distribution of liquid water absorption for July 2006 for the operational retrieval. Liquid water absorption peaks at the centers of ocean gyres. These regions are areas of low biological activity, and thus have very low turbidities, thus allowing long effective light paths through the ocean surface layer.

3.4 Glyoxal retrieval

In this section we test the sensitivity of the retrieval to fit window selection using real OMI spectra. All retrievals were performed using the pre-fitted liquid water optical depths described in the previous section. For each retrieval window, we also retrieved a set of orbits within a 5 day window of o10430 to determine the stripe offsets. The resulting mean SCDs as a function of window position are shown in Fig. 11. We show the mean SCD for all retrievals over land in o10430, as well as the Pacific Ocean region tested in for the liquid water retrieval. For the land case, the mean SCD is positive for lower wavelength bounds between 428 and 436 nm. For the Pacific Ocean sector, the region in fit window space containing positive SCD values shrinks, with negative mean SCDs retrieved for upper wavelength limits above 461 nm. This difference relative to the o10430 case is likely due to interference from liquid water absorption.

Significant differences exist between the SCD patterns seen in the real OMI spectra compared with the biases in the OSSE. The decrease in the mean SCD in the real spectra when the lower wavelength limit is extended below 435 nm encompasses a strong Fraunhofer line due to Hydrogen (434.047 nm), followed by two more lines associated with Iron and Calcium (430.790 and 430.774 nm respectively). The SCD decrease as the strong solar lines are included in the fit window could be a result of imperfect corrections for inelastic scattering, which was not simulated in the OSSE. The SCDs retrieved from OMI data are strongly negative using the 445–460 nm window that

was optimal for the OSSE. This could be a result of RCS uncertainties, which will have a larger impact on smaller fit windows. Errors due to RCS uncertainty are estimated using the first term in Eq. (13). Figure 12 shows the mean error on retrieved glyoxal induced by uncertainties in the RCSs of NO_2 and H_2O. Above 435 nm the mean estimated error from the NO_2 RCS increases rapidly, providing evidence that the negative SCDs retrieved for the shorter windows are impacted by RCS uncertainties.

For the operational retrieval, glyoxal is retrieved using a fit window set between 435 and 461 nm. The lower limit was selected to avoid the potential interference with the Ring effect. The upper wavelength limit was chosen as a balance between avoiding interference from the liquid water spectrum (favoring smaller windows) and reducing parameter error (favoring larger windows). This wavelength region is similar to previous studies (Lerot et al., 2010; Vrekoussis et al., 2010). The resulting glyoxal SCDs for July 2006 using the operational fit window are shown in Fig. 13. Retrieved SCDs remain negative over areas with strong liquid water absorption, even after the inclusion of the pre-fitted liquid water optical depths. Figure 14 plots the gridded SCDs in Fig. 13 against the liquid water optical depths in Fig. 10. There is a clear negative trend in glyoxal with increasing optical depth, with glyoxal columns scaling approximately linearly with liquid water optical depth. This behaviour is consistent with how errors in the liquid water RCS would impact the glyoxal SCD in the retrieval. These are expected to be large due to the fact that the reported uncertainties in the liquid water RCS are large (6–14 %), and the resolution of the RCS is 5 nm, far greater than the 0.65 FWHM of OMI.

4 Results and discussion

Figure 15 shows seasonally averaged glyoxal VCDs retrieved from OMI for 2007. The largest values are biomass burning regions, similar to retrievals from GOME-2 and SCIAMACHY (Vrekoussis et al., 2009; Lerot et al., 2010; Vrekoussis et al., 2010). There is also a large contribution from terrestrial biogenic sources, apparent in high

values of the Southeast US in summer and over Africa outside the biomass burning season. Assuming that all glyoxal lies within a a well mixed 2 km boundary layer, and an atmospheric scale height of 7.5 km, the Southeast US maximum corresponds to a surface mixing ratio of ~ 100 ppt. This is consistent with a mean glyoxal concentration of 83 ppt measured in Metter Georgia during June 1992 (Lee et al., 1995). The OMI retrievals are similarly broadly consistent with glyoxal concentrations measured from the ground in other areas of the United States. including Northern California (50–60 ppt for OMI vs. 25–70 ppt during BEARPEX 2007; Huisman et al., 2011) and Michigan (~ 35–40 ppt for OMI vs. 25 ppt during CABINEX 2009; Bryan et al., 2012).

Glyoxal columns in boreal forest regions of North America and Eurasia also show elevated values during summer, with VCDs in the range 2–4 × 10^{14} molecules cm^{-2}. These VCDs are slightly higher than those observed by GOME-2 and SCIAMACHY, which tend to peak at ~ 3 × 10^{14} molecules cm^{-2}(Vrekoussis et al., 2009; Lerot et al., 2010). Since these regions are predominantly composed of evergreen coniferous trees, these VCDs may be related to monoterpene emissions.

OMI shows persistently high glyoxal columns in the range 3–6 × 10^{14} molecules cm^{-2} over India and China. Elevated glyoxal levels are also seen in by GOME-2 and SCIAMACHY over these regions, with significantly elevated VCDs observed over most of Asia during summer (Vrekoussis et al., 2009; Lerot et al., 2010). For OMI this seasonality is weaker except over NE China. Here OMI columns tend to be lower on average by ~ 2 × 10^{14} molecules cm^{-2} compared to the other instruments. The VCDs from other retrievals may be greater due to the interference associated with boundary layer NO_2. The spatial pattern and timing of the broad summer glyoxal maximum over most of India and China follows observed water vapor patterns tied to the Indian and East Asian monsoons (Wang et al., 2014). Similarly, glyoxal columns are higher in the GOME-2 and SCIAMACHY retrievals over tropical ocean regions with elevated water vapor columns, and thus could represent a possible interference. Since all retrievals use water vapor RCSs calculated from the HITRAN database, differences likely arise from the choice of temperature and pressure used to derive the RCS. The sensitivity of our

retrieval to water vapor appears lower than the other retrievals, however further OSSE simulations accounting for H_2O temperature dependence are required to pinpoint the appropriate RCS choice.

