The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling

T. Stavrakou¹, J.-F. Müller¹, I. De Smedt¹, M. Van Roozendael¹, M. Kanakidou², M. Vrekoussis³, F. Wittrock³, A. Richter³, and J. P. Burrows³,⁴

¹Belgian Institute for Space Aeronomy, Avenue Circulaire 3, 1180, Brussels, Belgium
²ECPL, Department of Chemistry, University of Crete, Heraklion, Greece
³Institute of Environmental Physics, University of Bremen, Bremen, Germany
⁴Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

Received: 15 April 2009 – Accepted: 5 June 2009 – Published: 19 June 2009
Correspondence to: T. Stavrakou (jenny@aeronomie.be)

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

Tropospheric glyoxal and formaldehyde columns retrieved from the SCIAMACHY satellite instrument in 2005 are used with the IMAGESv2 global chemistry-transport model and its adjoint in a two-compound inversion scheme designed to estimate the continental source of glyoxal. The formaldehyde observations provide an important constraint on the production of glyoxal from isoprene in the model, since the degradation of isoprene constitutes an important source of both glyoxal and formaldehyde. Current modelling studies underestimate largely the observed glyoxal satellite columns, pointing to the existence of an additional land glyoxal source of biogenic origin. We include an extra glyoxal source in the model and we explore its possible distribution and magnitude through two inversion experiments. In the first case, the additional source is represented as a direct glyoxal emission, and in the second, as a secondary formation through the oxidation of an unspecified glyoxal precursor. Besides this extra source, the inversion scheme optimizes the primary glyoxal and formaldehyde emissions, as well as their secondary production from other identified non-methane volatile organic precursors of anthropogenic, pyrogenic and biogenic origin.

In the first inversion experiment, the additional direct source, estimated at 36 Tg/yr, represents 38% of the global continental source, whereas the contribution of isoprene is equally important (30%), the remainder being accounted for by anthropogenic (20%) and pyrogenic fluxes. The inversion succeeds in reducing the underestimation of the glyoxal columns by the model, but it leads to a severe overestimation of glyoxal surface concentrations in comparison with in situ measurements. In the second scenario, the inferred total global continental glyoxal source is estimated at 108 Tg/yr, almost two times higher than the global a priori source. The extra secondary source is the largest contribution to the global glyoxal budget (50%), followed by the production from isoprene (26%) and from anthropogenic NMVOC precursors (14%). A better performance is achieved in this case, as the updated emissions allow for a satisfactory agreement of the model with both satellite and in situ glyoxal observations.
1 Introduction

Glyoxal (CHOCHO) and formaldehyde (HCHO) are short-lived intermediate products in the oxidation of non-methane volatile organic compounds (NMVOCs) emitted by vegetation, fires and anthropogenic activities. They are also directly emitted during fossil fuel and biofuel combustion and biomass burning. Both compounds absorb in the UV-visible spectral region and have been measured by the SCIAMACHY satellite sensor since 2003, and more recently, by the OMI and GOME-2 instruments.

Although current modelling studies succeed in reproducing reasonably well the variability and the magnitude of the observed HCHO columns (Fu et al., 2007; Barkley et al., 2008; Stavrakou et al., 2009a), a large underestimation of the glyoxal columns has been reported (Myriokefalitakis et al., 2008; Fu et al., 2008). These discrepancies point to the existence of additional land and marine sources, and/or to an underestimation of the known sources. This study aims at determining the global distribution and strength of the glyoxal continental sources. To this purpose, we utilize synergistically HCHO and CHOCHO satellite columns in a two-compound inverse modelling framework based on the adjoint of the IMAGES model (Müller and Stavrakou, 2005).

In this approach, information gained from HCHO can help to constrain the sources of CHOCHO, since both gases have common precursors and are interrelated through chemistry (e.g. through the OH fields).

The grid-based inverse modelling approach used in this study is designed to estimate the fluxes emitted from every model grid cell for each emission category (Stavrakou and Müller, 2006). The adjoint model technique is used to minimize the bias between the model and the observations, while accounting for the errors on the observations and on the emission variables to be determined. This method has been recently used to constrain the pyrogenic and biogenic NMVOC sources of HCHO based on a 4-year satellite dataset (Stavrakou et al., 2009b). A similar setup will be used in this study in order to derive a “top-down” inventory of the glyoxal land fluxes.
In the current inversion framework, primary formaldehyde and glyoxal emission and secondary production from NMVOC precursors (anthropogenic, biogenic and pyrogenic) are optimized for the year 2005, as well as an additional missing source of glyoxal of biogenic origin. Two inversion experiments are investigated. In the first, the extra source is assumed to be directly emitted glyoxal, and in the second, glyoxal is formed through oxidation of an unknown glyoxal precursor (UVOC) with a lifetime of five days. The impact of the choice for the UVOC lifetime is assessed through sensitivity simulations. The magnitudes and distributions of the glyoxal sources inferred by the inversion case studies are discussed, and the optimized glyoxal concentrations are evaluated against in situ measurements.

The UVOC precursor of the second inversion should not be considered as a single molecular species, but as a generic species possibly representing a collection of compounds leading to glyoxal formation after a number of non-specified intermediate steps. These compounds could be directly emitted or produced in the oxidation of short-lived biogenic precursors like terpenoids. Detailed knowledge on the oxidation mechanisms and glyoxal yield of such compounds is still lacking. Moreover, the release of glyoxal from organic aerosols is another potentially significant source of glyoxal.