Average VCDs over the Sahara for GOME-2 and SCIAMACHY range between 1 and 2 × 10^{14} molecules cm^{-2}(Vrekoussis et al., 2009; Lerot et al., 2010). OMI values are close to zero since the Sahara is used as a reference sector in the stripe-correction algorithm. Acetylene represents the only known long lived source of glyoxal (Fu et al., 2008), with a global mean lifetime of 12 days (Xiao et al., 2007). The resulting background glyoxal concentrations calculated in GEOS-Chem are of the order of 10^{13} molecules cm^{-2}. GEOS-Chem acetylene fields in the upper troposphere are in good agreement with observations (González Abad et al., 2011; Xiao et al., 2007), suggesting that these low background VCDs are reasonable. The high background values in SCIAMACHY and GOME-2 likely reflect the choice of offset correction region. In the case of GOME-2, the retrieval background is corrected using a Pacific Ocean reference sector (Lerot et al., 2010). Since interference from liquid water is likely not fully accounted for due to the large uncertainties in its RCS, determining offsets in regions with significant liquid water absorption may positively bias the offset correction applied.

Figure 16 shows the ratio of glyoxal to HCHO VCDs (R_{GF}) computed for OMI for July to August 2007, using OMI HCHO retrievals from González Abad et al. (2014). The highest R_{GF} values are observed in regions associated with biomass burning, and boreal areas associated with monoterpene emissions. The higher R_{GF} values are in biomass burning and boreal forest regions. These reflect the large emission of glyoxal from biomass burning and the high yields of glyoxal from Monoterpenes (Fu et al., 2008). Intermediate R_{GF} values (~ 0.04) are observed in northeast China, where anthropogenic emissions are expected to dominate. Regions associated with strong Isoprene emissions, such as the Southeast United States and Northern equatorial Africa tend to have lower R_{GF} (< 0.04) in agreement with ratios observed in forested environments in the United States (DiGangi et al., 2012). The range of biogenic and anthropogenic R_{GF} values retrieved by Vrekoussis et al. (2010) appear to be at odds with

OMI and surface measurements, however these were derived from multi-year averages. It is therefore likely that the R_{GF} values are representative of multiple sources. For instance, the higher R_{GF} observed by Vrekoussis et al. (2010) attributed to biogenic regions in Africa likely arises from a combination of pyrogenic and biogenic emissions. In the Southeast United States Vrekoussis et al. (2010) observe $R_{GF} < 0.04$. This is consistent with the other observations assuming that isoprene is the dominant year round source of glyoxal and HCHO in the Southeast United States.

5 Conclusions

We have developed a glyoxal retrieval for the Ozone Monitoring Instrument (OMI) aboard the NASA Aura satellite. The new retrieval takes advantage of the higher spatiotemporal resolution of OMI ($13 \times 24\,km^2$ nadir pixels, daily global coverage) relative to previous satellite sensors used to retrieve glyoxal.

We began by testing the retrieval algorithm against simulated OMI spectra. The results show that retrievals that include only one (stratospheric) NO_2 RCS cannot sufficiently account for its temperature dependence. Not including a second NO_2 RCS at higher temperature leads to an overestimation of glyoxal over regions with high levels of boundary layer NO_x. The OSSE results are consistent with those derived from real spectra (Alvarado et al., 2013).

We then used the synthetic spectra to inform the design of a new cross-track bias correction for OMI glyoxal. It was shown that determining the cross track bias over the Sahara is sufficient for determining the correction for all orbits that use the same solar spectrum. The method was applied to real OMI spectra and significantly reduced the magnitude of the cross-track biases.

The sensitivity of the liquid water retrieval to fit window position was tested by systematically varying the lower and upper wavelength limits of the retrieval window. We found that upper window limits above 480 nm lead to strongly negative optical depths over desert regions. We attribute this to a strong absorption feature near 480 nm from

sandy surfaces. We determined that the interference with sand could be minimised by using a fit window of 385–470 nm, which avoids the sand spectral feature and incorporates a strong feature in the liquid water spectrum below 400 nm.

The sensitivity of the glyoxal retrieval position was also tested. It was found that retrieved SCDs systematically decrease for lower window limits below 435 nm, and speculate that this is due to interferences from the Ring effect. We compared SCDs retrieved over land and a remote ocean region containing strong liquid water absorption to show that liquid water interferes strongly for upper window limits above 461 nm. We estimated retrieval errors caused by errors in the RCSs of H_2O and NO_2 to show that errors induced by these cross sections increase significantly for lower window limits 435 nm. We determined the optimal window to be 435–461 nm, similar to previous windows used by satellites.

We present a year (2007) of glyoxal VCDs retrieved from OMI. The new results are broadly consistent with the few surface observations available. The background of the OMI retrieval is lower than previous studies, and the source over tropical oceans is less significant than in previous work. We present OMI glyoxal-to-HCHO ratios for the summer of 2007. The observed values over biogenic and anthropogenic source regions are consistent with ground based observations.

Acknowledgements. This study was funded by NASA through the Aura Science Team and by a Frank Knox Memorial Fellowship awarded to C. Chan Miller.

References

Abbot, D. S., Palmer, P. I., Martin, R. V., Chance, K. V., Jacob, D. J., and Guenther, A.: Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space, Geophys. Res. Lett., 30, 1886, doi:10.1029/2003GL017336, 2003. 6067

Aliwell, S. R., Van Roozendael, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P., Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J.-C.,

Pfeilsticker, K., and Pundt, I.: Analysis for BrO in zenith-sky spectra: an intercomparison exercise for analysis improvement, J. Geophys. Res.-Atmos., 107, ACH10.1–ACH10.20, doi:10.1029/2001JD000329, 2002. 6071

Alvarado, L., Richter, A., Vrekoussis, M., Wittrock, F., Hilboll, A., and Burrows, J.: Sensitivity study on glyoxal retrievals from OMI data (Poster), in: 2013 International DOAS Workshop, 2013. 6067, 6078, 6087

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res.-Atmos., 118, 2041–2057, doi:10.1002/jgrd.50059, 2013. 6067

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. 6096

Bryan, A. M., Bertman, S. B., Carroll, M. A., Dusanter, S., Edwards, G. D., Forkel, R., Griffith, S., Guenther, A. B., Hansen, R. F., Helmig, D., Jobson, B. T., Keutsch, F. N., Lefer, B. L., Pressley, S. N., Shepson, P. B., Stevens, P. S., and Steiner, A. L.: In-canopy gas-phase chemistry during CABINEX 2009: sensitivity of a 1-D canopy model to vertical mixing and isoprene chemistry, Atmos. Chem. Phys., 12, 8829–8849, doi:10.5194/acp-12-8829-2012, 2012. 6085

Chance, K.: Analysis of BrO measurements from the Global Ozone Monitoring Experiment, Geophys. Res. Lett., 25, 3335–3338, 1998. 6069

Chance, K.: Remote Sensing of the Atmosphere for Environmental Security: Spectroscopic Measurements of Tropospheric Composition from Satellite Measurements in the Ultraviolet and Visible: Steps Toward Continuous Pollution Monitoring From Space, NATO Security through Science, Springer Netherlands, 2006. 6067, 6078

Chance, K. and Kurucz, R.: An improved high-resolution solar reference spectrum for earth's atmosphere mea- surements in the ultraviolet, visible, and near infrared, J. Quant. Spectrosc. Ra., 111, 1289–1295, 2010. 6069

Chance, K., Kurosu, T. P., and Sioris, C. E.: Undersampling correction for array detector-based satellite spectrometers, Appl. Optics, 44, 1296–1304, 2005. 6070

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. Optics, 36, 5224–5230, 1997. 6070

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. 6067

DiGangi, J. P., Henry, S. B., Kammrath, A., Boyle, E. S., Kaser, L., Schnitzhofer, R., Graus, M., Turnipseed, A., Park, J-H., Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin III, R. L., Kim, S., Nakashima, Y., Wolfe, G. M., Kajii, Y., Apel, E.C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A., and Keutsch, F. N.: Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, Atmos. Chem. Phys., 12, 9529–9543, doi:10.5194/acp-12-9529-2012, 2012. 6067, 6086

Dirksen, R., Dobber, M., Voors, R., and Levelt, P.: Prelaunch characterization of the Ozone Monitoring Instrument transfer function in the spectral domain, Appl. Optics, 45, 3972–3981, 2006. 6070

Forster, P., Ramaswamy, V., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Nganga, J., R. Prinn, G. R., Schulz, M., and Dorland, R. V.: Changes in Atmospheric Constituents and in Radiative Forcing, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007. 6067

Fu, T.-M., Jacob, D. J., Palmer, P. I., Chance, K., 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. 6067

Fu, T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M., and Henze, D. K.: Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, J. Geophys. Res., 113, D15303, doi:10.1029/2007JD009505, 2008. 6074, 6086

González Abad, G., Allen, N. D. C., Bernath, P. F., Boone, C. D., McLeod, S. D., Manney, G. L., Toon, G. C., Carouge, C., Wang, Y., Wu, S., Barkley, M. P., Palmer, P. I., Xiao, Y., and Fu, T. M.: Ethane, ethyne and carbon monoxide concentrations in the upper troposphere and lower

stratosphere from ACE and GEOS-Chem: a comparison study, Atmos. Chem. Phys., 11, 9927–9941, doi:10.5194/acp-11-9927-2011, 2011. 6086

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated SAO OMI formaldehyde retrieval, Atmos. Meas. Tech. Discuss., 7, 1–31, doi:10.5194/amtd-7-1-2014, 2014. 6073, 6086

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, doi:10.5194/acp-6-3181-2006, 2006. 6067

Huisman, A. J., Hottle, J. R., Galloway, M. M., DiGangi, J. P., Coens, K. L., Choi, W., Faloona, I. C., Gilman, J. B., Kuster, W. C., de Gouw, J., Bouvier-Brown, N. C., Goldstein, A. H., LaFranchi, B. W., Cohen, R. C., Wolfe, G. M., Thornton, J. A., Docherty, K. S., Farmer, D. K., Cubison, M. J., Jimenez, J. L., Mao, J., Brune, W. H., and Keutsch, F. N.: Photochemical modeling of glyoxal at a rural site: observations and analysis from BEARPEX 2007, Atmos. Chem. Phys., 11, 8883–8897, doi:10.5194/acp-11-8883-2011, 2011. 6085

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G., and Wilson, J.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., 5, 1053–1123, doi:10.5194/acp-5-1053-2005, 2005. 6066

Kleipool, Q.: Transient signal flagging algorithm definition for radiance data, Tech. Rep. TN-OMIE-KNMI-717 TN-OMIE-KNMI-717 TN-OMIE-KNMI-717 TN-OMIE-KNMI-717 TN-OMIE-KNMI-717, Royal Netherlands Meteorological Institute, 2005. 6080

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res.-Atmos., 113, D18308, doi:10.1029/2008JD010290, 2008. 6073, 6077

Klimont, Z., Streets, D. G., Gupta, S., Cofala, J., Lixin, F., and Ichikawa, Y.: Anthropogenic emissions of non-methane volatile organic compounds in China, Atmos. Environ., 36, 1309–1322, doi:10.1016/S1352-2310(01)00529-5, 2002. 6067

Lee, Y.-N., Zhou, X., and Hallock, K.: Atmospheric carbonyl compounds at a rural southeastern United States site, J. Geophys. Res.-Atmos., 100, 25933–25944, 1995. 6085

Lelieveld, J. and Dentener, F. J.: What controls tropospheric ozone?, J. Geophys. Res.-Atmos., 105, 3531–3551, 2000. 6067

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. 6067, 6070, 6081, 6084, 6085, 6086

Lindström, P. and Wedin, P.: Methods and Software for Nonlinear Least Squares Problems, Tech. rep., Inst. of Information Processing, University of Umeå, Sweden, 1988. 6074

Lippmann, M.: Health effects of ozone: a critical review, JAPCA, 39, 672–695, 1989. 6066

Liu, X., Chance, K., Sioris, C. E., and Kurosu, T. P.: Impact of using different ozone cross sections on ozone profile retrievals from Global Ozone Monitoring Experiment (GOME) ultraviolet measurements, Atmos. Chem. Phys., 7, 3571–3578, doi:10.5194/acp-7-3571-2007, 2007. 6078

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10, 2521–2537, doi:10.5194/acp-10-2521-2010, 2010. 6077

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. 6067

McLinden, C. A., Olsen, S. C., Hannegan, B., Wild, O., Prather, M. J., and Sundet, J.: Stratospheric ozone in 3-D models: a simple chemistry and the cross-tropopause flux, J. Geophys. Res.-Atmos., 105, 14653–14665, doi:10.1029/2000JD900124, 2000. 6077