The next section provides a description of the IMAGESv2 model. Particular focus is given on the description of the chemistry of the glyoxal precursors (Sect. 2.3) and to the formation of secondary organic aerosols (Sect. 2.4) due to its importance as a glyoxal loss process. A short description of the glyoxal and formaldehyde observed columns is given in Sect. 2.5. The global glyoxal budget calculated with the model is discussed in Sect. 3. The setup for the inversion and the optimization results are presented in Sect. 4. Conclusions are drawn in the last section.
2 The model and the measurements

2.1 The IMAGESv2 CTM

The global IMAGESv2 chemistry-transport model (Müller and Stavrakou, 2005; Stavrakou et al., 2009a) is run at the horizontal resolution of $4^\circ \times 5^\circ$ and is vertically discretized at 40 sigma-pressure levels extending from the surface to 44 hPa. The model calculates the concentrations of 80 trace gases using a one-day time step. Advection is driven by monthly mean ECMWF reanalysed wind fields, whereas daily ECMWF fields are used for temperature, water vapor, cloud optical depths, and planetary boundary layer mixing. The model accounts for the impact of diurnal variations of the chemical compounds through correction factors computed via a diurnal cycle simulation with a 20-min time step. The diurnal profiles are also used to estimate the CHOCHO and HCHO concentration at the satellite overpass time (10:00 LT) from the daily averaged values calculated with a time step of one day. They depend on the location and the month. The comparison with ground-based observations uses the diurnally averaged glyoxal concentrations. The model is confronted with the data following a 4-month initialization. The calculated monthly averaged HCHO columns account for the sampling times of observations at each location.

2.2 NMVOC emissions

The model comprises emissions from 20 explicit NMVOCs (ethene, acetylene, isoprene, $\alpha$-pinene, glycolaldehyde, ethane, propane, propene, formaldehyde, glyoxal, acetaldehyde, methylglyoxal, acetone, methyl vinyl ketone, biacetyl, acetic acid, methanol, benzene, toluene and xylene). All NMVOCs are formaldehyde precursors, and among them, ethene, acetylene, isoprene, $\alpha$-pinene, glycolaldehyde, benzene, toluene, and xylene are precursors of glyoxal.

Biogenic emissions of isoprene are obtained from the MEGAN-ECMWF inventory (Müller et al., 2008). Their global estimate for 2005 is equal to 449 Tg. Emissions of
monoterpenes (represented by the model compound $\alpha$-pinene) are taken from Guenther et al. (1995) and amount to 144 Tg/yr. The biogenic source of methanol is calculated as in Jacob et al. (2005) and is estimated at 152 Tg/yr. Other biogenic sources are as in Müller and Stavrakou (2005).

The vegetation fire emissions are provided by the Global Fire Emission Database (GFEDv2) (van der Werf et al., 2006) and the speciation from Andreae (2007). Their global estimate amounts to 95 Tg for 2005, including 6.5 Tg ethene, 4.7 Tg glycolaldehyde, 5.9 Tg glyoxal, 1.6 Tg benzene, 0.94 Tg toluene and 0.28 Tg xylene.

Biofuel emissions are estimated using emission ratios to CO from Andreae (2007) and biofuel CO emissions from EDGARv3.3 inventory for 1997 (Olivier et al., 2001; Olivier, 2002). In addition, emissions associated with fossil fuel use and industrial activities are provided by the POET database (Olivier et al., 2003). For acetylene, biofuel emission (3.3 Tg/yr) and technological emissions (3 Tg/yr) have been adjusted in order to reproduce aircraft campaign measurements. The global anthropogenic emissions of ethene, glycolaldehyde and glyoxal are 5.4, 1.9, and 3.2 Tg, respectively. Aromatics emissions are obtained from the REanalysis of the TROpospheric chemical composition over the past 40 years (RETRO) database (Schultz et al., 2007). Their annual estimates are 3.5, 6.1 and 5.2 Tg for benzene, toluene and xylene, respectively.

### 2.3 NMVOC chemistry

The NMVOC degradation mechanism has been thoroughly presented in Stavrakou et al. (2009a). It has been designed so as to reproduce the HCHO yields of the quasi-explicit Master Chemical Mechanism (MCM) (Saunders et al., 2003) for pyrogenic and biogenic NMVOCs, and it has been further adjusted to account for the glyoxal production from the NMVOCs. Hereafter, updates of that mechanism regarding the glyoxal chemistry are briefly presented.

The degradation of isoprene is adapted from the Mainz Isoprene Mechanism (Pöschl et al., 2000) and has been further modified to include a production of two OH radicals in the reaction of first generation peroxy radicals from isoprene with HO$_2$ (Lelieveld
Further, the rate of the reaction of isoprene with OH is reduced by 30% in order to account for the isoprene/OH segregation effect, as recommended by Butler et al. (2008). Finally, the isoprene nighttime oxidation by NO$_3$ has been aligned with the MCM scheme. One-day box model simulations starting at 06:00 a.m. under high (low) NOx levels yield a molar yield of glyoxal of 9.9% (8.4%), in excellent agreement with the corresponding MCM yield of 10.1% (8.8%). For a simulation starting at 19:30 p.m., the molar glyoxal yield is estimated at 35%, very close to the MCM yield of 34%, under high or low NOx conditions. According to our chemical mechanism, reaction with OH represents 85% of the global isoprene sink, followed by ozonolysis (10%) and oxidation by NO$_3$ (5%). The calculated globally averaged molar glyoxal yield from isoprene is found to be 7.2%, by 16% higher than in Fu et al. (2008). The mean molar glyoxal yield from monoterpenes is 2.8%, 63.6% from acetylene, 5.7% from ethene, and 9.2% from glycolaldehyde. The formation yield of glyoxal from aromatics is taken to be the mean of the literature values reported in Arey et al. (2009): 33.3% for benzene, 22.3% for toluene and, taking into account for the relative contribution of o-, m-, and p-xylene in the total xylene emission (Middleton et al., 1990), the glyoxal yield from the lumped xylene compound is estimated at 17.3%. Note that laboratory studies conducted under close to atmospheric conditions suggest that the glyoxal yield from toluene is higher (30.6%, Volkamer et al., 2001) than assumed in this study.