Miller, C. C., Jacob, D., Furlow, M., Keutsch, F., Lerot, C., and DeSmedt, I.: Development of a Global Glyoxal Simulation for Constraining NMVOC Emissions, in: Gordon Research Conference on Biogenic Hydrocarbons, 2012. 6074

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K., 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.-Atmos., 113, D02307, doi:10.1029/2007JD008950, 2008. 6067

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res.-Atmos., 106, 14539–14550, doi:10.1029/2000JD900772, 2001. 6067, 6068, 6073

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., 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.-Atmos., 111, D12315, doi:10.1029/2005JD006689, 2006. 6067

Parrish, D. D., Ryerson, T. B., Mellqvist, J., Johansson, J., Fried, A., Richter, D., Walega, J. G., Washenfelder, R. A., de Gouw, J. A., Peischl, J., Aikin, K. C., McKeen, S. A., Frost, G. J., Fehsenfeld, F. C., and Herndon, S. C.: Primary and secondary sources of formaldehyde in urban atmospheres: Houston Texas region, Atmos. Chem. Phys., 12, 3273–3288, doi:10.5194/acp-12-3273-2012, 2012. 6067

Pfister, G. G., Emmons, L. K., Hess, P. G., Lamarque, J.-F., Orlando, J. J., Walters, S., Guenther, A., Palmer, P. I., and Lawrence, P. J.: Contribution of isoprene to chemical budgets: A model tracer study with the NCAR CTM MOZART-4, J. Geophys. Res.-Atmos., 113, D05308, doi:10.1029/2007JD008948, 2008. 6067

Pope, R. M. and Fry, E. S.: Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements, Appl. Optics, 36, 8710–8723, 1997. 6096

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. 6082

Rogers, C.: Inverse methods for atmospheric sounding, Vol. 2 of Atmospheric, Oceanic and Planetary Physics, World Scientific, 2000. 6074

Rothman, L., Gordon, I., Barbe, A., Benner, D., Bernath, P., Birk, M., Boudon, V., Brown, L., Campargue, A., Champion, J.-P., Chance, K., Coudert, L., Dana, V., Devi, V., Fally, S., Flaud, J.-M., Gamache, R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W., Mandin, J.-Y., Massie, S., Mikhailenko, S., Miller, C., Moazzen-Ahmadi, N., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Predoi-Cross, A., Rinsland, C., Rotger, M., Šimečková, M., Smith, M., Sung, K., Tashkun, S., Tennyson, J., Toth, R., Vandaele, A., and Auwera, J. V.: The {HITRAN} 2008 molecular spectroscopic database, J. Quant. Spectrosc. Ra., 110, 533–572, 2009. 6096

Sinreich, R., Ortega, I., and Volkamer, R.: Sensitivity Study of Glyoxal Retrievals at Different Wavelength Ranges (Poster), in: 2013 International DOAS Workshop, 2013. 6067

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, 2009. 6067

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res.-Atmos., 108, 8809, doi:10.1029/2002JD003093, 2003. 6067

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O_2–O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15, 15371–15381, doi:10.1039/C3CP50968K, 2013. 6096

Vandaele, A., Hermans, C., Fally, S., Carleer, M., Mérienne, M.-F., Jenouvrier, A., Coquart, B., and Colin, R.: Absorption cross-sections of NO_2: simulation of temperature and pressure effects, J. Quant. Spectrosc. Ra., 76, 373–391, 2003. 6078, 6096

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.-Atmos., 107, ACH3.1–ACH3.12, doi:10.1029/2001JD000971, 2002. 6096

Veihelmann, B. and Kleipool, Q.: Reducing along-track stripes in OMI Level 2 products, Tech. Rep. TN-OMIE-KNMI-785, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands, 2006. 6079

Vogel, L., Sihler, H., Lampel, J., Wagner, T., and Platt, U.: Retrieval interval mapping: a tool to visualize the impact of the spectral retrieval range on differential optical absorption spectroscopy evaluations, Atmos. Meas. Tech., 6, 275–299, doi:10.5194/amt-6-275-2013, 2013. 6068, 6078

Volkamer, R., Spietz, P., Burrows, J., and Platt, U.: High-resolution absorption cross-section of glyoxal in the UV–vis and {IR} spectral ranges, J. Photochem. Photobiol. A, 172, 35–46, 2005. 6096

Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: Temporal and spatial variability of glyoxal as observed from space, Atmos. Chem. Phys., 9, 4485–4504, doi:10.5194/acp-9-4485-2009, 2009. 6070, 6084, 6085, 6086

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. 6067, 6084, 6086, 6087

Wang, H., Liu, X., Chance, K., Gonzalez Abad, G., and Chan Miller, C.: Water vapor retrieval from OMI visible spectra, Atmos. Meas. Tech. Discuss., 7, 541–567, doi:10.5194/amtd-7-541-2014, 2014. 6085

Wei, W., Wang, S., Chatani, S., Klimont, Z., Cofala, J., and Hao, J.: Emission and speciation of non-methane volatile organic compounds from anthropogenic sources in China, Atmos. Environ., 42, 4976–4988, 2008. 6067

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, doi:10.1029/2006GL026310, 2006. 6067

Xiao, Y., Jacob, D. J., and Turquety, S.: Atmospheric acetylene and its relationship with CO as an indicator of air mass age, J. Geophys. Res.-Atmos., 112, D12305, doi:10.1029/2006JD008268, 2007. 6086

Table 1. Reference cross sections used in this study.

Molecule	Uncertainty[a] (%)	Reference
O_3 (228 K)	2	Brion et al. (1998)
NO_2 (231, 293 K)	1.1[b]	Vandaele et al. (2003)
Glyoxal (296 K)	3	Volkamer et al. (2005)
H_2O (280 K)	5	Rothman et al. (2009)
O_2-O_2 (293 K)	3.35	Thalman and Volkamer (2013)
H_2O (liquid, 295 K)	6–14[c]	Pope and Fry (1997)

[a] Relative uncertainties given in table are those reported in the literature references.
[b] Based on variance of the relative difference of two NO_2 RCS measurements in Vandaele et al. (2002).
[c] Reported uncertainty range is for 435–460 nm.

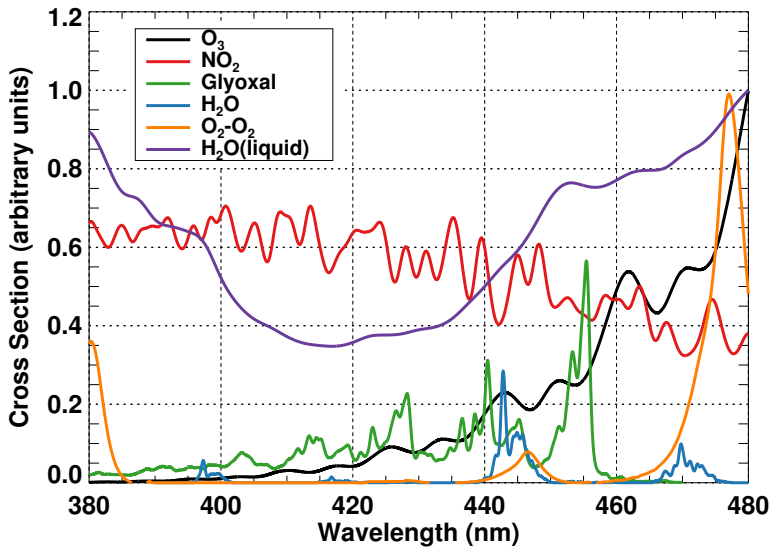

Figure 1. Absorption cross sections for molecules listed in Table 1 degraded to the resolution of OMI. Values have been normalized to O(1) for display purposes.

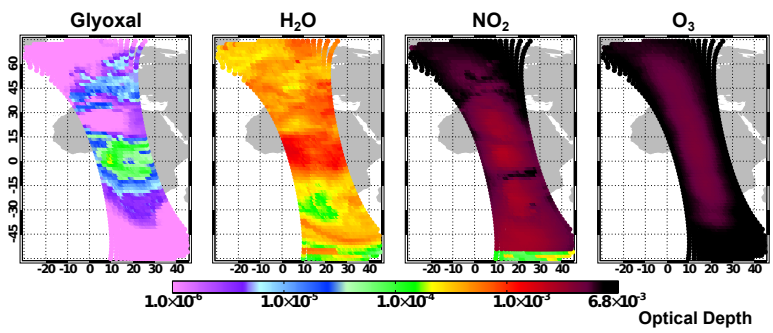

Figure 2. Simulated optical depths of important species in the glyoxal fitting region for 1 July 2006. The values correspond to the mean optical depths within the 430–460 nm spectral region. Results were simulated using the VLIDORT radiative transfer model ran using the viewing geometry of OMI combined with GEOS-Chem trace gas profiles.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

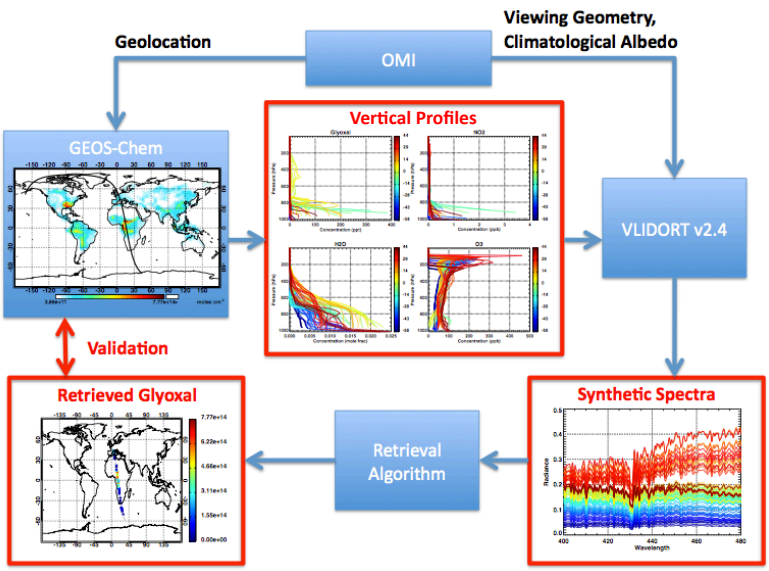

Figure 3. Schematic of the OSSE approach used to optimize the retrieval algorithm.

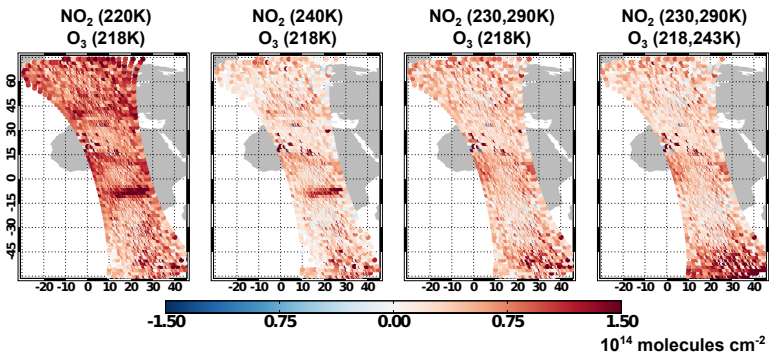

Figure 4. Differences between simulated and true glyoxal slant column densities (molecules cm^{-2}) retrieved using NO_2 and O_3 RCSs at different temperatures. Plot titles indicate temperatures of the RCSs used in the retrieval.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

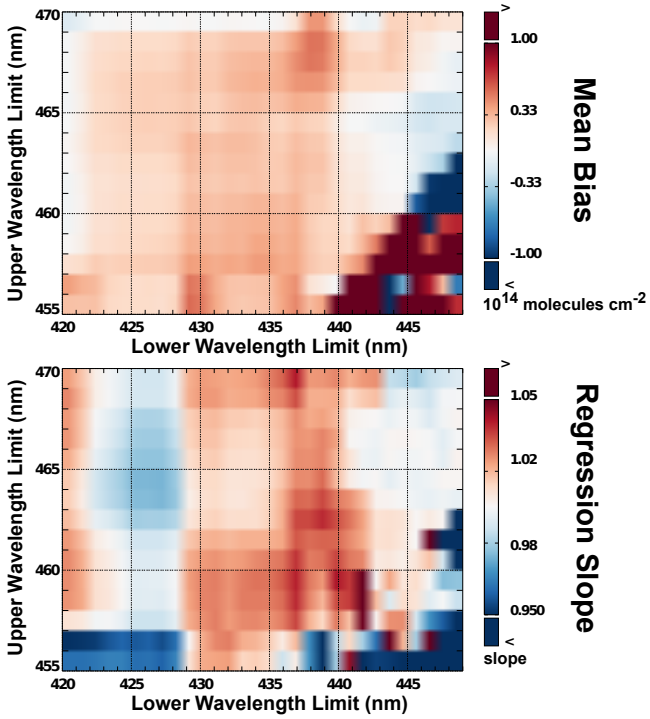

Figure 5. Sensitivity of glyoxal retrieval error to choice of fitting window. Results are from the OSSE described in the text. The top panel shows the mean bias between retrieved glyoxal columns and the true values. The bottom panel shows the slope of the linear regression of retrieved vs. true glyoxal columns.