Other identified precursors with significant glyoxal yields, like trimethylbenzenes, ethylbenzene, ethylobutenes, butenes, acrolein, etc. (Volkamer et al., 2005b) are not included in the model, either due to the lack of available emission data for these compounds on the global scale, or to their estimated minor contribution to the global annual glyoxal production. For instance, the global anthropogenic emission of trimethylbenzenes, estimated at 1 Tg/yr in the RETRO database (Schultz et al., 2007), would yield only about 0.04 Tg/yr of glyoxal, whereas the glyoxal produced from the degradation of methylbutenol is estimated at 0.35 Tg/yr (Fu et al., 2008).

The glyoxal photodissociation reactions combine absorption cross sections by Volkamer et al. (2005a) with literature quantum yields from Sander et al. (2006). The
2.4 Aerosol emissions and SOA formation

Secondary organic aerosols (SOA) are assumed to be formed rapidly and irreversibly as soon as glyoxal is processed in liquid clouds. The distribution of cloud volume fraction and liquid water content are obtained from ECMWF analyses for 2005.

The irreversible uptake of glyoxal by aqueous aerosols is assumed to proceed via a uniform uptake probability, \( \gamma = 2.9 \times 10^{-3} \) (Liggio et al., 2005). This value has been derived from laboratory experiments conducted under atmospheric conditions using ammonium sulphate seed aerosols at 55% relative humidity. Note however, that the rapidity and irreversibility of the glyoxal uptake has been questioned by Kroll et al. (2005) through experiments conducted in the absence of light and oxidants. Nevertheless, recent laboratory studies demonstrate that the uptake is greatly enhanced by the presence of oxidants and light (Volkamer et al., 2009). Another evidence for rapid glyoxal uptake is provided by direct measurements of gas-phase glyoxal (Volkamer et al., 2007), where the underestimation of the observed glyoxal concentrations can be fully accounted either through a parameterization by irreversible uptake into the aerosols with \( \gamma \) equal to 3.7 \( \times 10^{-3} \), or through reversible partitioning to aerosol liquid water with an effective Henry’s constant \( K_H \) estimated at 4 \( \times 10^9 \) M atm\(^{-1} \), which, although higher than the \( K_H \) in pure water by four orders of magnitude, is found to be consistent with recent experimental values (Ip et al., 2009). Due to the high uncertainty associated to the processes leading to the SOA formation from glyoxal, two sensitivity experiments are conducted assuming either no glyoxal uptake on aerosols, or that the uptake proceeds with the coefficient \( \gamma = 3.7 \times 10^{-3} \) (Volkamer et al., 2007), which is by 27% higher than in Liggio et al. (2005). The results of these sensitivity studies are discussed in Sects. 3 and 4.2.

Aqueous aerosols in IMAGESv2 include liquid inorganic aerosols and hydrophilic carbonaceous aerosols. Dust and sea-salt aerosols are neglected. Inorganic aerosols (the sulfate-ammonium-nitrate-water system) are calculated using EQSAM (EQuilib-
rium Simplified Aerosol Model) (Metzger et al., 2002), with deliquescence relative humidities (DRH) of Nenes et al. (1998). In order to account for the hysteresis of particle phase transitions (Wang et al., 2008), inorganic aerosols are assumed to be solid only when relative humidity never exceeded the DRH in the past 24 h. The description of sulfate formation follows Pham et al. (1995) with several improvements. Anthropogenic emissions (72.6 TgS/yr globally) are provided by EDGAR-FT2000 (Olivier et al., 2005). Non-eruptive volcanic emissions (7.2 TgS/yr, including 2.4 TgS/yr of non-SO$_2$ sulfur compounds emitted as SO$_2$ in the model) are provided by Andres and Kasgnoc (1998). The GFEDv2 database and the emission ratios of Andreae (2007) result in global emissions of 1.2 TgS/yr due to vegetation fires in 2005. Natural sulfur emissions include the oceanic DMS source obtained from Dentener et al. (2006) (18.3 TgS/yr) and a small biospheric emission of H$_2$S (0.5 TgS/yr) (Pham et al., 1995). The model includes the gas-phase oxidation of DMS and H$_2$S to form SO$_2$ (and methanesulfonic acid in the case of DMS) and the SO$_2$ oxidation by OH to form sulfate, with rates from Sander et al. (2006). The pH-dependent aqueous-phase oxidation of SO$_2$ by O$_3$ and H$_2$O$_2$ to form sulfate is included using kinetic data and Henry constants reported in Tie et al. (2001), with cloud volume fractions and liquid water content from ECMWF analyses.