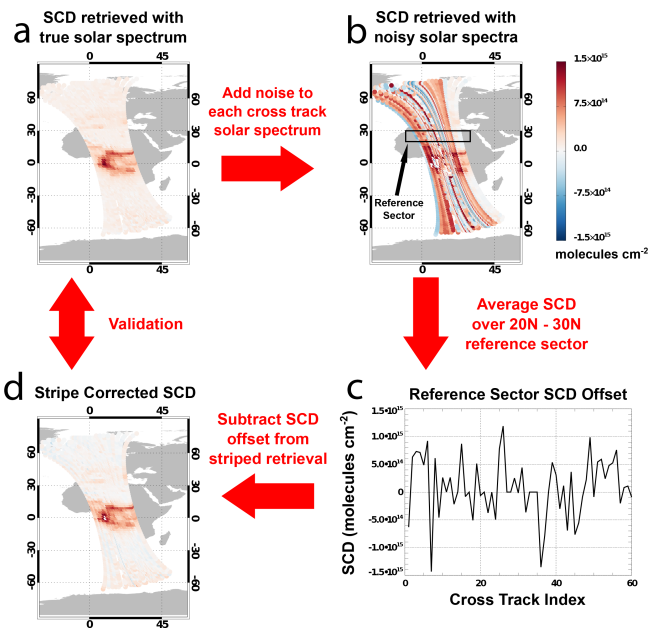

Figure 6. Simulation of stripes in the glyoxal retrieval OSSE. **(a)** Slant column densities (molecules cm^{-2}) retrieved from synthetic spectra, using a noise free solar spectrum. **(b)** Slant column densities (molecules cm^{-2}) retrieved from synthetic spectra, retrieved from synthetic spectra using solar spectra with a 3000 signal-to-noise ratio. **(c)** Mean slant column (molecules cm^{-2}) retrieved in the region 20–30° N, 10° W–30° E from **(b)**. **(d)** The stripe corrected retrieval, derived from subtracting the mean slant columns in **(c)** from the columns retrieved using the noisy solar spectra.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

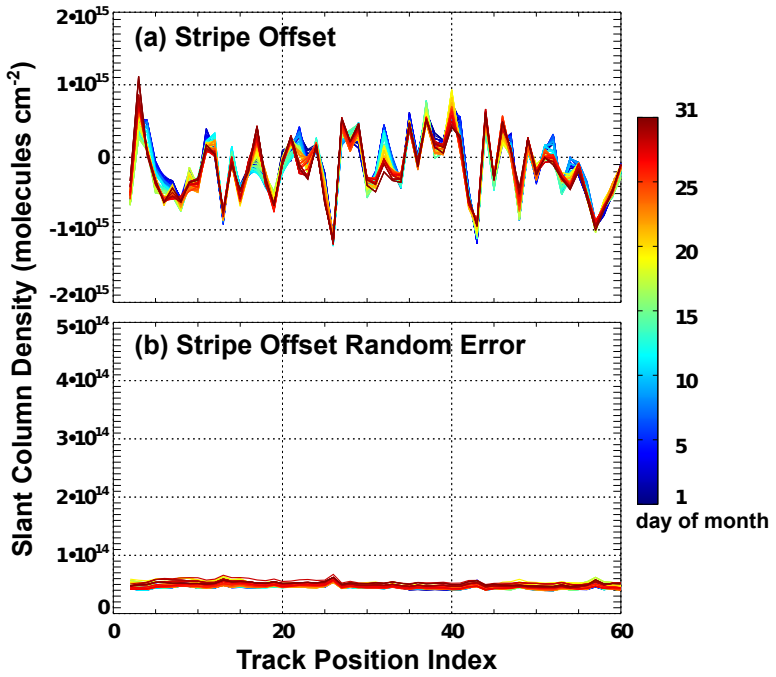

Figure 7. (a) Stripe offset (molecules cm^{-2}) derived from a running 5 day mean of the slant columns retrieved over the sahara (20–30° N, 10° W–30° E) for July 2006 as a function of cross track position. Lines are color coded by day of the month. **(b)** Estimate of random SCD error (molecules cm^{-2}) for **(a)** induced from noise in the measured radiances.

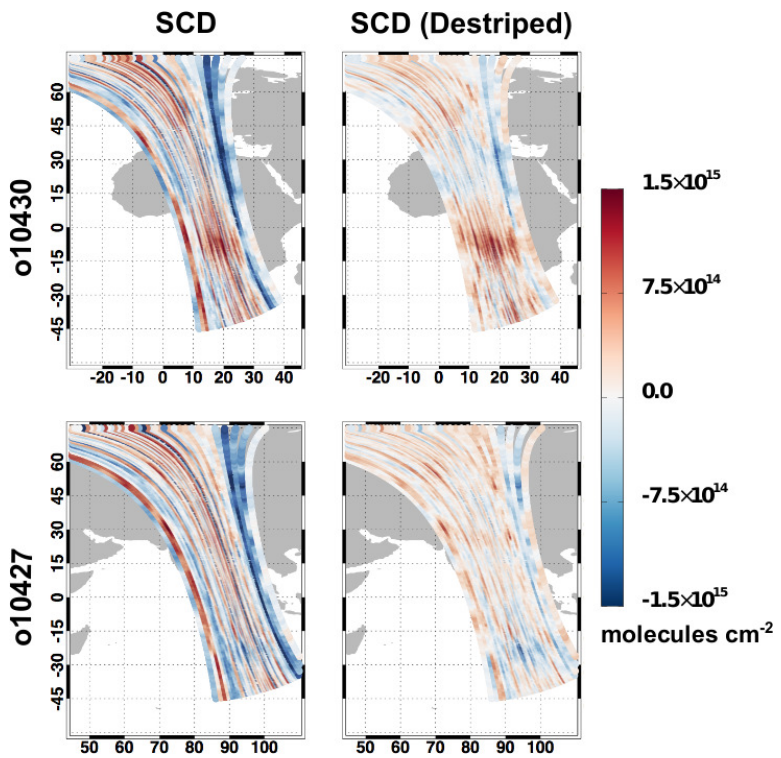

Figure 8. Left: retrieved glyoxal slant column densities (molecules cm^{-2}) using OMI data from orbits o10430 and o10427. Right: retrieved slant column density (molecules cm^{-2}) corrected for striping by subtracting the mean retrieved slant column between 20–30° N for each cross track position.

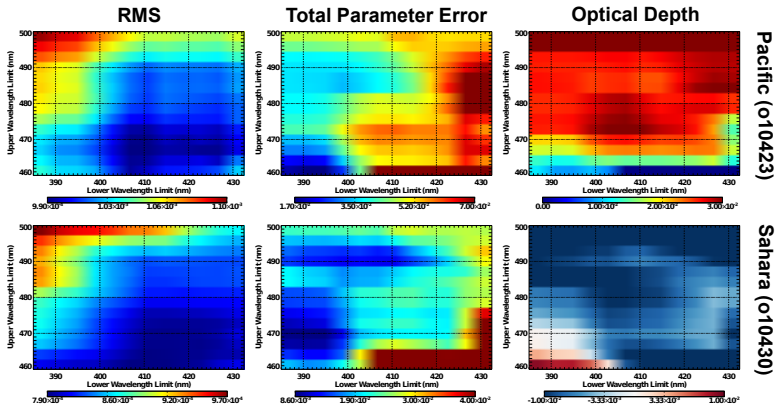

Figure 9. Retrieval window interval maps for spectra over the remote Pacific (top) and Sahara (bottom). Fitting RMS are adjusted for statistical degrees of freedom cf. Eq. (19). Liquid water optical depths and associated parameter errors are those at 460 nm.

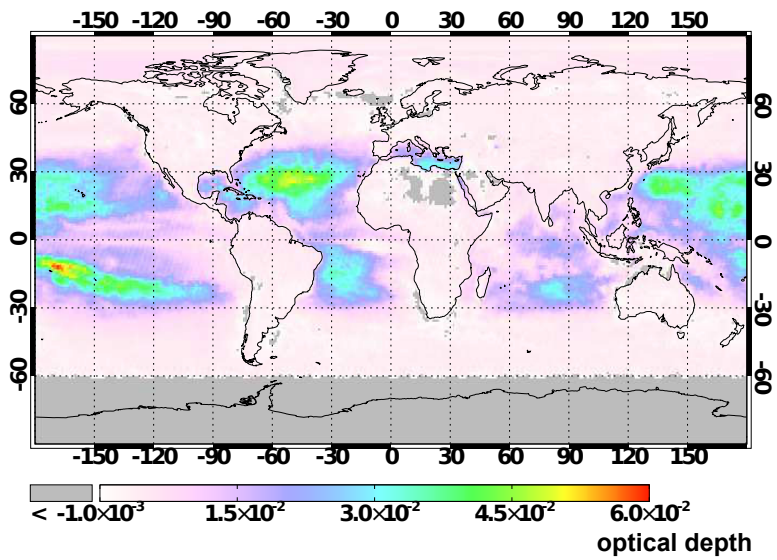

Figure 10. Retrieved liquid water optical depth at 460 nm for July 2006.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

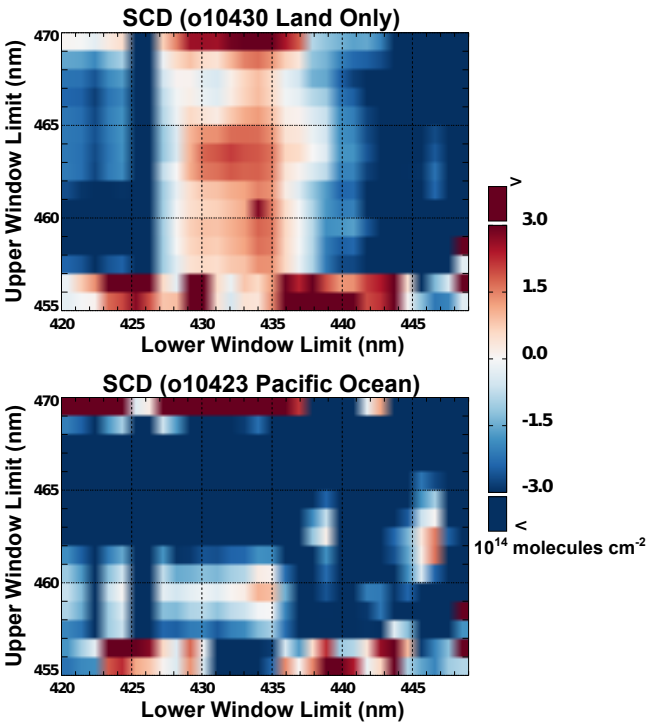

Figure 11. Mean glyoxal Slant column density retrieved as a function of window limits. Above – mean SCD retrieved for all retrievals over land from o10430. Below – mean SCD retrieved from 0–30° N over the pacific ocean (o10423).

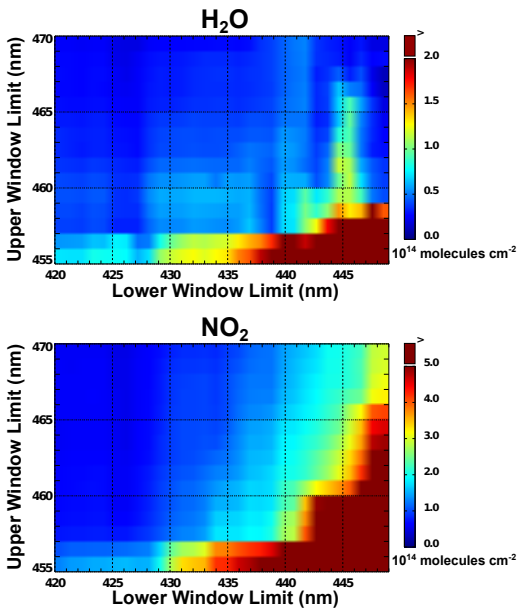

Figure 12. Estimate of mean glyoxal SCD error (molecules cm^{-2}) for all o10430 land pixels due to RCS uncertainties in H_2O (above) and NO_2 (below).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