Ammonia emissions due to domestic animals and synthetic fertilizers (19 TgN/yr globally) are obtained from Bouwman et al. (2002). Vegetation fires emitted 6.1 TgN H$_3$N in 2005, based on GFEDv2 and emission ratios of Andreae (2007). The other NH$_3$ emissions (totalling 19.4 Tg/yr) are provided by the GEIA inventory (Bouwman et al., 1997). The temporal variability of emissions due to soils and manure is driven by soil temperature (Park et al., 2004), whereas the seasonal variation of emissions due to crops and fertilizers follows the number of daylight hours (Adams et al., 1999).

The emissions of organic carbon (OC) and black carbon (BC) aerosols from anthropogenic activities (Bond et al., 2004) and biomass burning (van der Werf et al., 2006) are included in the model. Both OC and BC have hydrophilic and hydrophobic components. The conversion of hydrophobic to hydrophilic OC and BC due to aging is represented by an e-folding time of 6 h in tropical regions (Bellouin et al., 2007) and
one day elsewhere (Martin et al., 2003). The mass ratio of organic matter to OC is taken to 2.1, a value valid for rural areas (Pang et al., 2006). The effective diameter of dry organic aerosols is taken to 145 nm, based on field campaigns over Amazonia (Rissler et al., 2006) and Africa (Bellouin et al., 2007). The hygroscopic growth factor is calculated by setting the activity coefficient of water to 1.1, a value typical for α-pinene SOA (Compernolle et al., 2009) leading to growth factors in good agreement with field measurements over Amazonia during the dry season (Rissler et al., 2006; Mircea et al., 2005).

### 2.5 Observed CHOCHO and HCHO columns

We use glyoxal and formaldehyde column abundancies retrieved from the SCIAMACHY (Scanning imaging absorption spectrometer for atmospheric cartography) UV-visible sounder onboard ENVISAT satellite for 2005. The differential optical absorption spectroscopy (DOAS) technique (Platt et al., 1994) has been applied for both compounds. The spectral window found to be optimal for the fitting is 436–457 nm for CHOCHO (Wittrock et al., 2006) and 328.5–346 nm for HCHO (De Smedt et al., 2008). For CHOCHO, the Volkamer et al. (2005a) absorption cross sections are used, and for the HCHO retrieval the Cantrell et al. (1990) values shifted by 0.08 mm (Stavrakou et al., 2009b).

The columns are binned onto the model resolution and monthly averaged. The error on the CHOCHO columns is assumed to comprise an absolute error of $2 \times 10^{14}$ molecules cm$^{-2}$ and a 30% relative error, whereas the error on the HCHO columns includes the reported retrieval error (De Smedt et al., 2008) and a model error amounting to 20% of the column, based on sensitivity calculations with the IMAGESv2 model (Stavrakou et al., 2009a). Seasonally averaged observed glyoxal and formaldehyde columns are displayed in Fig. 1. A thorough discussion on the seasonal variability of the measured columns can be found in De Smedt et al. (2008), Wittrock et al. (2006), and Vrekoussis et al. (2009). As this study focuses on land sources, oceanic data of both compounds are not used to constrain the inversion. Moreover, data in the vicin-
ity of the south Atlantic anomaly (26.1° S, width=16.7°, 47.4° W, width=28.4°, Nichitiu et al., 2004) are also excluded.

3 The glyoxal global budget

The global a priori budget of glyoxal (in absence if any additional biogenic source) is summarized in Table 1. About 55%, 27% and 18% of the global glyoxal source are due to biogenic, anthropogenic, and pyrogenic emissions, respectively. The largest sink of glyoxal is photolysis (51%), followed by SOA formation (22%), OH oxidation (18%), and dry and wet deposition (8%). The global glyoxal source is estimated to 56 Tg/yr and the global lifetime to 2.5 h, in good agreement with literature values (45 Tg/yr, 2.9 h in Fu et al., 2008, 56 Tg/yr, 3 h in Myriokefalitakis et al., 2008). The global annual burden, estimated at 15.8 Gg/yr, is consistent with the Fu et al. (2008) estimate of 15 Gg/yr, and is found to be by 20% lower than in Myriokefalitakis et al. (2008). This discrepancy is mostly due to the SOA formation sinks which are not considered in Myriokefalitakis et al. (2008).

Directly emitted glyoxal from fires constitutes the major contribution to the total pyrogenic glyoxal production (60%), whereas the remainder is secondary production, mostly from acetylene oxidation (24%), aromatics (5%), glycolaldehyde (4%), and ethylene (7%). About 60% of the total anthropogenic glyoxal is generated through acetylene oxidation, followed by primary glyoxal (20%), and aromatics (15%). Biogenically emitted glyoxal is only of secondary origin, the isoprene oxidation being responsible for more than 90% of the total source, the remainder being formed from monoterpenes and ethene.

The estimated glyoxal sink associated with SOA formation in clouds (4.7 Tg/yr) is found to be in good agreement with the previous estimate of Fu et al. (2008), whereas the irreversible uptake on aerosols, amounting to 5.9 Tg/yr on OC/BC and to 1.8 Tg/yr on sulfates, is significantly higher than in Fu et al. (2008). The resulting mean lifetime against SOA conversion (ca. 11 h), is about two times shorter than in Fu et al. (2008).
The lifetimes against photolysis and OH oxidation are calculated to be 4.8 h and 14 h, respectively, in good agreement with the study of Myriokefalitakis et al. (2008). Note that the glyoxal aerosol sink is even higher when using the uptake coefficient taken from Volkamer et al. (2007) ($3.7 \times 10^{-3}$) and is estimated at 9.2 Tg/yr.

Annually averaged observed and simulated CHOCHO columns are compared in the top panels of Fig. 2. The model is found to underestimate the observations over all continents. The model/data biases and spatiotemporal correlations averaged over large regions are summarized in Table 2. On the global scale, the a priori simulation underestimates the data by 57%, and yields an overall correlation coefficient of 0.5. The correlation is stronger over Northern America and Asia (>0.6), but falls to zero over Europe. The model/data discrepancies over the oceans are much more important (Fig. 2), but their investigation lies out of scope of this study.

The global HCHO budget is estimated at 1660 Tg/yr. According to the model calculations, the global HCHO production is due to the oxidation of methane (60%), isoprene (30%), anthropogenic (7%), and pyrogenic NMVOCs (3%), whereas photolysis is the major sink (73%), followed by the OH oxidation (23%), and deposition (4%) (Stavrakou et al., 2009a). Systematic comparisons between the monthly mean observed and simulated HCHO columns over a large number of regions, presented in Stavrakou et al. (2009a), show that the model captures the spatial pattern of the observed HCHO columns and succeeds quite well in reproducing the intensity of the observed maxima.

Monthly mean HCHO columns from SCIAMACHY are compared to modelled columns (Fig. 3, right) averaged over several tropical regions. The seasonality and intensity of the observed HCHO columns are well reproduced by the model, except for an overestimation over Indonesia and an underestimation over southern Africa. The comparisons for glyoxal on the same figure confirm the model underestimation noted above. Furthermore, in many cases, the seasonality of the observed columns is severely misrepresented.
4 Inverse modelling

4.1 Setup for the inversion

The inversion is performed on a grid cell basis, i.e. the fluxes are updated at every grid cell and month (Stavrakou and Müller, 2006; Stavrakou et al., 2009b), except when the a priori emissions are very small (see below). Four emission categories are optimized: i) anthropogenic NMVOC emissions of glyoxal and its precursors (except for acetylene) ii) all pyrogenic NMVOCs, as in Stavrakou et al., 2009b), iii) identified biogenic NMVOC precursors (BVOC) (isoprene, monoterpenes, methanol, and ethene), and iv) unidentified biogenic VOCs (UVOC). The global a priori UVOC source is distributed as the leaf area index obtained from the Moderate Resolution Imaging Radiospectrometer (MODIS) satellite (http://modis.gsfc.nasa.gov/data/atbd/) and is scaled globally to 20 Tg/yr (Fig. 4, upper panel).

Emissions are not optimized in grid cells with very low emissions. An emission parameter is optimized in a grid cell only when the a priori maximum flux in this cell is higher than $10^{10}$ molecules cm$^{-2}$ s$^{-1}$ for the first three categories and $5 \times 10^{8}$ molecules cm$^{-2}$ s$^{-1}$ for the latter. The thresholds has been adjusted so as to avoid the existence of large disparities in the number of unknowns to be optimized from the different categories. The total number of control parameters is equal to ca. 30,000. The errors on the emission parameters are assumed to be factor of two. Correlations between the a priori errors on the control parameters are introduced through the off-diagonal elements of the error covariance matrix, in the same way as in Stavrakou et al. (2009b) for the inversion using HCHO data.

Among the above-mentioned species, isoprene, monoterpenes, glycolaldehyde, and ethene are precursors of both glyoxal and formaldehyde. Although the chemical feedbacks between glyoxal and formaldehyde are weak, performing a simultaneous inversion of both species constitutes the best way to take advantage from the constraints provided from both satellite datasets, while accounting for all chemical interdependencies.
Two inversion experiments are conducted. In inversion A, UVOC is directly emitted glyoxal, whereas in inversion B, UVOC forms glyoxal upon oxidation by OH with a molar yield of unity and a reaction rate of $2.86 \times 10^{-12}$ molecules $^{-1}$ cm$^3$ s$^{-1}$, which is adjusted so that the global lifetime of UVOC is equal to five days. Sensitivity calculations with shorter and longer lifetimes for UVOC will also be addressed.

4.2 Results

Inferred annual mean glyoxal columns from both inversions are illustrated in Fig. 2. In both cases the inversion reproduces reasonably well the observations, although an underestimation persists due to the a priori constraints on the emissions. Inversion B, however, provides a better match with the observed patterns and magnitudes over the continents, as well as near the coasts over the tropical oceans, due to the secondary glyoxal production from UVOC inside the transported air masses. The columns inferred from inversion A (Fig. 3, green) remain too low over southeast Asia and miss the observed seasonality, whereas the deduced columns from inversion B allow for an excellent model/data agreement over all displayed regions. Overall, the globally averaged relative bias over land is decreased from $-57\%$ in the a priori, to $-27\%$ and $-19\%$ in scenarios A and B, respectively, whereas the spatiotemporal correlation coefficient increases from 0.5 in the a priori to 0.73 and 0.77 in cases A and B. The lowest biases ($< -10\%$) and highest correlations ($>0.8$) are obtained over Tropical America in both inversions, and over Asia in inversion B.

The a priori and optimized annual global distribution of the UVOC source is displayed in Fig. 4. The a priori flux and the ratios of the optimized by the a priori fluxes for the different emission categories are illustrated in Figs. 5 and 6. The column enhancement of inversion A (Fig. 2) is achieved by strengthening the primary UVOC source to 36 Tg/yr, and by moderately increasing anthropogenic emissions as well as biogenic isoprene fluxes (Table 1). The optimized isoprene source amounts to 476 Tg/yr, in excellent agreement with a previous inversion constrained by HCHO columns (Stavrakou et al., 2009b). A similar increase in the isoprene source is deduced from the experiment B,
whereas the secondary UVOC source is estimated at 56 Tg/yr. The optimized tropical UVOC source (40 Tg/yr) represents two thirds of the total UVOC source and is strongly enhanced with respect to its a priori value.

A decrease by about 20% is found for the pyrogenically emitted NMVOCs in both inversions. The inferred global terrestrial source amounts to 94 Tg/yr in inversion A and to 108 Tg/yr in inversion B. At the continental scale, the derived mean annual isoprene flux decrease by 25% over Northern America and the strong increase over southern Africa are similar to the inversion results using only HCHO columns (Stavrakou et al., 2009b). The a posteriori anthropogenic emissions are enhanced compared to the a priori, especially over the Far East, where the increase exceeds a factor of two in inversion A. Further, the pyrogenic emissions inferred over northern Africa are decreased by 50%, in qualitative agreement with previous studies (Myriokefalitakis et al., 2008).

The uncertainties associated to the aforementioned glyoxal source estimates are large, but difficult to quantify. They are mostly due to possible biases in the satellite retrievals, to uncertainties in the chemical mechanisms of the NMVOCs, and to the representation of the glyoxal sinks in the model.

Comparisons with surface glyoxal concentration measurements at continental mid-latitude sites of the Northern Hemisphere are provided in Table 3 and Fig. 7. Whereas a close-to-zero average bias is calculated in the a priori simulation (without UVOC source), inversion A leads to a strong overestimation (+34 ppt), which would have been even larger if the settings of the inversion (e.g. a priori errors on the emission parameters) were chosen so as to allow for a better match with the satellite data. The computed bias after optimization B is very close to the a priori, whereas its slightly negative value is explained by the inferred decrease of the isoprene emissions over Northern America, as illustrated on Fig. 6. Further, the root mean square deviation in inversion B is by 16% lower than in the a priori. These comparisons suggest that the space-based glyoxal columns are compatible with the available in situ measurements when the extra glyoxal source is assumed to be secondary. However, in situ glyoxal measurements in tropical continental regions, certainly critical in order to evaluate the
The sensitivity tests C and D use the setup of inversion B. Aerosol uptake of glyoxal is ignored in inversion C and a higher uptake probability is assumed in inversion D (see Sect. 2.4). The results show that our conclusions regarding the magnitude of the UVOC source and the agreement with the independent measurements of Table 3 do not depend significantly on the assumptions made for the uptake of glyoxal by aerosols. More precisely, when ignoring the SOA formation sink on aerosols, the inferred global glyoxal and UVOC flux amount to 95 Tg/yr and 45 Tg/yr, respectively, and the global glyoxal lifetime is by 16% longer (2.9 h) compared to the lifetime of inversion B. The agreement with the ground-based measurements is similar to that of inversion B, except for the mean average bias (Table 3) which becomes positive in this inversion (1.1 pptv) due to the longer residence time of glyoxal near the surface. When the higher value $\gamma=3.7 \cdot 10^{-3}$ for the uptake coefficient is used, the estimated global annual UVOC and glyoxal source, and the glyoxal lifetime are quite similar to those of inversion B, namely, 57 Tg, 110 Tg and 2.6 h, respectively, whereas the predicted SOA sink on aerosols amounts to 13.2 Tg/yr. In this case, the average bias with the observations of Table 3 is calculated to be −2 pptv.

However, the results show a stronger sensitivity to the choice of the UVOC lifetime. Sensitivity tests where the UVOC lifetime ranges between one (inversion E) and 20 days (inversion F) give results qualitatively similar to inversion B, with inferred global terrestrial glyoxal source estimated between 96–138 Tg/yr, and UVOC flux between 42 and 84 Tg/yr. These inversions underestimate the satellite data by 13–21%, whereas the agreement with in situ measurements worsens compared to the case B results. Inversions with UVOC lifetimes shorter than one day give results more similar to the inversion A, and lead to a significant overestimation of the surface glyoxal mixing ratio. The main results of the performed tests are summarized in Fig. 8.

Both optimizations improve significantly the modelled HCHO column, bringing it much closer to the observations (Fig. 3). This is achieved mostly by decreasing the Indonesian biogenic fluxes by up to 40% and strongly increasing the isoprene emis-
sions over southern Africa, whereas less significant changes are inferred over other regions (Stavrakou et al., 2009b).

5 Conclusions

The continental source of glyoxal has been constrained by glyoxal and formaldehyde columns retrieved from the SCIAMACHY sensor for 2005, based on a two-compound grid-based inversion scheme using the IMAGESv2 CTM and its adjoint. Formaldehyde provides a useful constraint on the emissions of isoprene, an important precursor of both glyoxal and formaldehyde. Fluxes of anthropogenic, pyrogenic, identified biogenic NMVOCs, and uncharacterized organic compounds (UVOC), initially distributed as the MODIS leaf area index, are determined by the inversion. Two possibilities have been explored. The UVOC is either direct glyoxal flux (case A) or an unknown glyoxal precursor with a lifetime of five days (case B).

The continental glyoxal source (56 Tg/yr in the a priori), is estimated at 94 Tg/yr in inversion A, including a primary UVOC source of 36 Tg/yr, and at 108 Tg/yr in inversion B, including a secondary UVOC source of 54 Tg/yr. The isoprene source, major contributor to the global glyoxal source, exhibits only a slight increase with respect to its priori value on the global scale, but large deviations at the continental scale, in agreement with an inversion study constrained by formaldehyde columns only (Stavrakou et al., 2009b). Global anthropogenic fluxes are enhanced over the Far East, especially in inversion A, but stay close to their a priori values over other regions. Pyrogenic emissions are reduced by 25% on the global scale, the decrease being more important over the African continent. The emission updates enable significant reduction of the model/data discrepancies for the glyoxal and formaldehyde, demonstrating the capability to reconcile the two datasets in a single scheme. The global mean model/data bias, −57% in the a priori simulation, is reduced to −27% and to −19% in inversions A and B, respectively, whereas the globally calculated spatiotemporal correlation coefficient increases from 0.5 to 0.77 in inversion B.
Validation using in situ glyoxal measurements in rural settings suggests that the missing glyoxal source is secondary (inversion B), since a large additional direct emission of glyoxal (inversion A) leads to a large overestimation of surface glyoxal concentrations. However, the lack of in situ measurements in the continental tropics makes difficult to evaluate the model results over these strongly emitting regions. Note also that errors in the satellite retrievals, the chemical degradation mechanisms, and the a priori spatiotemporal flux distribution may introduce significant uncertainties in the estimated sources. Further, the inferred glyoxal source depends on the choice of the UVOC lifetime and could range between 96 and 138 Tg/yr, according to sensitivity calculations, whereas it is found to be weakly sensitive to the choice of the uptake coefficient in the parameterization of SOA formation sink from aerosols. This sink, estimated between 10 Tg/yr and 13 Tg/yr, is found to be much higher than in a previous model study (Fu et al., 2008). The large relevance of SOA formed from aqueous aerosols has been also pointed out in Volkamer et al. (2009). More work is definitely needed in order to improve the representation of these processes in large scale models.

More glyoxal measurements, both space-based and in situ, would be a valuable contribution to our understanding of the emission mechanisms, the precursors of this compound, as well as the chemical mechanism controlling the local glyoxal concentrations.

Acknowledgements. This work has been supported by the PRODEX programme of the ESA funded by the Belgian Science Policy, and has been in part facilitated by funding by the University of Bremen, the State of Bremen, the German Aerospace, DLR, and EU through the ACCENT Network. We thank the DLR and ESA for the provision of the level 0 data from SCIAMACHY, used in this study. M. Vrekoussis acknowledges the A. v. Humboldt foundation and the European Union (Marie Curie) for consecutive research fellowships.
References


ACPD
9, 13593–13628, 2009

Estimation of the glyoxal land source
T. Stavrakou et al.


Estimation of the glyoxal land source

T. Stavrakou et al.


Estimation of the glyoxal land source
T. Stavrakou et al.


13616
Table 1. Global annual glyoxal source calculated with IMAGESv2 for 2005 in the a priori simulation and in inversions A and B. A priori and optimized flux estimates are expressed in Tg/yr.

<table>
<thead>
<tr>
<th>Sources</th>
<th>A priori</th>
<th>Inversion A</th>
<th>Inversion B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthopogenic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>primary</td>
<td>3.2</td>
<td>5</td>
<td>3.4</td>
</tr>
<tr>
<td>secondary</td>
<td>11.9</td>
<td>13.3</td>
<td>12.1</td>
</tr>
<tr>
<td>Pyrogenic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>primary</td>
<td>5.9</td>
<td>4.7</td>
<td>4.3</td>
</tr>
<tr>
<td>secondary</td>
<td>3.9</td>
<td>2.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Biogenic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>32.5</td>
<td>32</td>
</tr>
<tr>
<td>UVOC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>94</td>
<td>108</td>
</tr>
<tr>
<td>Sinks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photolysis</td>
<td>28.7</td>
<td>41.9</td>
<td>62.4</td>
</tr>
<tr>
<td>OH oxidation</td>
<td>10</td>
<td>14.1</td>
<td>18</td>
</tr>
<tr>
<td>SOA in clouds</td>
<td>4.7</td>
<td>9.3</td>
<td>10</td>
</tr>
<tr>
<td>SOA in aerosols</td>
<td>7.7</td>
<td>11.6</td>
<td>10.4</td>
</tr>
<tr>
<td>Dry deposition</td>
<td>2.7</td>
<td>12.1</td>
<td>3.9</td>
</tr>
<tr>
<td>Wet deposition</td>
<td>1.8</td>
<td>4.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Lifetime (hr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>2.4</td>
<td>2.5</td>
</tr>
<tr>
<td>Burden (Gg)</td>
<td>15.8</td>
<td>25</td>
<td>32.3</td>
</tr>
</tbody>
</table>
### Table 2. Mean model/data biases and spatiotemporal correlation coefficients per region calculated for the a priori simulation, and after the inversions A and B.

<table>
<thead>
<tr>
<th>Regions</th>
<th>A priori</th>
<th>Inversion A</th>
<th>Inversion B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>r</td>
<td>Bias</td>
</tr>
<tr>
<td>N. America (30–70 N, 60–170 W)</td>
<td>−48.6</td>
<td>0.66</td>
<td>−21.4</td>
</tr>
<tr>
<td>Tropical America (30 S–20 N, 30–80 W)</td>
<td>−63.7</td>
<td>0.43</td>
<td>−3.2</td>
</tr>
<tr>
<td>Europe (35–65 N, 10 W–55 E)</td>
<td>−62.8</td>
<td>0.00</td>
<td>−32.2</td>
</tr>
<tr>
<td>Africa (40 S–30 N, 20 W–45 E)</td>
<td>−54.2</td>
<td>0.47</td>
<td>−31.0</td>
</tr>
<tr>
<td>Asia (10 S–55 N, 60–160 E)</td>
<td>−54.8</td>
<td>0.62</td>
<td>−23.8</td>
</tr>
<tr>
<td>Oceania (10–40 S, 110–170 E)</td>
<td>−17.3</td>
<td>0.57</td>
<td>−6.3</td>
</tr>
<tr>
<td>Globe</td>
<td>−56.9</td>
<td>0.50</td>
<td>−26.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location</th>
<th>Time</th>
<th>Observ.</th>
<th>A priori</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgia (32.5N, 82W) [1]</td>
<td>Jul–Aug 1991</td>
<td>18.0</td>
<td>65.1</td>
<td>132.6</td>
<td>55.4</td>
</tr>
<tr>
<td>Georgia (32.5N, 82W) [1]</td>
<td>Jun 1992</td>
<td>83.0</td>
<td>48.8</td>
<td>132.6</td>
<td>45.9</td>
</tr>
<tr>
<td>Nashville (37N, 87W) [2]</td>
<td>Jul 1995</td>
<td>67.0</td>
<td>89.0</td>
<td>64.7</td>
<td>58.5</td>
</tr>
<tr>
<td>Pinnacles (39N, 78W) [3]</td>
<td>Sept 1990</td>
<td>44.0</td>
<td>48.9</td>
<td>138.1</td>
<td>51.0</td>
</tr>
<tr>
<td>Blodgett Forest (39N, 120W) [4]</td>
<td>Aug–Sep 2000</td>
<td>27.0</td>
<td>25.4</td>
<td>59.4</td>
<td>31.1</td>
</tr>
<tr>
<td>Tomakomai (42.7N, 141.6E) [5]</td>
<td>Sep 2003</td>
<td>26.0</td>
<td>7.4</td>
<td>45.3</td>
<td>19.3</td>
</tr>
<tr>
<td>Uryu, Japan (44.3N, 142.2E) [6]</td>
<td>Aug 2002</td>
<td>18.0</td>
<td>13.2</td>
<td>70.8</td>
<td>27.2</td>
</tr>
<tr>
<td>Pabsthum (52.5N, 12.5E) [7]</td>
<td>Aug 1998</td>
<td>38.0</td>
<td>18.8</td>
<td>67.6</td>
<td>26.3</td>
</tr>
</tbody>
</table>

Average bias (pptv) 0.14 34.3 -1.3
RMSD 6.2 15.7 5.2
**Fig. 1.** Averages of the observed CHOCHO (left, in $10^{14}$ molecules cm$^{-2}$) and HCHO (right, in $10^{15}$ molecules cm$^{-2}$) vertical columns for winter (DJF), spring (MAM), summer (JJA) and autumn (SON) 2005. Blank regions correspond to an absence of valid data due to large zenith angles, south Atlantic Anomaly, or missing cloud information.
Fig. 2. Annually averaged CHOCHO vertical columns observed (top left), simulated with IMAGESv2 (top right), and deduced from inversion (bottom panels). The global annual CHOCHO source for each case is given inset. Units are $10^{14}$ molecules cm$^{-2}$. 
Fig. 3. Monthly mean CHOCHO (left) and HCHO (right) column abundancies simulated with IMAGESv2 and observed by SCIAMACHY (black dots) in 2005. Black solid lines correspond to the a priori simulation, green and red to inversions A and B, respectively. CHOCHO columns are expressed in $10^{14}$ molecules cm$^{-2}$ and HCHO columns in $10^{15}$ molecules cm$^{-2}$.
Fig. 4. Global annual distribution of the additional glyoxal precursor (UVOC) source inferred by the inversions. Units are 10^{10} molecules cm^{-2} s^{-1}.
Fig. 5. Global annual a priori anthropogenic flux of glyoxal precursors, and the ratio of the optimized by the a priori estimate inferred by the inversion studies.
Fig. 6. Global annual a priori pyrogenic NMVOC and biogenic isoprene emissions (upper panel) and the ratios of the optimized by the a priori fluxes deduced by the inversion B (lower panel).
Fig. 7. In situ measurements of glyoxal vs. a priori (in black) and optimized (in green for inversion A and red for inversion B) glyoxal concentrations (see Table 3).
Fig. 8. Global glyoxal source and UVOC source inferred from the performed sensitivity inversions: B (standard with UVOC lifetime equal to five days and uptake coefficient equal to $2.9 \cdot 10^{-3}$), C (as B, but assuming no aerosol uptake of glyoxal), D (as B, but with uptake coefficient equal to $3.7 \cdot 10^{-3}$), E (as B, with UVOC lifetime of one day), F (as B, with UVOC lifetime of 20 days).