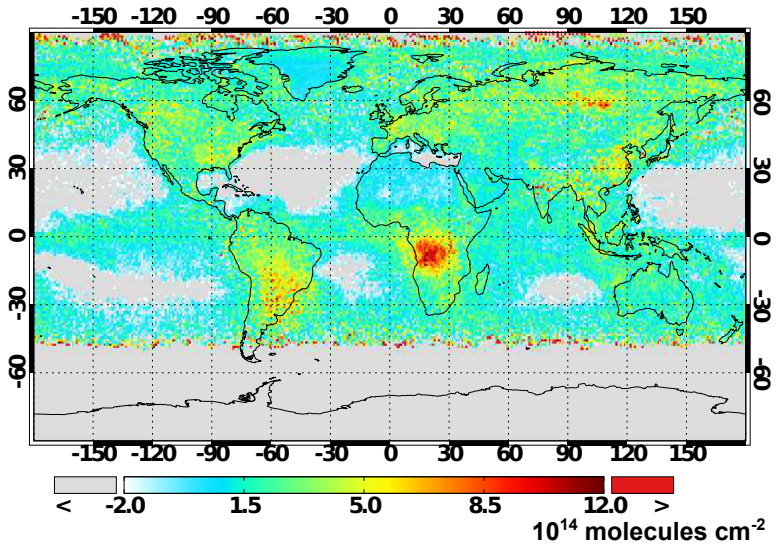

Figure 13. Monthly mean glyoxal slant column densities retrieved for July 2006 (molecules cm^{-2}).

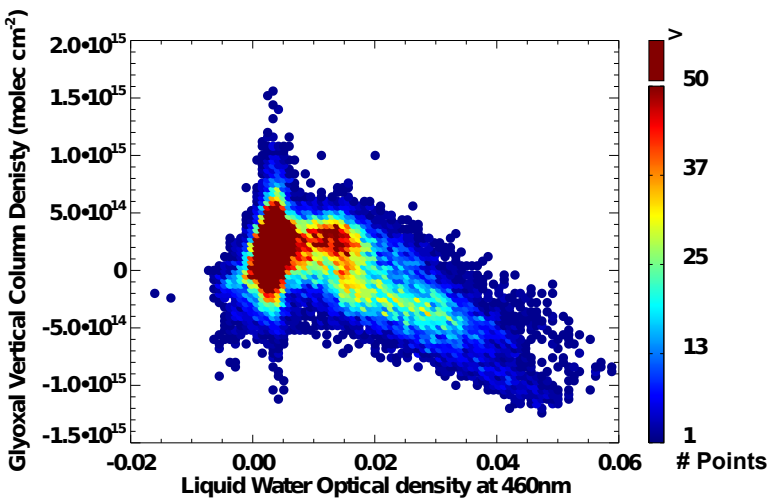

Figure 14. Scatterplot of glyoxal slant column densities (molecules cm^{-2}) retrieved for July 2006 and liquid water optical depth at 460 nm. Points are color coded by the number of pixels falling within the grid cell corresponding to the point in the figure.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

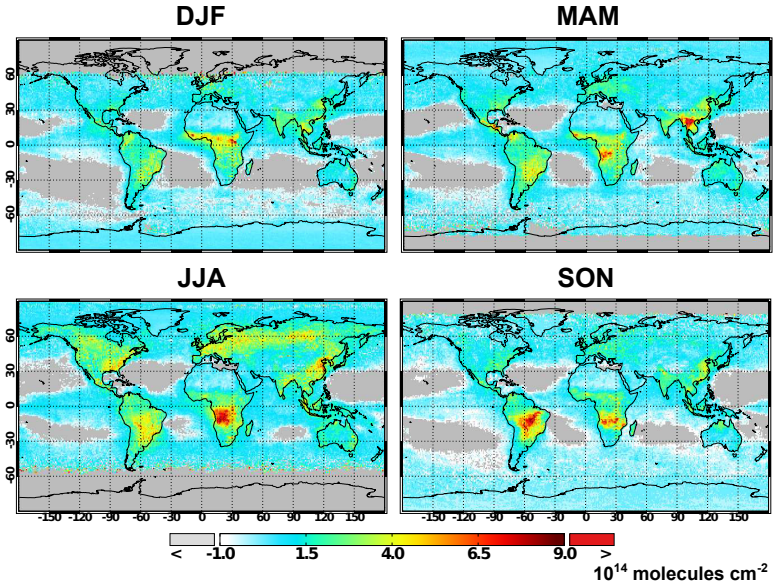

Figure 15. Seasonally averaged glyoxal VCDs (molecules cm^{-2}) retrieved from OMI for 2007.

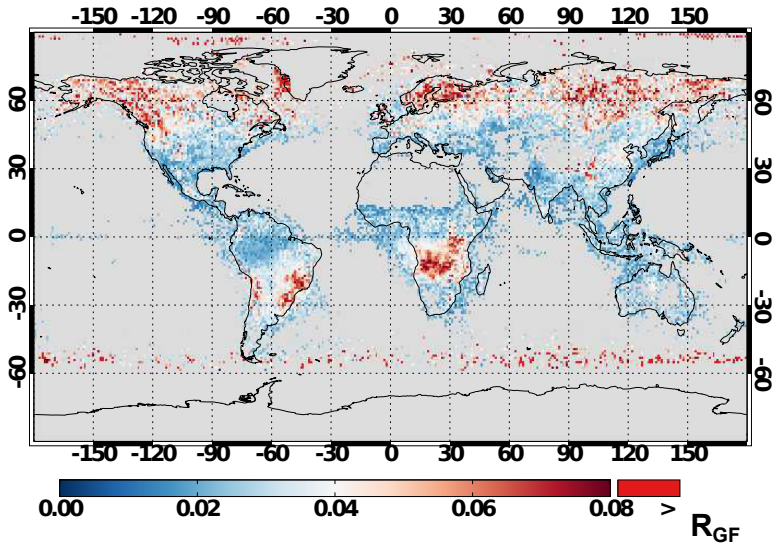

Figure 16. Ratio of glyoxal to formaldehyde VCD from OMI averaged for June to August 2007.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |