Bias correction of OMI HCHO columns based on FTIR and aircraft measurements and impact on top-down emission estimates

Jean-François Müller[1], Trissevgeni Stavrakou[1], Glenn-Michael Oomen[1], Beata Opacka[1], Isabelle De Smedt[1], Alex Guenther[2], Corinne Vigouroux[1], Bavo Langerock[1], Carlos Augusto Bauer Aquino[3], Michel Grutter[4], James Hannigan[5], Frank Hase[6], Rigel Kivi[7], Erik Lutsch[8], Emmanuel Mahieu[9], Maria Makarova[10], Jean-Marc Metzger[11], Isamu Morino[12], Isao Murata[13], Tomoo Nagahama[14], Justus Notholt[15], Ivan Ortega[5], Mathias Palm[15], Amelie Röhling[6], Wolfgang Stremme[4], Kimberly Strong[8], Ralf Sussmann[16], Yao Té[17], and Alan Fried[18]

[1]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[2]University of California Irvine, Irvine, CA, USA
[3]Instituto Federal de Educaçao, Ciência e Tecnologia de Rondônia (IFRO), Porto Velho, Brazil
[4]Instituto de Ciencias de la Atmósfera y Cambio Climático, Universidad Nacional Autónoma de México (UNAM), Mexico City, México
[5]Atmospheric Chemistry, Observations & Modeling, National Center for Atmospheric Research (NCAR), Boulder, CO, USA
[6]Karlsruhe Institute of Technology (KIT), Institute for Meteorology and Climate Research (IMK-ASF), Karlsruhe, Germany
[7]Finnish Meteorological Institute (FMI), Sodankylä, Finland
[8]Department of Physics, University of Toronto, Toronto, Canada
[9]Institut d'Astrophysique et de Géophysique, Université de Liège, Liège, Belgium
[10]Saint Petersburg State University, Atmospheric Physics Department, St. Petersburg, Russia
[11]Observatoire des Sciences de l'Univers Réunion (OSU-R), UMS 3365, Université de la Réunion, Saint-Denis, France
[12]National Institute for Environmental Studies (NIES), Tsukuba, Ibaraki 305-8506, Japan
[13]Graduate School of Environment Studies, Tohoku University, Sendai 980-8578, Japan
[14]Institute for Space-Earth Environmental Research (ISEE), Nagoya University, Nagoya, Japan
[15]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[16]Karlsruhe Institute of Technology, IMK-IFU, Garmisch-Partenkirchen, Germany
[17]LERMA-IPSL, Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, 75005 Paris, France
[18]Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA

Correspondence: Jean-François Müller (jfm@aeronomie.be)

Abstract.

Spaceborne formaldehyde (HCHO) measurements constitute an excellent proxy for the sources of non-methane volatile organic compounds (NMVOCs). Past studies suggested substantial overestimations of NMVOC emission in state-of-the-art inventories over major source regions. Here, the QA4ECV (Quality Assurance for Essential Climate Variables) retrieval of HCHO columns from OMI (Ozone Monitoring Instrument) are evaluated against (1) FTIR (Fourier-transform infrared) column observations at 26 stations worldwide, and (2) aircraft in situ HCHO concentration measurements from campaigns conducted over the U.S. in 2012–2013. Both validation exercises show that OMI underestimates high columns and overestimates low columns. The linear regression of OMI and aircraft-based columns gives $\Omega_{\mathrm{OMI}} = 0.651\,\Omega_{\mathrm{airc}} + 2.95 \cdot 10^{15}$ molec.cm^{-2}, with Ω_{OMI} and Ω_{airc} the OMI and aircraft-derived vertical columns, whereas the regression of OMI and FTIR data gives $\Omega_{\mathrm{OMI}} = 0.659\,\Omega_{\mathrm{FTIR}} + 2.02 \cdot 10^{15}$ molec.cm^{-2}. Inverse modelling of NMVOC emissions with a global model based on OMI

columns corrected for biases based on those relationships leads to much-improved agreement against FTIR data and HCHO concentrations from 11 aircraft campaigns. The optimized global isoprene emissions ($\sim 445\,\mathrm{Tg\,yr^{-1}}$) are 25% higher than those obtained without bias correction. The optimized isoprene emissions bear both striking similarities and differences with recently published emissions based on spaceborne isoprene columns from the CrIS (Cross-track Infrared Sounder) sensor. Although the interannual variability of OMI HCHO columns is well understood over regions where biogenic emissions are dominant, and the HCHO trends over China and India clearly reflect anthropogenic emission changes, the observed HCHO decline over the Southeastern U.S. remains imperfectly elucidated.

1 Introduction

The atmospheric oxidation of non-methane volatile organic compounds exerts multiple influences on tropospheric ozone (e.g. Chameides et al., 1988; Archibald et al., 2020), fine particulate matter (e.g. Spracklen et al., 2011; Miao et al., 2021) and the oxidizing capacity of the atmosphere (e.g. Seinfeld and Pandis, 1988; Martinez et al., 2003; Bates and Jacob, 2019). Quantifying those impacts is made difficult by the large number and diversity of NMVOCs, with chemical lifetimes ranging from a few minutes to several months (Shu and Atkinson, 1994; Atkinson, 2000) and by important gaps in our understanding of their emissions and chemical degradation mechanisms. Even biogenic isoprene, the most abundantly emitted NMVOC at the global scale, has very uncertain emissions due (among others) to the strong variability of emission rates among different plant species (Guenther et al., 2006) and to the large uncertainties in the climate and vegetation maps used to calculate the fluxes in state-of-the-art emission models (Arneth et al., 2011; Sindelarova et al., 2014). In addition, field campaigns in various environments such as cities (e.g. Karl et al., 2018) and remote areas (e.g. Read et al., 2012; Wang et al., 2012; Lawson et al., 2015; Travis et al., 2020) suggest that current inventories of anthropogenic and natural emissions of many NMVOCs, and particularly oxygenated volatile organic compounds (OVOCs) are incomplete. The comparison of measured total OH reactivity with the sum of contributions from measured individual compounds over forests has revealed the presence of a "missing OH reactivity" that is partly, but not always completely, explained by unobserved oxidation products of known VOC precursors (Di Carlo et al., 2004; Sinha et al., 2010; Yang et al., 2016; Nölscher et al., 2016; Sanchez et al., 2021), thereby providing additional indication that primary emissions are missed in emission inventories. For biomass burning as well, the contribution of unidentified organic compounds to the total NMVOC fluxes might be of the same order, or even larger, than the explicitly identified compounds (Akagi et al., 2011; Andreae, 2019).

Since chemistry transport models do not include unidentified NMVOCs, they are expected to underestimate the total NMVOC flux. Nevertheless, chemistry-transport models (CTMs) using state-of-the-art emission inventories (Guenther et al., 1995, 2006) were found to overestimate the column abundances of formaldehyde, a major NMVOC oxidation product, against retrieved columns from spaceborne UV-visible sounders such as the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography/Chemistry (SCIAMACHY, Stavrakou et al., 2009; Barkley et al., 2011) and the Ozone Monitoring Instrument (OMI, Millet et al., 2008; Barkley et al., 2011; Fortems-Cheiney et al., 2012; Stavrakou et al., 2015; Bauwens et al., 2016). Furthermore, the largest model biases (exceeding -40%) were found over high-emission areas such as Amazonia,

the Central African rainforest, and the Southeastern U.S. (Barkley et al., 2011; Marais et al., 2014; Bauwens et al., 2016),

45 where isoprene oxidation is believed to be by far the largest source of formaldehyde (Palmer et al., 2003; Stavrakou et al., 2009). Note that earlier studies using HCHO retrievals from the pioneering Global Ozone Monitoring Experiment (GOME, Palmer et al., 2003; Abbot et al., 2003; Shim et al., 2005; Palmer et al., 2006; Fu et al., 2007; Stavrakou et al., 2009) led to mixed conclusions with both model underestimations and overestimations, primarily due to large differences between different retrievals (De Smedt et al., 2008; Stavrakou et al., 2009). We restrict the following discussion to studies based on subsequent,

50 higher-resolution sounders, in particular OMI.

The general model overestimation of spaceborne HCHO abundances over source regions was mainly attributed to an over-estimation of isoprene emissions estimated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN, Guenther et al., 2006, 2012), and inverse modelling of emissions – i.e. the derivation of improved emissions constrained by observational data while accounting for the error covariances of the data and the a priori emissions – has led to a reduction

55 of biogenic emissions globally (e.g., Bauwens et al., 2016). The overestimation of isoprene emissions was also concluded in several studies based on model comparisons with isoprene concentration measurements, in particular over tropical ecosystems (e.g., Houweling et al., 1998; Bey et al., 2001). However, those comparisons are uncertain, due (among others) to the strong variability of isoprene concentrations and to their dependence on OH radical levels, likely too low in those models due to the neglect of OH-recycling mechanisms which were not yet identified (Lelieveld et al., 2008; Paulot et al., 2009; Peeters

60 et al., 2014; Wennberg et al., 2018; Novelli et al., 2020). A more recent evaluation indicates a model underestimation against aircraft measurements of many reactive VOCs, including formaldehyde and isoprene, over North America (Chen et al., 2019). In addition, recent isoprene flux measurements by eddy covariance from airborne platforms did not suggest overestimations of MEGAN-calculated emissions over the Amazon forest (Gu et al., 2017), the Southeastern U.S. (Yu et al., 2017), and California (Misztal et al., 2014).

65 Whereas validation studies of the Harvard GOME retrieval showed a good consistency of spaceborne columns against aircraft in situ observations (Martin et al., 2004; Millet et al., 2006), evaluation of OMI HCHO columns against three aircraft campaigns (Boeke et al., 2011) also indicated a good agreement (-3% bias) with respect to the mean aircraft-based columns, but displayed a larger relative bias (-17%) for higher columns ($> 5 \times 10^{15}\,\mathrm{molec.\,cm}^{-2}$). The latter was tentatively attributed to the preferential sampling of polluted plumes by the aircraft, since the highest HCHO columns were measured in the vicinity

70 of Mexico City and an airshed near Houston, Texas. However, similar, or even larger, biases were found in the comparison of HCHO columns from several spaceborne sensors, including OMI, against in situ measurements of the SEAC^{4}RS campaign conducted over a wide area covering much of the Southeastern U.S. during summer, a region with high HCHO abundances of primarily biogenic origin (Zhu et al., 2016). The lowest bias (-20%) was realized by the OMI BIRA-V14 retrieval (De Smedt et al., 2015) and could be further reduced (to -12%) when adopting the aircraft vertical shape profiles in the calculation of air

75 mass factors of the OMI retrieval.

The first evaluations of HCHO satellite columns using Fourier-transform infrared (FTIR) measurements were conducted in remote environments (Jones et al., 2009; Vigouroux et al., 2009) and suggested a fairly good agreement. In contrast with this, and in qualitative agreement with the results of Boeke et al. (2011), the recent evaluation of TROPOMI (TROPOspheric Moni-

toring Instrument) HCHO columns against FTIR measurements from a network of 25 stations worldwide showed a pronounced

80 dependence of the bias of TROPOMI columns with the column magnitude, with overestimations (averaging $+25\%$) found for very low columns ($< 2.5 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$) and underestimations ($-30.8\%$) for high columns ($> 8 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$). The largest underestimation (-36%) was found for the station with the highest average HCHO column, Porto Velho in Amazonia ($29 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$). Comparable biases can be expected for OMI as for TROPOMI, given the similarity between the two instruments (De Smedt et al., 2021).

85 Acknowledging the potentially important consequences of such biases on top-down VOC emission estimates based on spaceborne HCHO columns, we aim (1) to assess the biases of HCHO columns from a recent OMI retrieval (De Smedt et al., 2018) using both FTIR data (using a methodology similar to Vigouroux et al. (2020)) and in situ aircraft measurements from several campaigns in the U.S., and (2) investigate the consequences of those biases for inverse modelling of NMVOC emissions based on OMI data.

90 The manuscript is structured as follows. Section 2 describes the OMI retrieval, the network of FTIR data and the airborne datasets used in this work; Section 3 presents the validation methodology, the model setup and the emission inversions; Section 4 presents the correction of OMI biases using FTIR and aircraft data, and proposes a bias-correction formula for use in inverse modelling; Section 5 presents an assessment of top-down VOC emissions based on OMI, with and without bias correction; it also provides an evaluation of the optimized results against independent data, and examines the long-term variability and

95 trends of VOC emissions based on the OMI dataset between 2005 and 2016; finally, Section 6 presents the conclusions of this study.

2   Description of observational datasets

2.1   The OMI HCHO columns

The Ozone Monitoring Instrument (OMI) was launched in 2004 aboard the Aura mission in a low-Earth polar orbit crossing the

100 Equator around 13:30 LT. OMI is a nadir spectrometer that measures the solar radiation backscattered by the Earth's atmosphere and surface between 270 and 500 nm (Levelt et al., 2006). OMI has a 2600 km wide swath (divided into 60 across-track rows), providing near-global coverage every day. Due to an anomaly affecting the CCD detector, an increasing number of rows had to be filtered out, resulting in gradual degradation of the coverage to about 50% (Torres et al., 2018). The OMI ground pixel size varies from $13 \times 24\,\mathrm{km^2}$ at nadir to $28 \times 150\,\mathrm{km^2}$ at the edges of the swath.

105 The OMI HCHO dataset (v1.2, https://doi.org/10.18758/71021031) was developed within the QA4ECV project of the 7th Framework Programme of the European Union (EU-FP7). The retrieval algorithm for HCHO is described in De Smedt et al. (2018). It is based on a three-step DOAS (differential optical absorption spectroscopy) method. First, the fit of the slant columns is performed in a spectral window between 328.5 and 359 nm, with HCHO cross sections from Meller and Moortgat (2000). The slit function of each OMI row is adjusted on a daily basis as part of the wavelength calibration procedure, and the

110 absorptions cross sections are convolved accordingly. The DOAS reference spectrum is updated every day with averaged Earth radiances measured in the equatorial Pacific. The fit therefore provides a differential slant column, which corresponds to the

HCHO excess over source regions in comparison to the remote background. In a second step, the slant columns are converted to tropospheric vertical columns using a lookup table of vertically-resolved air mass factors calculated at 340 nm using the VLIDORT v2.6 radiative transfer model (Spurr, 2008). Surface albedo is obtained from the monthly OMI climatology at 0.5° resolution (Kleipool et al., 2008). Daily a priori vertical profiles are obtained from the TM5 analysis, at 1° spatial resolution (Williams et al., 2017). The standard AMF calculations uses the effective cloud fraction and cloud top pressure from the Fresco v7 cloud product (Veefkind et al., 2016), treating clouds as Lambertian reflectors and applying the independent pixel approximation (Martin et al., 2002; Boersma et al., 2004). However, in this work, the cloud correction is switched off, except for a strict filtering (effective cloud fraction > 0.2). Clear-sky air mass factors (AMF) are used in lieu of the cloud-corrected AMFs. This choice ensures an optimal consistency with the TROPOMI HCHO dataset (De Smedt et al., 2021). Indeed, the TROPOMI HCHO retrieval is inherited from the QA4ECV algorithm with the aim to generate a consistent time series of early afternoon observations. Finally, to correct for any global offset and for stripes arising between the rows, a background correction is applied on daily basis using the HCHO slant columns over the Pacific Ocean. The TM5 HCHO model columns derived in the same region are finally added to compensate for the background HCHO concentrations. Several diagnostic variables are provided together with the measurements. Besides the cloud filter, the recommended processing is applied, in particular with respect to the row anomaly (De Smedt et al., 2018). For every pixel, column averaging kernels and a priori profiles are provided, as well as the random and systematic components of the tropospheric column uncertainty.

2.2   Ground-based FTIR HCHO data

The FTIR stations participating in the present study are listed in Table S1 (see also Fig. 1(b)). Most stations are affiliated to the Network for the Detection of Atmospheric Composition Change (NDACC). The InfraRed Working Group (IRWG) of NDACC (https://www2.acom.ucar.edu/irwg) requires standardized high-quality instruments, namely high-resolution spectrometers mostly from the same manufacturer (Bruker 120HR or 125HR), and homogenized retrieval algorithms: PROFITT9 (Hase et al., 2006) and SFIT4 (Pougatchev et al., 1995). These codes retrieve information on atmospheric composition from solar absorption measurements performed under clear-sky conditions in the infrared spectral region, and provide total columns as well as low-vertical resolution profiles of many atmospheric species.

Harmonized retrieval settings have been set up and used within the whole IRWG in order to build a consistent ground-based FTIR HCHO data set for robust interpretation of satellite and model validation (Vigouroux et al., 2018). This data set has proven its value for the validation of TROPOMI HCHO (Vigouroux et al., 2020, see also the quarterly reports https://mpc-vdaf.tropomi.eu), GOME-2 (https://acsaf.org/valreps.php) or more recently of OMPS (Kwon et al., 2023). We refer to Vigouroux et al. (2018) for details about the HCHO retrieval settings. The most critical aspects with regard to harmonization within the network are the spectral signatures and the spectroscopic parameters. The HCHO fitted spectral signatures belong to the ν_1 and ν_5 vibrational bands, around 3.6 μm. The spectroscopic database is the atm16 line list from G. Toon (http://mark4sun.jpl.nasa.gov/toon/linelist/linelist.html), which corresponds to HITRAN 2012 (Rothman et al., 2013) for formaldehyde. Low-vertical resolution formaldehyde profiles are obtained using Tikhonov regularization (Tikhonov, 1963). However, the degrees of freedom for signal (DOFS) are low for HCHO (from 1 to 1.6) due to its weak spectral signature, im-

plying that mainly a total column can be retrieved. The sensitivity of the measurement is highest in the free troposphere, but a good sensitivity is achieved near the surface (see Fig. 4 of Vigouroux et al. (2018)). The FTIR HCHO uncertainty is calculated according to Rodgers (2000). The systematic uncertainty, mostly due to uncertainty on the spectroscopic parameters, is $\sim 13\%$ (Vigouroux et al., 2018). The random uncertainties can be as low as 1.0×10^{14} molec. cm^{-2} (8%) for clean sites such as Eu-
150 reka and up to 5.3×10^{14} molec. cm^{-2} (7%) for polluted stations such as Paris. Only the Mexico City site has a larger random uncertainty (11.1×10^{14} molec. cm^{-2}) because it operates a low-resolution spectrometer (Vertex 80). Additionally, note that significant random and systematic smoothing uncertainties also exist (Vigouroux et al., 2018); however, those uncertainties vanish when the FTIR averaging kernels are used to smooth the model profiles, as done in this study (Rodgers and Connor, 2003).

2.3 Airborne HCHO data

155

Table 1 lists the aircraft campaign data used in this study. Among those, four HCHO datasets from three campaigns conducted in 2012-2013 over the United States are used to evaluate OMI HCHO columns, as described in Sect. 3.3. All campaigns are used to evaluate the inverse modelling results. More specifically,

- the DC3 (Deep Convective Clouds and Chemistry) NASA mission sampled the atmospheric composition over the Central
160 U.S. in May-June 2012 (Barth et al., 2015). Formaldehyde mixing ratios were acquired from two aircraft platforms, the NASA DC-8 and the NSF/NCAR Gulfstream V (GV), equipped with similar infrared absorption spectrometers employing difference frequency generation (DFG) laser sources: the DFGAS (Difference Frequency Generation Absorption Spectrometer) instrument on the DC-8 (Weibring et al., 2007) and the more sensitive CAMS (Compact Atmospheric Multispecies Spectrometer) instrument on the GV (Richter et al., 2015). More details regarding the instruments can
165 be found in Fried et al. (2016). A comparison between the DFGAS and CAMS measurements indicates a very good agreement between the two instruments, with the DFGAS measurements being slightly higher (by 11%, Fried et al. (2016)).

- the SEAC^{4}RS (Studies of Emissions, Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys) campaign was conducted over the southeastern U.S. in August-September 2013 on board the NASA DC-8 aircraft (Toon
170 et al., 2016). The mission took place place over regions rich in biogenic VOC emissions. Two different instruments were used: CAMS and the ISAF (In situ Airborne Formaldehyde) instrument (Cazorla et al., 2015). The ISAF data were found to be about 10% higher than the CAMS values (Zhu et al., 2016). Only the CAMS data are used here.

- the DISCOVER-AQ (Deriving Information on Surface conditions from Column and Vertically Resolved Observations Relevant to Air Quality) was a multi-year aircraft mission led by NASA (Crawford and Pickering, 2014). The
175 HCHO measurements were acquired using the DFGAS technique (Weibring et al., 2007). The campaign over California (January-February 2013) is used in the OMI HCHO validation. The campaigns over Maryland (2011) and Colorado (2014) are used for model evaluation.

(a)

(b)

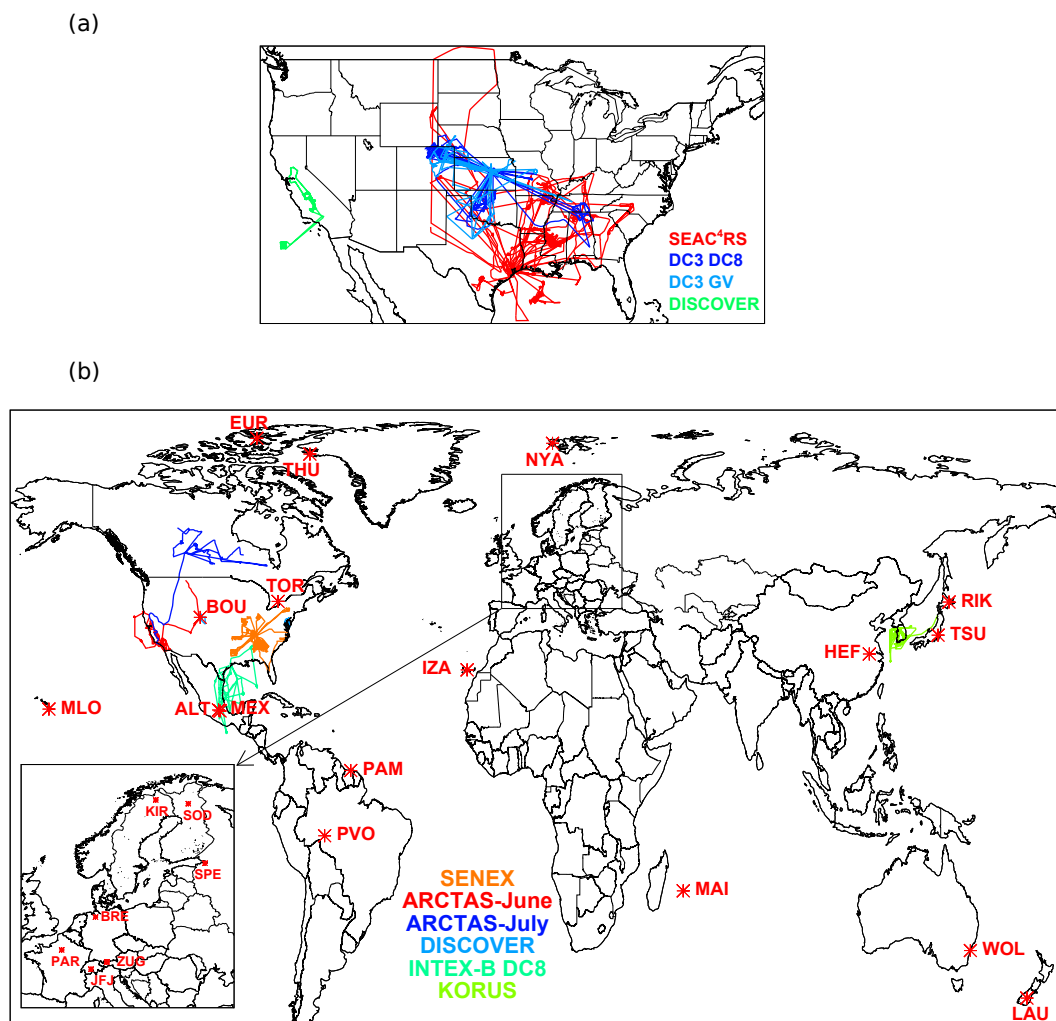



Figure 1. (a) Flight tracks of the aircraft missions (DC3, SEAC[4]RS, DISCOVER-AQ California) used as constraints in the aircraft-based inversion, (b) flight tracks of the additional aircraft campaigns used for global model evaluation, and location of FTIR stations used in this work to evaluate OMI HCHO columns. The station coordinates and codes are provided in Table S1. Note that the stations Garmisch and St.Denis are not shown, given their close proximity to the stations Zugspitze (ZUG) and Maïdo (MAI), respectively.

- the INTEX (Intercontinental Chemical Tranport Experiment) campaigns took place in 2004 and 2006. We use here the measurements of the March 2006 INTEX-B campaign, also named MILAGRO (Megacity Initiative: Local and Global Research Observations) (Molina et al., 2010). The NASA DC8 measured chemical composition over Mexico, Texas and the Gulf of Mexico below ca. 10 km altitude. HCHO was measured using Tunable Diode Laser Spectroscopy (TDLAS) (Fried et al., 2008). The INTEX-B campaign of April-May 2006 is not used here, as it was conducted primarily over the Pacific Ocean, where HCHO levels are insensitive to emissions over land.

180

Table 1. Aircraft campaign datasets used in this work for OMI and model evaluation. Datasets 1-4 are used to determine the OMI biases. All datasets are used to evaluate the top-down emission inversions.

Dataset number	Aircraft dataset	Period	Measurement technique	Reference
1	DC3 (DC8)	May-June 2012	DFGAS	Weibring et al. (2007)
2	DC3 (GV)	May-June 2012	CAMS	Richter et al. (2015)
3	SEAC^{4}RS	Aug.-Sep. 2013	CAMS	Richter et al. (2015)
4	DISCOVER-AQ California	Jan.-Feb. 2013	DFGAS	Weibring et al. (2007)
5	MILAGRO (DC8)	March 2006	TDLAS	Fried et al. (2008, 2011)
6	ARCTAS-CARB	June 2008	TDLAS	Fried et al. (2008)
7	ARCTAS-B	July 2008	TDLAS	Fried et al. (2008)
8	DISCOVER-AQ Maryland	July 2011	DFGAS	Weibring et al. (2007)
9	DISCOVER-AQ Colorado	July-Aug. 2014	DFGAS	Weibring et al. (2007)
10	SENEX	June-Jul 2013	ISAF	Cazorla et al. (2015)
11	KORUS-AQ	May-June 2016	CAMS	Richter et al. (2015); Fried et al. (2020)

- the ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites) campaigns took place in 2008 (Jacob et al., 2010). Whereas the ARCTAS-A campaign (not used here) targeted the springtime atmospheric composition over the Arctic, the ARCTAS-CARB (June) and ARCTAS-B (July) campaigns sampled the tropospheric composition during the summer between 34 and >80° N. Only data below 60° N are used here. The TDLAS technique (see above) was employed to measure HCHO during both missions.

- the Southeast Nexus (SENEX) campaign (Warneke et al., 2016) used the NOAA WP-3D aircraft to sample the lower troposphere (below ca. 6 km altitude) over the Southeast U.S. in June 2013. HCHO was measured using the ISAF instrument (see above).

- the KORUS-AQ (Korea-United States Air quality) campaign investigated air composition over Korea and surrounding areas in May-June 2016. Formaldehyde was measured throughout the troposphere by the NASA DC8 using CAMS (see above).

For all campaigns, we exclude HCHO measurements from urban plumes (identified as [NO$_2$]> 4 ppbv or [NO$_x$]/[NO$_y$]$>$ 0.4) and biomass burning plumes ([CH$_3$CN]> 225 pptv), as well as nighttime data (before 9 AM or after 18 PM local time) and measurements over oceans. Data above 4 km altitude are also excluded. Furthermore, as the average HCHO mixing ratio from the CAMS instrument were about 13% lower than those of the ISAF instrument during SEAC^{4}RS, we increase the CAMS values by 6.5% in order to bring the SEAC^{4}RS data midway between CAMS and ISAF, while also reducing the bias between the GV data (CAMS) and DC8 data (DFGAS) from the DC3 mission.

The measurements are publicly available via data archive centers (see Data availability). The flight tracks are shown in Fig. 1.

2.4 Spaceborne isoprene columns from CrIS

The global modelling results of Sect. 5 will be evaluated against emissions derived from spaceborne isoprene column measurements (Sect. 5.3). The first global isoprene observations from space were derived through the application of an efficient

205 machine-learning algorithm to thermal infrared radiances measured by the Cross-track Infrared Sounder (CrIS) (Wells et al., 2020). Isoprene column abundances were derived from on-peak/off-peak brightness temperature differences at the ν_{28} absorption feature. Isoprene emissions for three months (January, April and July 2013) were derived from the CrIS measurements and the GEOS-Chem chemistry-transport model through the application of an iterative mass balance algorithm. Prior to the inversion of isoprene fluxes, NO_x fluxes were inverted using spaceborne NO_2 column data from the OMI sounder. This NO_x

210 adjustment was found critical due to the high sensitivity of isoprene columns to OH (and hence NO_x) fields in the model. The isoprene inversion pinpointed several regions where emission errors for both isoprene and NO_x cause major prediction biases in models.

3 Methodology

3.1 Validation using FTIR measurements

215 The validation of OMI HCHO data using the FTIR ground-based network follows the methodology applied for TROPOMI by Vigouroux et al. (2020). Only the collocation criteria are chosen differently, due to the lower precision of OMI compared to TROPOMI columns: we average the OMI pixels within 50 km around each station, which leads to a mean value of 13 pixels per collocation. Only pixels satisfying the recommended quality criteria (Sect. 2.1) are averaged. In addition, the collocated pair is kept only if at least 10 valid OMI pixels are available for averaging.

220 We apply the formalism of Rodgers and Connor (2003) in order to account for the different a priori vertical profiles in the FTIR and OMI retrievals (single climatological profile from the WACCM model and daily profiles from TM5, respectively) and for the different vertical sensitivity of both instruments. For each FTIR profile in time coincidence (within 3 hours of the OMI overpass time) and for each individual OMI pixel (within 50 km of the station), the OMI a priori profile is substituted for the FTIR profile (see Eq. 2 in Vigouroux et al., 2020), and the corrected FTIR profile is smoothed after regridding (and

225 extrapolation if needed) to the satellite grid using the OMI averaging kernel (Eq. 3 of Vigouroux et al., 2020). The vertical re-gridding ensures that the same upper boundary for the tropospheric column definition is used in both products (the one from OMI). In addition, a scaling factor is applied to the OMI and smoothed FTIR columns to correct for the altitude difference between the OMI pixel and the the station (Eq. 4 of Vigouroux et al., 2020). That factor is taken as the fraction of the a priori OMI column that lies above the station altitude.

230 In summary, each coincident pair consists of the average of the smoothed and scaled FTIR columns within 3 hours of the OMI overpass and the average of the individually-scaled OMI pixels within 50 km of the station.

3.2 HCHO simulation using the MAGRITTEv1.1 CTM

The Model of Atmospheric composition at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE v1.1) is a chemistry-transport model based on the previous IMAGES (Intermediate Model for the Annual and Global Evolution of Species) model (Muller and Brasseur, 1995; Stavrakou et al., 2018). MAGRITTE v1.1 calculates the distribution of 182 chemical species, among which 141 compounds undergo transport (advection, deep convection, and turbulent mixing in the boundary layer) in the model. The chemical mechanism includes a detailed description of isoprene and other biogenic volatile organic compounds oxidation (BVOCs) mechanisms (Müller et al., 2019). In particular, it incorporates an up-to-date representation of isoprene peroxy radical unimolecular reactions and other recent mechanistic advances relevant to BVOC oxidation (Müller et al., 2019). The photolysis rates are interpolated from tabulated values calculated using the TUV photolysis estimation package (Madronich and Flocke, 1998). Most model parameterizations, including the chemical mechanism for anthropogenic and pyrogenic organic compounds are obtained from the IMAGES model (Stavrakou et al., 2009; Bauwens et al., 2016). MAGRITTE can be run either as a global model, at $2° \times 2.5°$ resolution, or as regional model, at $0.5° \times 0.5°$ resolution. In regional mode, the boundary conditions of the lateral borders are provided off-line by the global model. The chemical concentrations are calculated on a σ-pressure coordinate grid, with 40 vertical levels distributed within the troposphere and the lower stratosphere (below the 44 hPa level).

Meteorological fields are obtained from the ERA5 ECMWF reanalyses (Hesbach et al., 2020). The effect of diurnal variation on the photolysis rates and kinetic rate constants are taken into account through correction factors calculated from model simulations with a 20-min time step. These correction factors are used to calculate the diurnal cycle of HCHO columns required for the comparisons with ground-based and aircraft measurements.

Anthropogenic emissions of CO, NOx, SO_2 and carbonaceous aerosols are taken from the HTAPv2 (Hemispheric Transport of Air Pollution version 2) inventory (Janssens-Maenhout et al., 2015) for 2010. The anthropogenic NOx emissions over the U.S. are adjusted (Travis et al., 2016; Müller et al., 2019) to match observed NO_x concentration and HNO_3 deposition data. The speciated emissions of NMVOCs are obtained from the EDGARv4.3.2 inventory (Huang et al., 2017) between 2005 and 2012, and are taken equal to their 2012 values afterwards. The global annual anthropogenic NMVOC source is estimated at 162.3 Tg in 2005 and increases annually by $> 1\%$ to reach 179.8 Tg in 2012. Vegetation fire emissions are provided from the GFED4s database (van der Werf et al., 2017), with vertical injection profiles from Sofiev et al. (2013). The global biomass burning flux is estimated to range between 78 and 107 $\mathrm{Tg\,yr^{-1}}$, and the average global flux over 2005–2017 was 90.5 $\mathrm{Tg\,yr^{-1}}$. Isoprene and monoterpene fluxes are calculated by the MEGAN model embedded in the MOHYCAN canopy environment model (Müller et al., 2008; Guenther et al., 2012; Bauwens et al., 2018) at $0.5° \times 0.5°$ based on the ERA5 reanalysis meteorological fields and Leaf Area Index (LAI) data from MODIS Collection 6 reprocessed by Yuan et al. (2011). The CO_2 inhibition effect is accounted for using the parameterization of Possell and Hewitt (2011). The effects of soil moisture stress are neglected, since previous model evaluations against OMI data have shown a deterioration of temporal correlation when accounting for the soil moisture activity factor calculated using MEGANv2.1 (Guenther et al., 2006) and ECMWF soil moisture fields (Bauwens et al., 2018; Stavrakou et al., 2018). The global annual isoprene flux ranges between

414 Tg (in 2008) and 452 Tg (in 2016). The average annual flux amounts to 433 Tg yr^{-1} over 2005-2017. Biogenic methanol is also calculated according to MEGAN as described in Stavrakou et al. (2011). The monthly fluxes of isoprene and methanol emissions are available online at http://emissions.aeronomie.be. The annual monoterpene fluxes range between 109 Tg (in 2008) and 120 Tg (in 2016). Biogenic acetaldehyde and ethanol emissions (amounting to 22 Tg yr^{-1} for each compound) are

270 calculated as described in Millet et al. (2010). Biogenic emissions of C_2H_4 (scaled to 4 Tg yr^{-1} globally), HCHO (4 Tg yr^{-1}) and CH_3COCH_3 (28 Tg yr^{-1}) are also obtained from MEGAN (Guenther et al., 2012) (http://eccad.aeris-data.fr). Finally, oceanic emissions are included for acetaldehyde (56 Tg yr^{-1}), methanol (49 Tg yr^{-1}) and acetone (63 Tg yr^{-1}) (Müller et al., 2018).

3.3 Aircraft-based inversion

275 The regional MAGRITTE model with its inverse modelling capability is used to generate HCHO model distributions that closely approximate the aircraft observations from the campaigns DC3, SEAC^{4}RS and DISCOVER-AQ California (Table 1). This is realized by adjusting the monthly NMVOC emissions used in the model and minimizing a cost function which quantifies the overall discrepancy between the model-calculated mixing ratios and the observations. The cost (J) is expressed as

$$J(\mathbf{f}) = \frac{1}{2}\big[(H(\mathbf{f}) - \mathbf{y})^T \mathbf{E}^{-1}(H(\mathbf{f}) - \mathbf{y}) + \mathbf{f}^T \mathbf{B}^{-1}\mathbf{f}\big], \tag{1}$$

280 where $\mathbf{f}$ denotes the vector of dimensionless emission parameters, $H(\mathbf{f})$ is the chemistry-transport model operating on the control variables, $\mathbf{y}$ is the observation vector, and $\mathbf{E}$, and $\mathbf{B}$ are the covariance matrices of the errors on the observations and the emission parameters $\mathbf{f}$, respectively. For each observation, the model is sampled at the same day, hour, pixel and altitude as the measurement. However, the observation vector $\mathbf{y}$ and its model counterpart $H(\mathbf{f})$ consist of campaign-averaged (observed or modelled) mixing ratios in each model pixel (0.5°×0.5°) for which observations are available. In this way, different model

285 pixels contribute similarly to the cost function, despite the spatially heterogeneous sampling of air composition by the aircraft. In this way, we avoid giving excessive weight to intensively surveyed areas (e.g. the Houston ship channel during the SEAC^{4}RS campaign, Toon et al., 2016), since those areas were often chosen due to special features (e.g. high pollution levels) and might not be representative at a larger scale.

The monthly averaged emission flux for a given category (anthropogenic, pyrogenic or biogenic) is expressed as

$$G(\mathbf{x}, t, \mathbf{f}) = \sum_{j=1}^{m} \exp(f_j)\phi_j(\mathbf{x}, t) \tag{2}$$

290 with ϕ_j the a priori emission distributions detailed in Section 3.2. The emission for a given category and pixel is not optimized when its maximum monthly value over the course of the year is lower than 10^9 molec. cm^{-2} s^{-1}. This threshold is sufficiently low that the emission of most pixels are optimized over the contiguous U.S. for the biogenic and anthropogenic categories. The total number of optimized parameters is $\sim 2.2 \times 10^5$.

The matrix $\mathbf{E}$ is assumed diagonal, and includes instrumental errors as well as representativity and model errors. The total

295 uncertainty is derived by quadrature addition of a 15% relative uncertainty and a 200 pptv absolute error. The 15% error is

slightly higher than both the estimated instrumental systematic uncertainty (e.g., 12.4% for the TDLAS instrument) (Fried et al., 2008) and the typical bias between different measurement techniques (see above). The 200 pptv contribution is higher than the limit of detection, typically lower than 100 pptv (Fried et al., 2008) but is intended to give more weight to higher HCHO mixing ratios in the cost function.

300   The diagonal elements of $\mathbf{B}$ are taken equal to 1.1^2, i.e. the errors on all emission parameters are assumed to be a factor 3 $(e^{1.1})$. Anthropogenic emission parameters from pixels in the same country are weakly correlated (coefficient of 0.1), whereas parameters for different countries are assumed uncorrelated. For biogenic and pyrogenic emissions, a decorrelation length of 100 km is used. The cost function is minimised using an quasi-Newton optimization algorithm involving the calculation of the gradient of the cost function by the adjoint of the model (Müller and Stavrakou, 2005). The convergence criterion is a reduction

305   of the norm of the gradient of the cost J by an order of magnitude. Typically, this criterion is reached after 20 iterations.

The model domain includes the contiguous U.S. ($10\text{-}54°\text{N}$, $65\text{-}130°\text{W}$). Simulations start on July 1st 2011 and last 2.5 years.

The optimized HCHO distributions are used to calculate, for each campaign, a campaign-average gridded HCHO column field accounting for the sampling times and averaging kernels of the OMI retrievals. Those columns are then compared to the corresponding OMI columns at the locations of the aircraft measurements aggregated onto the model grid. Model pixels with

310   less than 10 OMI measurements, or less than 5 aircraft measurements below 4 km are excluded from analysis.

3.4   Inversion constrained by satellite HCHO columns

The methodology for optimizing global NMVOC emissions based on OMI data is similar to that presented in the previous section, except (1) the global model is used, (2) monthly-averaged bias-corrected OMI HCHO columns binned onto the model resolution ($2°\times2.5°$) are used as contraints, (3) a decorrelation length of 300 km is assumed for a priori error correlations in

315   the biogenic and pyrogenic sectors, to account for the coarser model resolution (4) separate inversions are performed for each year between 2005 and 2017, and each simulation starts on July 1st of the year preceding the target simulation year. For each optimization, the number of optimized parameters is $\sim 10^5$. The convergence criterion (reduction by a factor of 30 of the norm of the gradient of J) is attained after typically $15-20$ iterations.

The OMI column uncertainty is obtained by quadrature addition of the OMI retrieval uncertainty and an absolute error taken

320   to be $2 \times 10^{15}\,\text{molec. cm}^{-2}$. The retrieval error consists of a systematic and a random component, but the latter is greatly reduced upon averaging. In order to limit the noise, monthly averages based on less than 20 valid measurements are excluded from analysis. The systematic uncertainty is typically $35-55\%$ over source areas in tropical regions and during summertime at mid-latitudes. In winter at mid-latitudes, the retrieval error usually exceeds 80%. The OMI monthly averages are compared to the corresponding MAGRITTE monthly averages. Those are calculated from daily values accounting for the number of

325   measurements and averaging kernel for each day (also binned onto the model resolution) and for the sampling time ($\sim$13h30 LT) of observation at each location.

The determination of VOC emissions from satellite HCHO data has several limitations. Although the fluxes from three emission categories are inverted simultaneously through the minimization of the cost function (Eq. 1), the distinction between these categories is uncertain, in particular at places and times where more than one category is dominant. The optimization realizes

the separation largely based on the a priori magnitude and spatio-temporal patterns of the emissions, through the correlation between a priori errors on the emission parameters. Therefore, errors in the a priori emission distributions might cause errors in the attribution of emissions between different categories. Fortunately, a single emission category is very often dominant over continental areas, e.g. anthropogenic emissions are strongly dominant over northeastern China and biogenic emissions are dominant over the Eastern U.S. and most tropical forests. However, biomass burning is a highly episodic source which generally coincides with biogenic source areas, resulting in uncertain top-down emissions for both biogenic and pyrogenic sources. The same is also true in areas (e.g. India) where both anthropogenic and biogenic emissions are significant. In those regions, the total top-down VOC emissions are much better resolved than individual categories.

In addition, the top-down VOC emissions have uncertainties related to the multiple factors that might affect the abundance of HCHO, besides the magnitude of the emissions. This includes, for example, the background HCHO abundance, largely determined by OH radical levels, the oversimplified speciation of VOCs in large-scale models, the VOC chemical oxidation mechanisms, the deposition of VOC oxidation intermediates, the diurnal cycle of emissions, especially for biomass burning, the vertical transport processes control the vertical profile of HCHO, and the NO_x levels, which influence the yields of HCHO from many important VOCs. Although a few of those uncertainties were partially addressed in previous studies (e.g. Oomen et al., 2023), an exhaustive quantitative study of those uncertainties would be a daunting task and is beyond the scope of the present study.

4 Results

4.1 OMI HCHO bias determined using FTIR data

Due to the higher noise of OMI data (compared to TROPOMI), a relatively poor correlation is obtained between individual coincident pairs, with a Pearson's coefficient of 0.55, whereas a correlation coefficient of 0.81 was found in the evaluation of TROPOMI data against FTIR data (Vigouroux et al., 2020). We use therefore the monthly means of coincident pairs to derive a more robust linear relationship between OMI and FTIR columns. The scatter plot of the coincident monthly-averaged FTIR and OMI columns is shown in Fig. 2. The comparison of monthly means shows a Pearson's correlation coefficient of 0.67. The regression using the Theil-Sen estimator (Sen, 1968) yields a positive OMI constant bias (intercept of 2.02×10^{15} molec. cm^{-2}) and a negative OMI proportional bias (slope of 0.659). The positive OMI bias for clean sites and negative bias for polluted sites is similar as for TROPOMI validation (Vigouroux et al., 2020). We tested alternative choices for the collocation distance: a higher value (100 km instead of 50 km) degrades the correlation and yields a slightly lower slope (0.63) than our reference regression. Lower distances (e.g. 20 km) lead to an excessively low number of OMI pixels to be averaged ($\ll 10$) and therefore to poor correlation with FTIR.

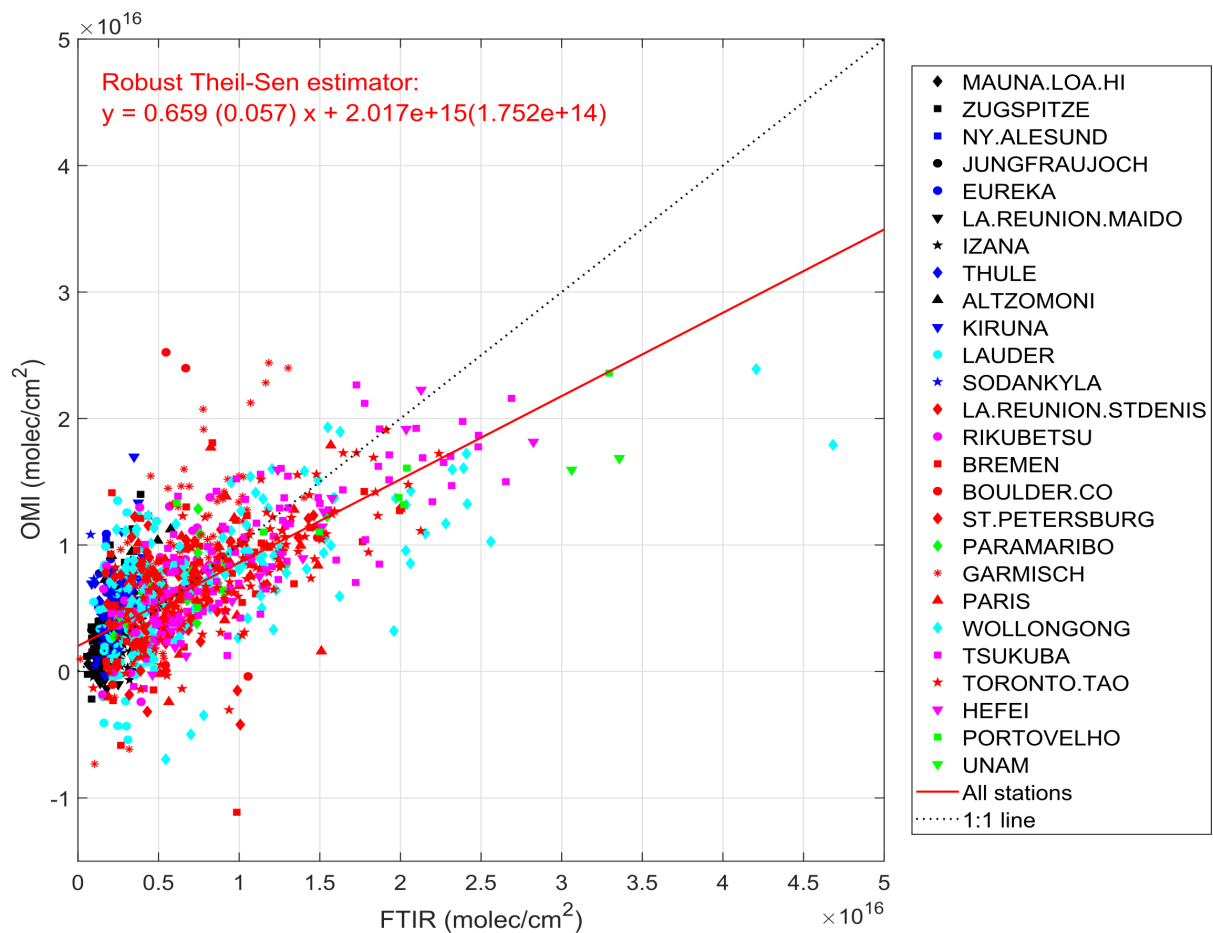



Figure 2. Scatter plot of co-located OMI and FTIR HCHO monthly columns over 2005–2017. The red line represents the Theil-Sen regression. Its slope and intercept (with their 1-σ errors between brackets) are given inset in red.

4.2 OMI HCHO bias determined using aircraft data

360   Here the OMI data are being evaluated against aircraft in situ measurements, using the regional MAGRITTE model as transfer standard. Figure 3(a-c) displays the distribution of observed HCHO mixing ratios from the three aircraft campaigns. The highest concentrations ($> 3\,\mathrm{ppbv}$) are seen in the Southeast during summertime, whereas lower, but still significant levels ($1-3\,\mathrm{ppbv}$) are observed over the Central U.S. (Colorado, Nebraska) during summer and over California in winter. The model with a priori emissions reproduces very well the high summer values over the Southeast, with very little bias found against SEAC[4]RS data

365   (Fig. 3(d-f)). Over California and Central U.S., substantial model underestimations are found, reaching typically a factor of 2. Similar results were obtained by Chen et al. (2019) with the GEOS-Chem model, indicating a good model agreement in regions dominated by biogenic VOC emissions according to MEGAN (i.e. the Southeast U.S.), and large underestimations

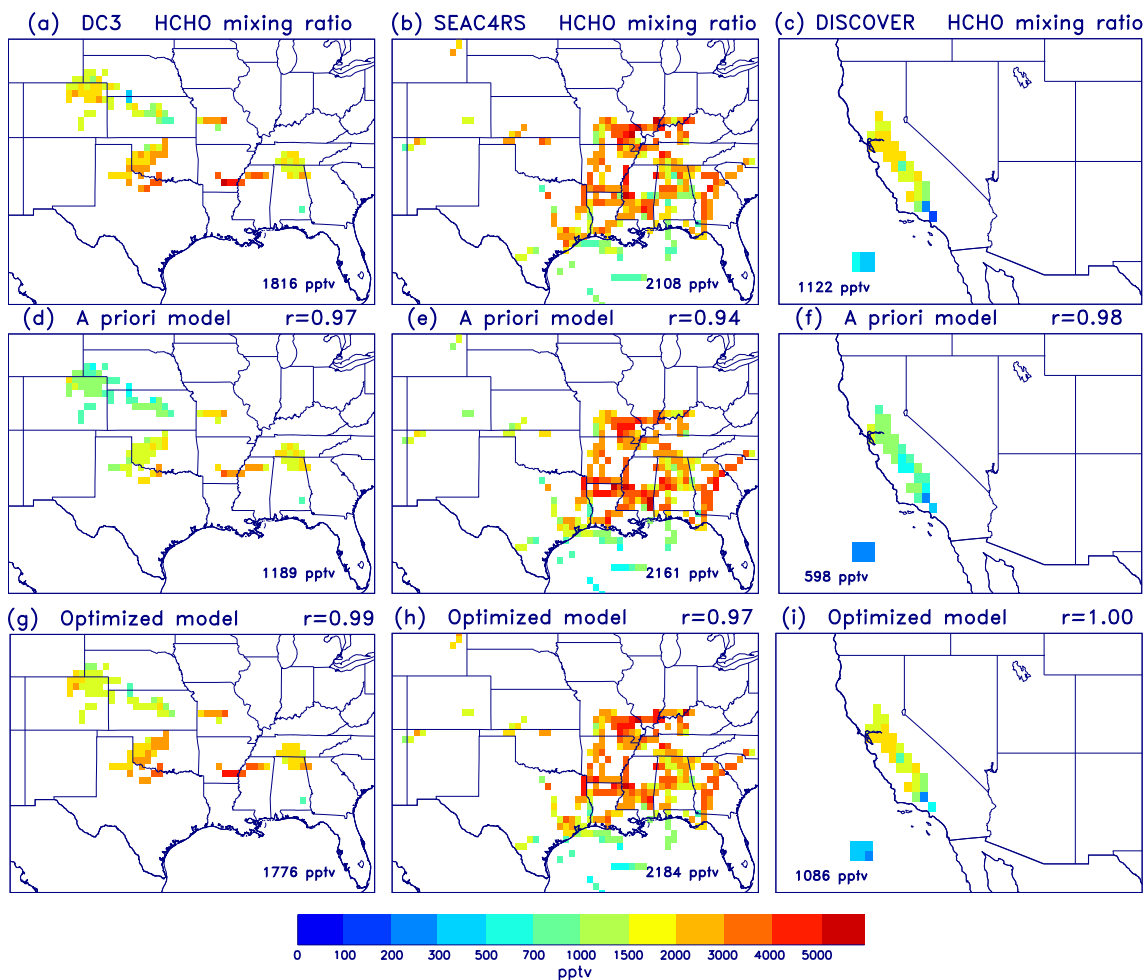



Figure 3. Measured HCHO mixing ratio ($pptv$) below 4 km altitude from the campaigns (a) DC3, (b) SEAC[4]RS, and (c) DISCOVER-AQ California, and corresponding model distributions using a priori emissions (d, e, f) and with optimized emissions (g, h, i). Urban and biomass burning plumes were filtered out as described in the text. The data are gridded to the model resolution. The averaged measured or modelled mixing ratio is given inset for each campaign.

over California and the Central U.S., as shown by extensive comparisons with aircraft campaigns. Not only HCHO, but many VOCs are similarly underestimated in those areas (Chen et al., 2019), suggesting the presence of missing VOC sources in emission inventories. Urban VOC emissions are likely strongly underestimated (Karl et al., 2018), likely partly due to the use of volatile chemical products (VCPs) (McDonald et al., 2018) and the underestimation of urban biogenic VOC emissions (Gu et al., 2021). Over Colorado, Nebraska and Texas (among others), fossil fuel exploitation is a large source of alkanes and other VOCs (Pétron et al., 2014; Franco et al., 2016; Tzompa-Sosa et al., 2019) likely responsible for a large part of the model discrepancy.

In contrast with the prior simulation, the optimized model using adjusted VOC emissions reproduces very well the observed distributions (Fig. 3(g-i)), with spatial correlation coefficients of 0.97 or more for each campaign, and average biases of at most 3%. This agreement is achieved through a substantial increase of anthropogenic emissions throughout the U.S., whereas biogenic emissions are decreased over much of Southeast U.S. and biomass burning emissions undergo a moderate increase (see Table S2 and Fig. S1 in the Supplement). Since the emission parameters are severely underconstrained by the inversion due to the poor coverage of the observations over the model domain, the solution found by the inversion has limited reliability and is strongly dependent on the a priori emission distribution and inversion setup, in particular the covariance matrix of the emission parameters. Despite this caveat, the optimization reproduces very well both the HCHO horizontal distribution (Fig. 3(g-i)) and the vertical profile shape of HCHO mixing ratios for all campaign datasets, as seen on Fig. 4. While panels (a-d) in Fig. 4 display the campaign-averaged profiles for the datasets used in the emission inversion, Fig. 4(e) shows the profiles for the SENEX campaign which took place in 2013 but was not used as constraint in the inversion. Although the observed profile shape is correctly simulated by the model, the overall agreement is significantly lower for this campaign (-22% bias) than for the campaign datasets used in the inversion. Given the significant overlap of the SENEX and SEAC^{4}RS spatial coverages (see Fig. 1), the underestimation of SENEX HCHO by the model suggests a shift of the summertime biogenic VOC emission peak towards spring, since SENEX and SEAC^{4}RS were conducted (primarily) in June and August, respectively. This finding, which is in agreement with the previous determination of isoprene emissions based on satellite (GOME) data by Palmer et al. (2006), will be further examined using OMI and the global model in Sect. 5.

For the campaigns used in the inversion, a regression of the observed and simulated concentrations yields a slope of almost 1 and a correlation coefficient of 0.93 (Fig. 5(a)). In constrast with this, Fig. 5(b) shows that the OMI HCHO columns are significantly biased with respect to co-located columns calculated using the optimized model distributions (Sect. 3.3). High OMI HCHO columns ($> \sim 12 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$) are underestimated by up to ca. 20% for columns $\sim 20 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$, whereas low columns ($< \sim 8 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$) are generally overestimated. This underestimation of high columns is relatively similar with the slight underestimation (12%) of OMI HCHO columns from the BIRA V14 product (De Smedt et al., 2015) against SEAC^{4}RS HCHO data determined by Zhu et al. (2016). It should be stressed, however, that Zhu et al. (2016) filtered out very negative HCHO columns ($< -5 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$) from the V14 product, which contributed to increase the campaign-averaged values and reduce the OMI underestimation. Without this filter, the BIRA V14 data are significantly lower than the QA4ECV columns (see Fig. S2). The underestimation of high columns is also qualitatively consistent with the mean OMI bias of -17% derived by Boeke et al. (2011) for OMI columns $> 5 \times 10^{15}\,\mathrm{molec.cm^{-2}}$, based on comparisons with several aircraft campaigns conducted in 2008.

A linear regression (Theil-Sen) of OMI and aircraft constrained model columns (Fig. 5(b)) yields

$$\Omega_{\mathrm{OMI}} = 0.651\,\Omega_{\mathrm{airc}} + 2.95 \cdot 10^{15}, \tag{3}$$

where Ω_{OMI} and Ω_{airc} are the HCHO columns ($\mathrm{molec.\,cm^{-2}}$) from OMI and from the aircraft-constrained model simulation, respectively. The regression parameters bear some uncertainty due to aircraft measurement uncertainties. As discussed in Sect. 2.3, the intercomparison of DFGAS and CAMS observations suggests a potential DFGAS overestimation by up to 11%

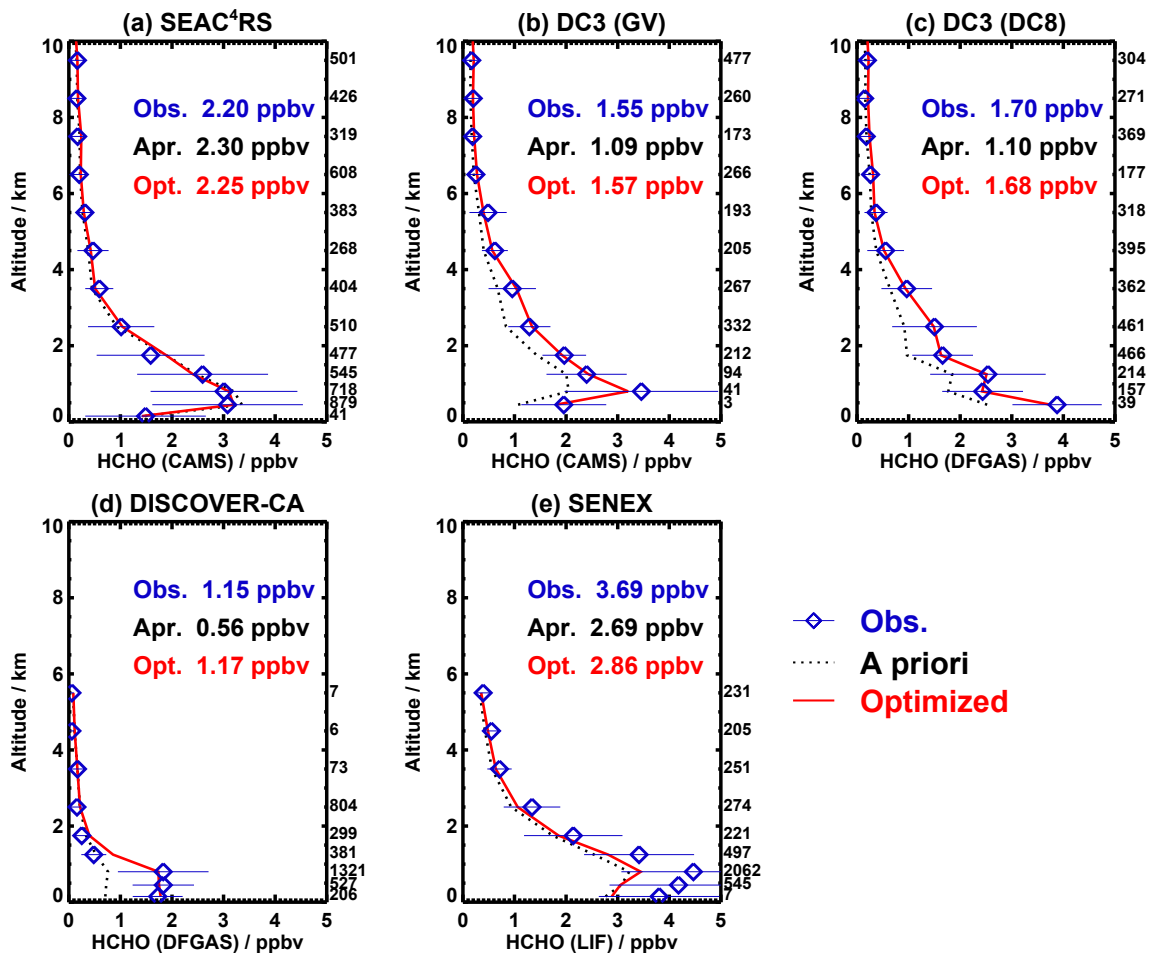



Figure 4. Measured and modelled profiles of HCHO mixing ratios over land (in ppbv) for (a) SEAC[4]RS, (b) DC3 (GV), (c) DC3 (DC8), (d) DISCOVER-AQ California, and (e) SENEX. Urban and biomass burning plumes were filtered out as described in the text. The averaged measured and modelled (a priori and optimized) mixing ratios below 4 km altitude are given inset for each dataset. The number of observations are given to the right of each plot. The boundaries of the altitude bins are 0, 0.3, 0.6, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10 km. The error bars represent the standard deviation of the measurements.

and a potential bias of our adjusted CAMS concentrations of up to $\pm 6.5\%$. Assuming that the optimized model columns

410 are proportional to the observed concentrations, the impact of those potential biases can be estimated through sensitivity regressions. The resulting slope is found to range between 0.6 and 0.7, while the intercept lies in the range $(2.6 - 3.3) \times 10^{15}\,\mathrm{molec.\,cm}^{-2}$.

The regression given by Eq. 3 is shown as a black line in Fig. 5(b) and compared with the result of the regression based on FTIR data (see Sect. 4.1), shown as a dashed red line. The slopes of the two regressions are almost identical, but the FTIR-

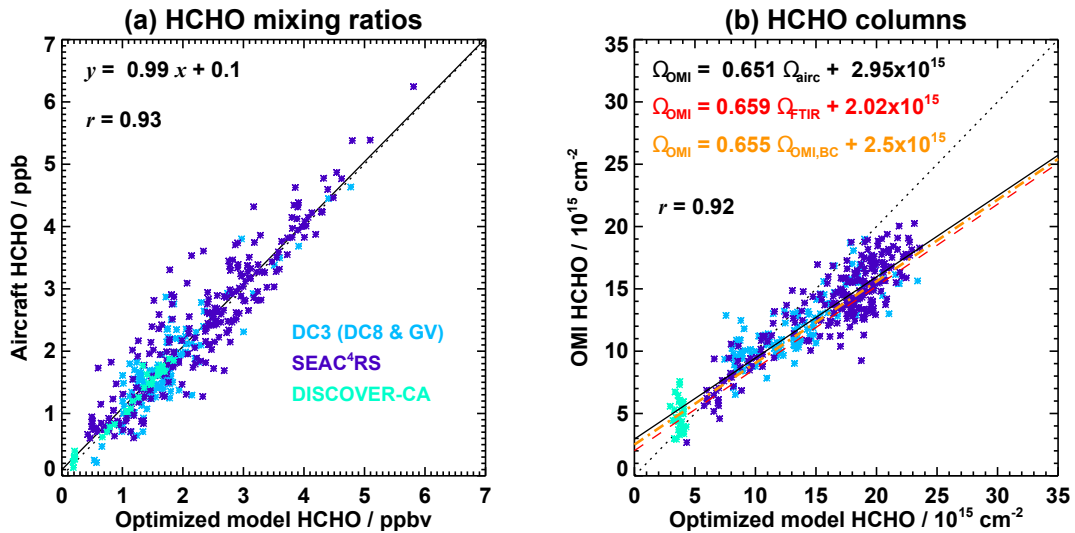



Figure 5. Scatter plots of (a) modelled and observed HCHO mixing ratios (daytime, below 4 km altitude) from three aircraft campaigns (DC3, SEAC[4]RS and DISCOVER-AQ California), and (b) modelled and OMI total HCHO columns at the same model pixels as in panel (a). The modelled values are constrained by the aircraft measurements through an emission optimization as described in text. The correlation coefficient and regression parameters using the Theil-Sen estimator are given in each panel (black font) and shown as black solid line. The slope and intercept of the regression of OMI columns vs. FTIR data (see Sect. 4.1) is given in panel (b) (red font) and shown as dashed red line. The adopted bias correction (relationship between OMI columns Ω_{OMI} and bias-corrected columns $\Omega_{OMI,BC}$) is also given (orange) and shown as orange dash-dotted line.

415 based fit has a smaller intercept and suggests a larger negative bias of high OMI columns than the aircraft-based evaluation. For example, for an OMI column of 18×10^{15} molec. cm^{-2}, the FTIR and aircraft datasets suggest underestimations by factors of 1.35 and 1.29, respectively. There might be several reasons for this, e.g. systematic uncertainties in the measurements and representativeness issues. More importantly, identical results are not expected given the different locations sampled by the two techniques. The OMI bias depends on the magnitude of the column, but likely also on other parameters. In absence of 420 additional information, we adopt a bias-correction of OMI data based on both aircraft and FTIR data, shown as orange line in Fig. 5(b). It is obtained by averaging the slopes and intercepts of the regressions of Fig. 2 and Eq. 3. The bias-corrected columns ($\Omega_{OMI,BC}$) are calculated with

$$\Omega_{OMI,BC} = (\Omega_{OMI} - 2.5 \cdot 10^{15})/0.655. \tag{4}$$

The resulting correction enhances columns above $\sim 7 \times 10^{15}$ molec. cm^{-2} and decreases columns below that value. The 425 contrasts between high- and low-emission regions will therefore be strengthened by the bias correction.

Table 2. Global emission optimizations conducted in this work.

Label	Description	Years
OPT1	OMI-based inversion, without bias correction	2006, 2008, 2011–2017
OPT2	OMI-based inversion, with bias correction	2005–2017
OPT3	as OPT2, higher prior emission errors (factor 4)	2012–2013

Note that the standard QA4ECV, cloud-corrected v1.2 product shows very similar biases with respect to aircraft data as the cloud-uncorrected product evaluated above. The evaluation of the standard product (CF< 0.4, with cloud correction) yields a slope of 0.68 and an intercept of $2.6 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$ (see Fig. S2).

5 Inferring emissions from bias-corrected OMI HCHO columns

5.1 Impact of bias correction on top-down VOC emissions

The global emission inversions are labeled as indicated in Table 2. OPT1 is conducted without any bias correction to the OMI HCHO columns, whereas OPT2 uses bias-corrected columns obtained from Eq. 4. OPT3 is a sensitivity test aimed at determining the influence of prior errors on the emission parameters. The errors on all emission parameters are taken to be a factor of $\sim 4\ (e^{1.4})$ in this case (instead of a factor of $e^{1.1} =\sim 3$ in OPT1 and OPT2, see Sect. 3.3).

Figure 6 shows seasonally-averaged HCHO columns from OMI (bias-corrected), the a priori simulation and the OPT2 simulation. Although the a priori simulation reproduces many features of the observations, there are noticeable differences, such as overestimated columns over Australia and Paraguay in December-January-February (DJF) and underestimated columns over southern Africa, Europe, East and South Asia in both DJF and June-July-August (JJA). The overestimation over Australia might be partly due to the high emission rates used in MEGANv2.1, based on measurements for young eucalypt trees which may emit more isoprene than adult trees (Emmerson et al., 2016).

As expected, the optimization realizes a much better agreement with the data, especially in areas with strong signal, such as Tropical regions during DJF and JJA, and Eastern U.S., China and India during summer. Model underestimations remain significant over regions with relatively low columns, such as Western U.S. and Europe, where the relative errors on the data are larger, particularly in winter; the retrieval error is typically $40-50\%$ over high-emission areas (e.g. Tropical forests) and $50-100\%$ over low-emission regions (e.g. Western Europe).

The improved simulation of HCHO columns from Fig. 6 is realized through significant changes in the amount and distribution of VOC emissions. Figure 7 displays the distribution of the emission ratios (optimized flux/a priori flux) per emission category, for both optimizations OPT1 and OPT2 (2011–2017 average). The mean annual emission totals in large regions are given in Table 3 for OPT2 (for 2005–2017), and in Table S3 for both OPT1 and OPT2 over 2011–2017.

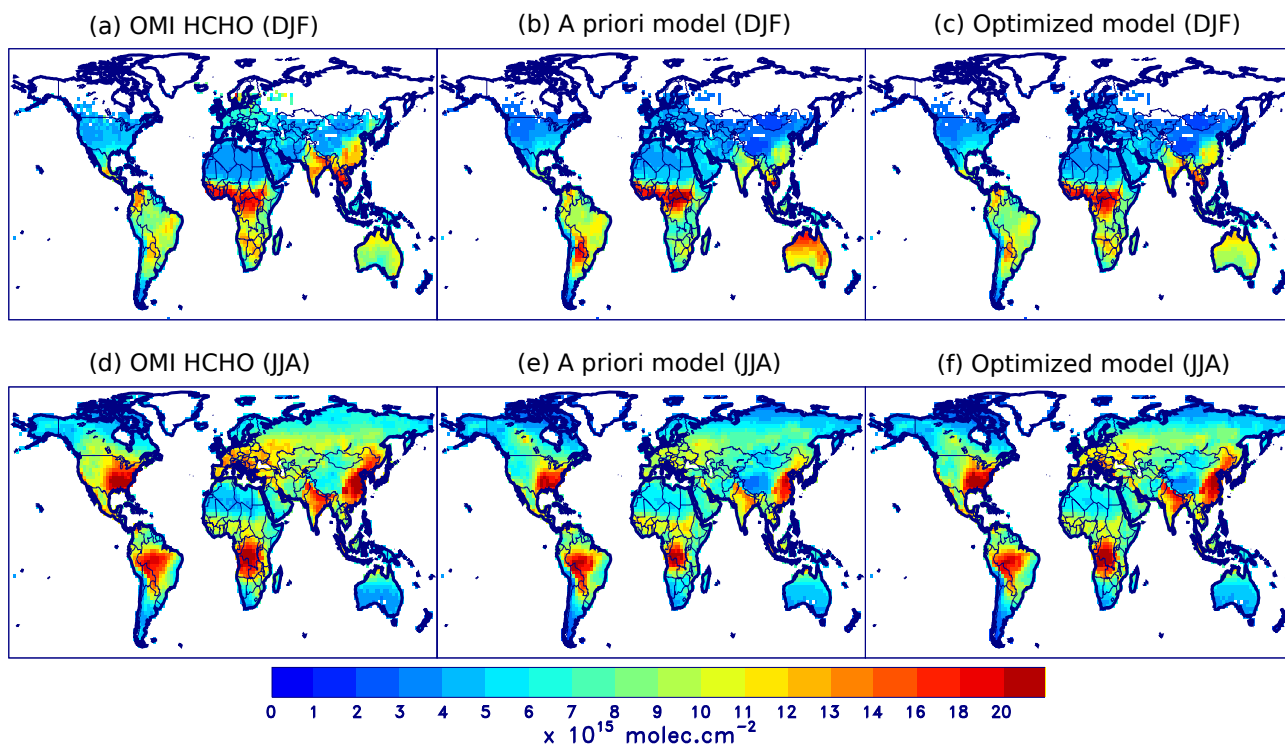



Figure 6. 2005-2017 average of HCHO columns (10^{15} molec. cm^{-2}) in December-January-February (DJF) from (a) OMI (bias-corrected as described in the text), (b) the a priori model, and (c) the model with optimized emissions. (d)-(e), same as (a)-(c) but for June-July-August (JJA).

450   The bias correction of OMI columns has a large impact on the inferred top-down emissions (Fig. S4), especially for the biogenic emission category, which is the dominant contribution to HCHO columns over continental regions (Stavrakou et al., 2009). Without bias correction, global isoprene emissions are decreased from 430 Tg yr^{-1} in the a priori from MEGAN (2011–2017 period, Table S3) to 362 Tg yr^{-1} in optimization OPT1 (i.e. a 16% decrease), whereas OPT2 increases those emissions by about 5%, to 451 Tg yr^{-1}. The impact is strongest over South Asia (+43% difference between OPT1 and OPT2). It is also

455   striking over many other regions, such as the Eastern U.S. and Southern China (Fig. 7), where biogenic emissions are decreased by up to over 30% in OPT1 but are increased in OPT2. In other regions such as the high latitudes and Australia, the difference is lower or even negative. This result is expected, given the lower columns in these regions (typically below 10^{16} molec. cm^{-2}), implying that the bias correction leads to only slight HCHO increases or even to HCHO decreases.

The top-down global isoprene emissions from OPT2 (445 Tg yr^{-1} for 2005–2013) are a factor of 1.64 higher than the top-

460   down emissions derived by inverse modelling of OMI data by Bauwens et al. (2016) (272 Tg yr^{-1}). The optimized emissions of the OPT3 run (with higher a priori emission errors) are even slightly higher (by 3%) than in OPT2. Note that our a priori (MEGAN) emissions are also higher than in Bauwens et al. (2016), especially over Australia and Africa, due to the soil

(a) Emission ratios (Optimized/prior) with bias correction

(b) Emission ratios (Optimized/prior) without bias correction

0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.4 1.7 2.0 3.0 4.0

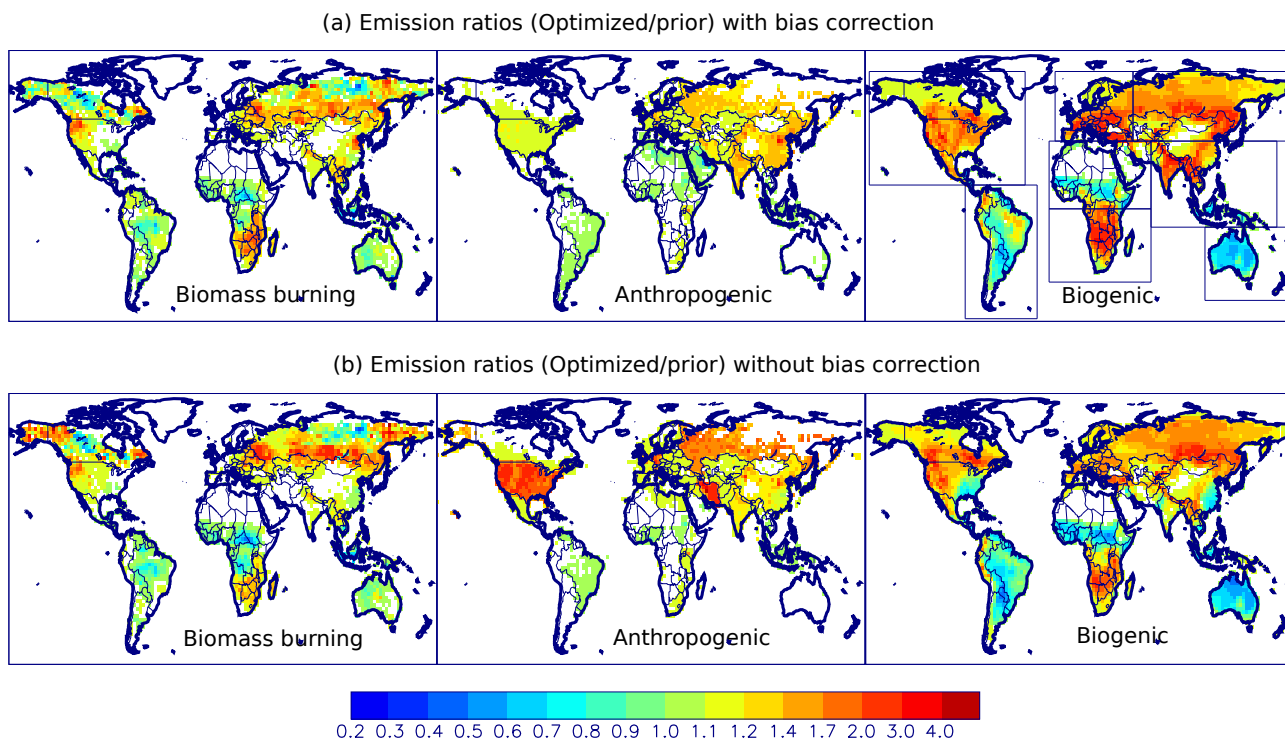



Figure 7. Ratios of top-down emissions to a priori emissions (2011–2017 average) for biomass burning VOCs (left panel), anthropogenic VOCs (middle), and biogenic VOCs (right), for optimizations (a) OPT2, constrained by bias-corrected columns, and (b) OPT1, constrained by uncorrected columns. Pixels with very small emission changes ($< 1\%$) are left blank. The regions used for calculation of total emissions given in Table 3 are shown as boxes on the top right panel.

moisture stress impact that was accounted for by Bauwens et al. (2016) but not in our study (Sect. 3.2). The bias correction of OMI data applied in this work explains only a part (factor 1.25) of the large discrepancy between the top-down estimates.
465   The main reason for the remainder is the different OMI retrieval (BIRA-V14) (De Smedt et al., 2015) used as constraint in Bauwens et al. (2016). The BIRA-V14 HCHO columns are indeed significantly lower than the QA4ECV data used here, as already noted by Wells et al. (2020). As shown on Fig. S3, QA4ECV columns are typically higher than the V14 product by $10 - 50\%$ over continents, and by up to 80% in parts of Southern Africa. This explains why the OPT1 isoprene emissions over Southern Hemisphere Africa, $60\,\mathrm{Tg\,yr^{-1}}$, are about a factor of 2 higher than the top-down estimate from Bauwens et
470   al. (2016), whereas the two top-down estimates are in much better agreement over South America and Australia, where the QA4ECV and V14 HCHO columns are more similar (Fig. S3).

     The effect of the bias correction on the optimized emissions is obtained by comparing the OPT1 and OPT2 top-down emissions. The bias correction increases the top-down biogenic emissions (by a factor of 1.25 at global scale) and biomass burning emissions (factor of 1.13), and decreases the anthropogenic emissions at mid- and high latitudes, except over China

Table 3. Mean a priori and OMI-based (OPT2) emission estimates (Tg yr^{-1}) per source category for different world regions and globally. Regions are defined in Fig. 7. The means are taken either over 2005-2017 (OPT2 simulation period) or 2005-2013 (for comparison with Bauwens et al. (2016)). B16: Bauwens et al. (2016). NH: Northern Hemisphere; SH: Southern Hemisphere.

	North America	South America	Europe	NH Africa	SH Africa	North Asia	South Asia	Oceania	**Global**
Biomass burning NMVOC emissions									
GFED4s 2005–2017	5.9	15.0	1.7	16.3	25.6	6.8	14.4	3.9	**90**
OPT2 2005–2017 (this work)	5.4	12.5	2.2	12.7	30.2	8.2	12.4	3.7	**87**
Isoprene emissions									
MEGAN-MOHYCAN 2005–2017	35	149	6.7	85	44	9.3	39	60	**430**
OPT2 2005–2017 (this work)	44	129	11.4	77	76	15	53	36	**443**
OPT2 2005–2013 (this work)	45	129	11.2	78	76	14	51	36	**445**
MEGAN-MOHYCAN 2005-2013 (B16)	32	141	6.8	50	29	9.4	36	38	**343**
OMI-based 2005–2013 (B16)	26	97	8.4	35	28	11	31	36	**272**
Anthropogenic NMVOC emissions									
EDGAR 2005–2017	20.6	12.6	17.0	39.4	11.8	11.3	52.2	1.1	**166**
OPT2 2005–2017 (this work)	21.4	11.5	18.8	34.8	12.8	14.7	59.8	1.1	**175**
Total NMVOC emissions									
A priori 2005–2017	62	176	25	141	82	28	106	65	**696**
OPT2 2005–2017 (this work)	71	154	32	128	120	38	126	43	**730**

475 (Table S3 and Fig. 7). This is primarily due to the generally low HCHO columns during winter in those regions, typically below $6.5 \times 10^{15} \,\text{molec. cm}^{-2}$, for which the bias correction decreases the columns (Eq. 4). The exacerbated seasonal variation of the HCHO columns due to the bias correction favors the biogenic emissions to the detriment of anthropogenic sources at these latitudes. This result is highly dependent on the HCHO column uncertainties, however. The scatter plots of Fig. 2 and Fig. 5 shows a large dispersion for low columns, indicating high uncertainty in the bias correction. For this reason, the optimization

480 results should be considered with caution in low-column areas.

5.2 Evaluation of optimized HCHO against aircraft and FTIR data

The optimizations OPT1 and OPT2 were conducted over years (Table 2) for which aircraft campaign data are available, over the U.S., Canada, Mexico and South Korea (Fig. 1 and Table 1). Figure 8 displays the mean observed and modelled vertical HCHO profiles for all campaigns. For most campaigns, the model agreement is improved when the bias correction is applied

485 to the satellite columns. On average for all campaigns, the large negative bias found for both the a priori simulation (-27%) and the OPT1 run (-28%) is strongly reduced in the OPT2 run (-9%) (Table S4). The root mean square deviation (RMSD)

is also decreased, from 0.85 and 0.86 ppbv in the a priori and OPT1 run to 0.55 ppbv in the OPT2 run. In particular, OPT2 presents very little biases over high-column areas (e.g. DC3, SEAC[4]RS, SENEX). The observed horizontal distribution of the vertically-averaged HCHO mixing ratios is also very well matched by the model over high-emission regions in the Eastern

490   U.S., with Pearson's correlation coefficients ranging between 0.96 and 0.98 after optimization for those campaigns (Fig. S5). The results of the OPT3 optimization are similar to, or even slightly better (for SEAC[4]RS and for DC3 DC-8) than those of OPT2 (Table S4).

In contrast with this good performance, OPT2 presents large underestimations over low-emission areas over the Western U.S., namely the DISCOVER campaigns in Colorado (-43% bias) and California (-53%). Low OMI columns were measured

495   during the latter campaign (Jan.–Feb. 2013), in the range $(3-6) \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$, such that the bias correction further decreased the columns. The OMI columns were higher during DISCOVER-Colorado ($\sim 7 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$ near Boulder, Colorado in July–Aug. 2014), but the modelled columns remained much underestimated (by $0-40\%$) after emission optimization, due to the large OMI column uncertainties.

Figure 9 displays the mean observed and modelled seasonal evolution of HCHO columns at 11 FTIR stations located in

500   source regions. The other stations (primarily high-altitude, high-latitude or maritime stations) show little sensitivity to VOC emission changes. The complete time series (2005–2017) of monthly observed and modelled (OPT2) HCHO columns are shown in the Supplement (Figs. S6 and S7). The impact of the emission optimizations is relatively small at mid- and high latitudes, and largest at the two South American stations (Paramaribo and Porto Velho). The OPT2 run performs better than OPT1, with lower RMSD values at all sites and lower biases at 8 out of 11 sites (Table S5). The OPT2 run also performs

505   better than the prior simulation at all sites except Kiruna and Sodankyla in northern Scandinavia, where the moderate biogenic emission increase inferred by the optimization (Fig. 7) enhances the small positive model bias which was already present in the a priori simulation.

At the other European stations and at North American sites, the emission increase of the OPT2 run improves the agreement with the measurements, but fails to close the gap completely (except at Garmisch), in particular during summer. This discrep-

510   ancy is mostly explained by the model underestimation of OMI columns after optimization, as seen in Fig. 6 and Fig. S8. The exception is Mexico City, where the model successfully matches the OMI columns (Fig. S8) but underestimates the FTIR columns by about a factor of 3. This poor model performance can be attributed to the coarse model resolution ($2° \times 2.5°$) being unable to resolve the megacity emissions. At Boulder, the model underestimation of both OMI and FTIR columns (by up to a factor of 2 in summer) is consistent with the model underestimation against airborne HCHO data from DISCOVER-Colorado,

515   by almost a factor of 2 (Fig. 8).

The comparison at the Amazonian sites of Paramaribo (Suriname) and Porto Velho (Rondônia, Southwestern Brazil) validates the VOC emission decrease (by $\sim 20\%$) inferred by OPT2 during the dry season (July–November). The stronger emission decrease of the OPT1 run in Rondônia (factor of 2) leads to a significant model underestimation of FTIR HCHO columns at Porto Velho. At both sites, biogenic emissions are also strongly decreased in the first half of the year (up to a factor of 3 at

520   Porto Velho in February 2017). This decrease at Porto Velho is due to very low HCHO columns observed by OMI in the first months of 2017, especially February (Fig. S8).

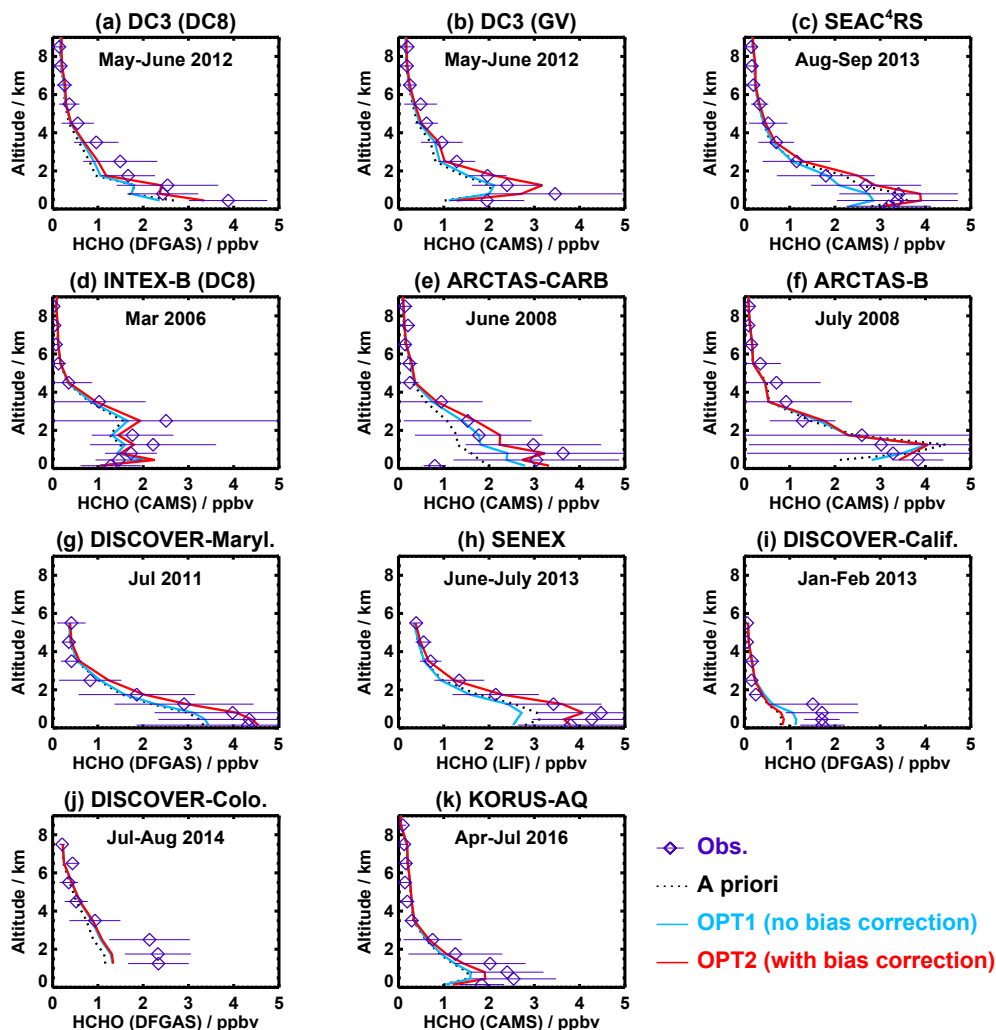



Figure 8. Measured and modelled profiles of HCHO mixing ratios over land (in ppbv) for (a) DC3 (DC8), (b) DC3 (GV), (c) SEAC^{4}RS, (d) MILAGRO, (e) ARCTAS-CARB, (f) ARCTAS-B, (g) DISCOVER-AQ Maryland, (h) SENEX, (i) DISCOVER-AQ California, (j) DISCOVER-AQ Colorado, and (k) KORUS-AQ. Optimization constrained by either uncorrected OMI columns (OPT1, in blue) or bias-corrected columns (OPT2, in red). The error bars represent the standard deviation of the measurements. Comparison statistics are provided in Table S4.

5.3 Evaluation of top-down isoprene emissions

The major source regions of biogenic isoprene according the MEGAN model are South America, Sub-Saharan Africa, Australia, the Eastern U.S. and South Asia, including South China (Fig. 10a). The fluxes display a pronounced seasonality reflecting
525 primarily their dependence on temperature and radiation (Guenther et al., 2006), with clear maxima during summertime, except

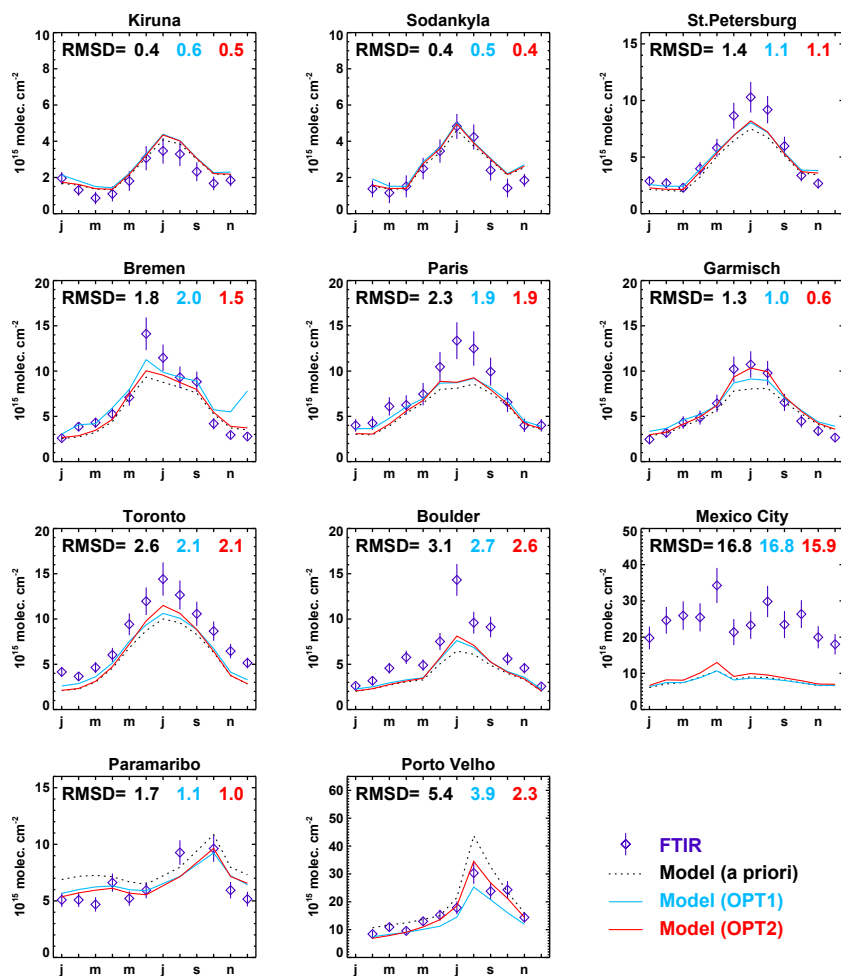
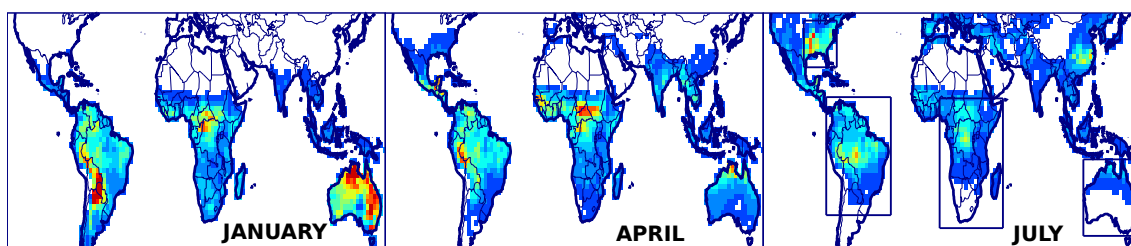



Figure 9. Average seasonal cycle of HCHO columns at FTIR stations (2011-2017). Symbols with error bars represent the mean FTIR columns and related systematic uncertainties. Black dotted curve: model with a priori emissions; blue curve: OPT1 simulation constrained by uncorrected OMI columns; red curve: OPT2 simulation constrained by bias-corrected OMI columns. The root mean square deviation (RMSD) of the 3 model runs is given inset for each station.

near the Equator. In January 2013, the highest fluxes are predicted in a region centered over Paraguay in South America as well as over northern and eastern Australia; in April, near the Peru-Brazil border region and especially over the Central African Republic (CAR) and South Sudan; and in July, in the U.S. states bordering the Mississippi River below 38°N latitude.

The OMI-based optimization (OPT2) leads to substantial changes in the distribution of isoprene emissions. As seen on 530 Fig. 10, sharp declines are predicted over most hotspots mentioned above, most noticeably those over Paraguay, the CAR, Australia and the Peru-Brazil border. Emission decreases are also found over the Guayanas and Indonesia, especially in January. Other regions emerge as important source regions, most prominently the Congo Basin and a vast area in southwestern Africa

(a) A priori (MEGAN) isoprene fluxes

(b) OMI-based (bias-corrected) isoprene fluxes

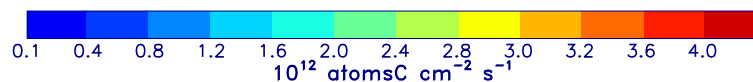
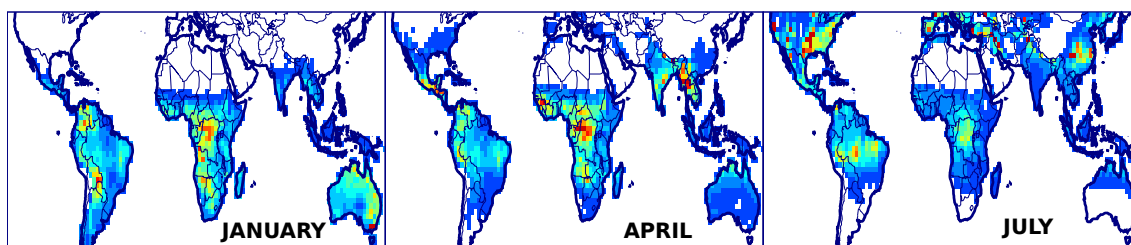



Figure 10. (a) A priori (MEGAN) isoprene monthly fluxes (10^{12} atoms C cm^{-2} s^{-1}) in year 2013. (b) Optimized isoprene fluxes based on (bias-corrected) OMI HCHO columns, also for 2013. The boxes shown in the July a priori flux subplot show the regions for which emission totals are displayed in Fig. 11.

spanning Angola, Namibia and Botswana. Although the distributions shown on Fig. 10 are for the year 2013, similar features and emission updates are derived for all years (see Sect. 5.4 for a discussion of interannual variability).

Those emission updates are in good qualitative agreement with optimized isoprene emissions based on isoprene column densities from the spaceborne Cross-track Infrared Sounder (CrIS) for the same year (2013) (Wells et al., 2020). Those authors applied an iterative mass balance technique to optimize the emissions of isoprene in the GEOS-Chem model, using as constraints the first global retrieval of isoprene columns from space. Due to the critical role played by hydroxyl radical (OH) levels in this emission inversion, and given the strong influence of nitrogen oxides (NO$_x$) on OH, the NO$_x$ emissions in GEOS-Chem were optimized based on spaceborne NO$_2$ data before conducting the isoprene inversion. The CrIS-derived monthly-averaged emissions (Supplementary Figs. 12–15 in Wells et al., 2020) display striking similarities and differences with the OMI-constrained distributions of Fig. 10b. Common features include the decreases over Paraguay, the Guyanas, the Peru-Brazil border, and the CAR and South Sudan region, as well as the increases over Congo and southwestern Africa as well as Colombia and Central U.S. states spanning from Texas to Illinois. Despite very different a priori isoprene emissions over Australia (e.g. respectively 10.8 and 4.2 Tg in January in our inventory and in Wells et al., 2020), the CrIS-based total emissions over the continent are slightly higher than our OMI-based values (Fig. 11). Closer examination shows a good agreement

in the Eastern part, whereas the CrIS-based emissions appear to be higher in the North and lower in the Central part of the country, compared to our estimates (Fig. 10). The OMI-based emissions are significantly lower than both the a priori estimate and the CrIS-based fluxes over South America in January and April, whereas the opposite holds over Africa (Fig. 11). The discrepancy is largest over Brazil, where OMI columns as low as 6×10^{15} molec. cm^{-2} are frequently observed during the wet season (December–May). Those low values drive a decrease in biogenic emissions, reaching ca. $20 - 30\%$ over Western and Northern Brazil. As discussed in Sect. 5.2, the optimized model underestimates FTIR data in February–May at Porto Velho, due to excessively low OMI columns (in February). Although the Porto Velho comparison was for a different year (2017), it suggests that the wet season OMI data might be less reliable, possibly due to high cloudiness in the area. In July, a very good agreement is found between the CrIS-based and OMI-based emissions over South America, in terms of distribution and total. Over Africa, the discrepancy is largest in January over the CAR, in a region with strong fire emissions in several biomass burning inventories including GFEDs (Pan et al., 2020). Both pyrogenic and biogenic VOC emissions are strongly reduced in this region by the OMI-based inversion, as in several previous inverse modeling studies (e.g., Bauwens et al., 2016; Müller et al., 2018), but the interference of strong pyrogenic emissions makes the top-down isoprene estimate particularly uncertain, and likely contributes to the differences with CrIS-based results. Interference from biomass burning probably plays a role in other regions as well, especially Southern Africa in June–September, the southern part of Amazonia in August–September, southeast Asia in February–April, and boreal forests during June–August. In most other regions and periods, the discrimination between pyrogenic and biogenic emissions is less critical, given the general dominance of biogenic VOCs over biomass burning as source of HCHO (Stavrakou et al., 2009, 2018). In addition, the derivation of isoprene emissions from CrIS data has also its limitation, most importantly the role played by the assumed isoprene vertical profile in the retrieval of the total column (Wells et al., 2022) and the strong impact of OH radical concentrations on CrIS-derived emissions (Wells et al., 2020).

5.4   Trends of HCHO columns and VOC emissions

Figure 12 displays the temporal evolution of annually-averaged (bias-corrected) OMI and modelled columns over large regions. The contours of those regions are shown on Fig. 13, which displays the distribution of percentage trends of HCHO columns (panels a-c) and total VOC emissions (d-e) between 2005 and 2016. All trends are obtained using a simple least-squares regression method.

Already without emission optimization, the modelled columns are strongly correlated temporally with the observations over regions where biogenic VOC emissions and biomass burning are the dominant sources of HCHO. Interannual variability over those regions is primarily related to climate variables (temperature and radiation) through their influence on biogenic emissions, biomass burning and background HCHO (Stavrakou et al., 2018), and the good model performance suggests that the inventories used in the model (MEGAN-MOHYCAN and GFED4s) are generally adequate. The year 2017 stands out in the time series over several regions in the Southern Hemisphere, namely Amazonia, Southern Africa and Australia (Fig. 12). The temporal correlation between OMI and the a priori model deteriorates when including year 2017, and a further degradation is seen when also including 2018 (not shown). This degradation is found not only for those large regions but also on maps of the correlation coefficient for the different time periods (not shown). The average correlation coefficient over the Southern Hemisphere (over

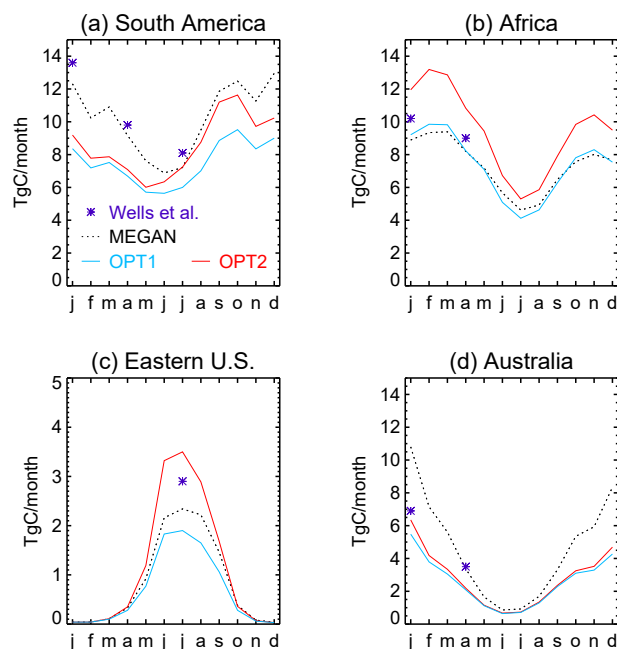



Figure 11. Seasonal variation of isoprene monthly fluxes (2013) in (a) South America (82.5–32.5 W, 32 S–14 N), (b) Africa (5–52.5 E, 36 S–14 N), (c) Southeastern U.S. (75-100 W, 26–42 N), (d) Australia (115–155 E, 10–40 S). Results shown for the a priori inventory (MEGAN, dotted line), the OMI-based optimizations (OPT1 in blue, OPT2 in red) and the CrIS-based optimization results from Wells et al. (2020) (asterisks).

$2° \times 2.5°$ pixels with mean columns higher than $5 \times 10^{15}\,\mathrm{molec.\,cm^{-2}}$) is decreased from 0.62 for 2005–2016 to 0.57 for 2005–2017 and 0.56 for 2005–2018. This deterioration is likely related to instrumental degradation and/or to changes in the version of the model (TM5) used for estimating the air mass factors. We therefore restrict the remaining discussion to the period 2005–2016.

585   The observed and modelled trends of yearly HCHO columns are positive over most areas (Fig. 12 and Fig. 13a-c). Those trends in HCHO are partly explained by emission trends (Fig. 13d-e and Fig. 14), e.g. the rapid increase of anthropogenic VOC emissions over many Asian countries ($> 1\%\,\mathrm{yr^{-1}}$ over e.g. India and China), the rapid intensification of boreal forest fires over Western Canada and Eastern Siberia (up to $\sim 4\%\,\mathrm{yr^{-1}}$), and comparatively slower changes ($< 1\%\,\mathrm{yr^{-1}}$) in biogenic VOC emissions over many areas (e.g. over Amazonia, Southern Africa and Australia) (Fig. 14). Complex patterns of changes are

590   seen over Africa, where a few regions (e.g. Nigeria and Ethiopia) experience strong increasing trends presumably associated to anthropogenic emissions, whereas other regions see significant decreases which might be partly due to declining biomass burning activity in response to socioeconomic development (e.g. over Chad, Cameroon and parts of Sahel), as reported in previous studies (Andela and van der Werf, 2014; Hickman et al., 2021).

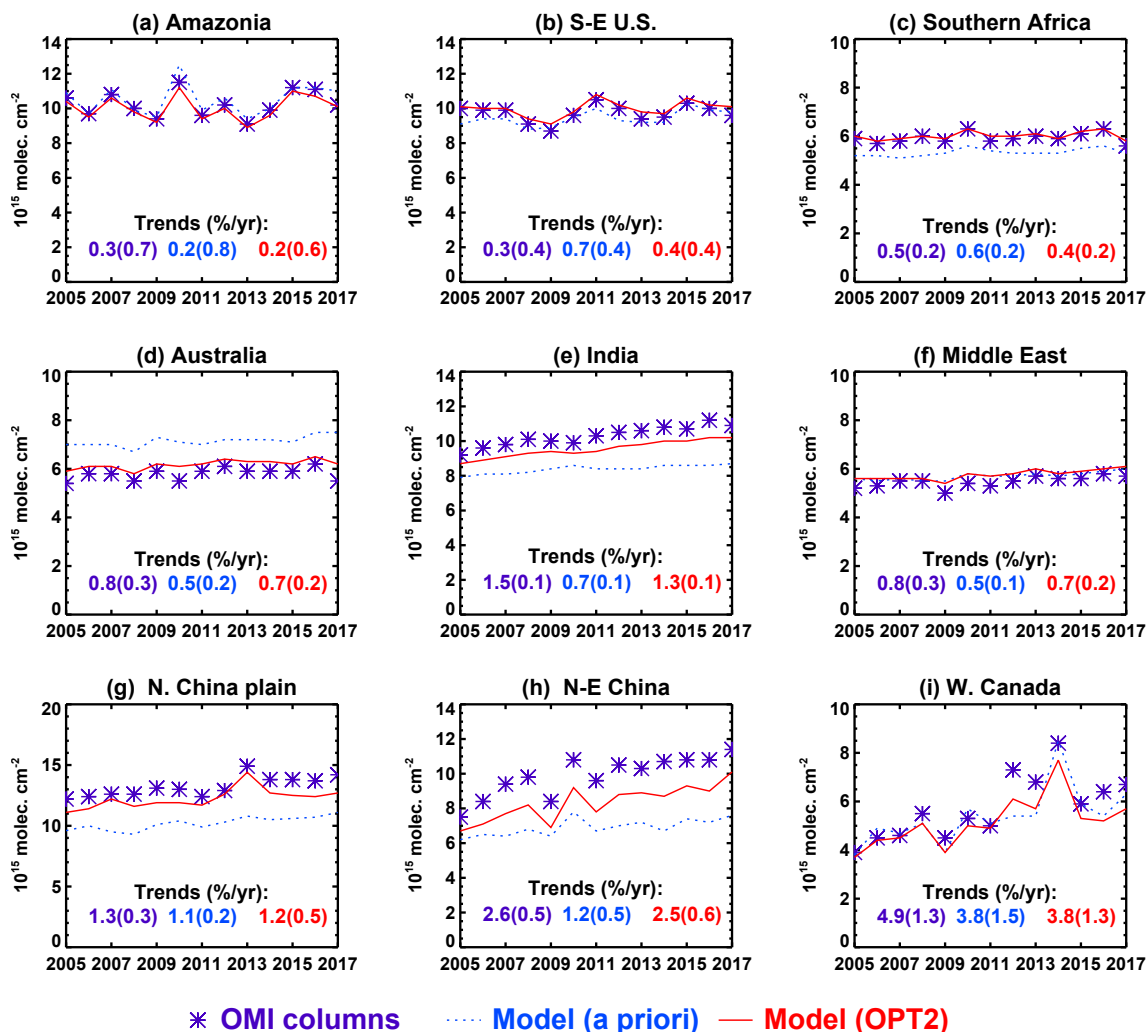



Figure 12. Annually-averaged (bias-corrected) HCHO columns (2005–2017) averaged over (a) Amazonia (40–75° W, 20° S–5° N), (b) S-E U.S. (75–100° W, 26–36° N), (c) Southern Africa (20° W–65° E, 0–40° S), (d) Australia (110–155° E, 10–38° S), (e) India (67–92° E, 10–30° N), (f) Middle East (40–55° E, 26–42° N), (g) North China plain (112–122° E, 32–40° N), (h) N-E China (122–132° E, 42–50° N), (i) W Canada (100–120° W, 54–64° N). Linear regression trends (2005–2016) are given inset, with their 1-σ uncertainty between brackets. OMI: violet symbols; a priori model in blue, optimized model (OPT2) in red.

The trends derived by the a priori model in regions with strong anthropogenic influence (India, Middle East, northern China) are underestimated, partly because the a priori anthropogenic emissions are kept constant in the model after 2012 and taken equal to their 2012 values (from EDGAR). The largest discrepancy between the model and OMI is seen over India and Northeastern China (North of 42°N, see Fig. 13c), where the observed trends are about twice larger as in the model. As a result, the top-down emission trends in those regions are among the highest globally, reaching $2.5 - 3\%$ per year over 2005–

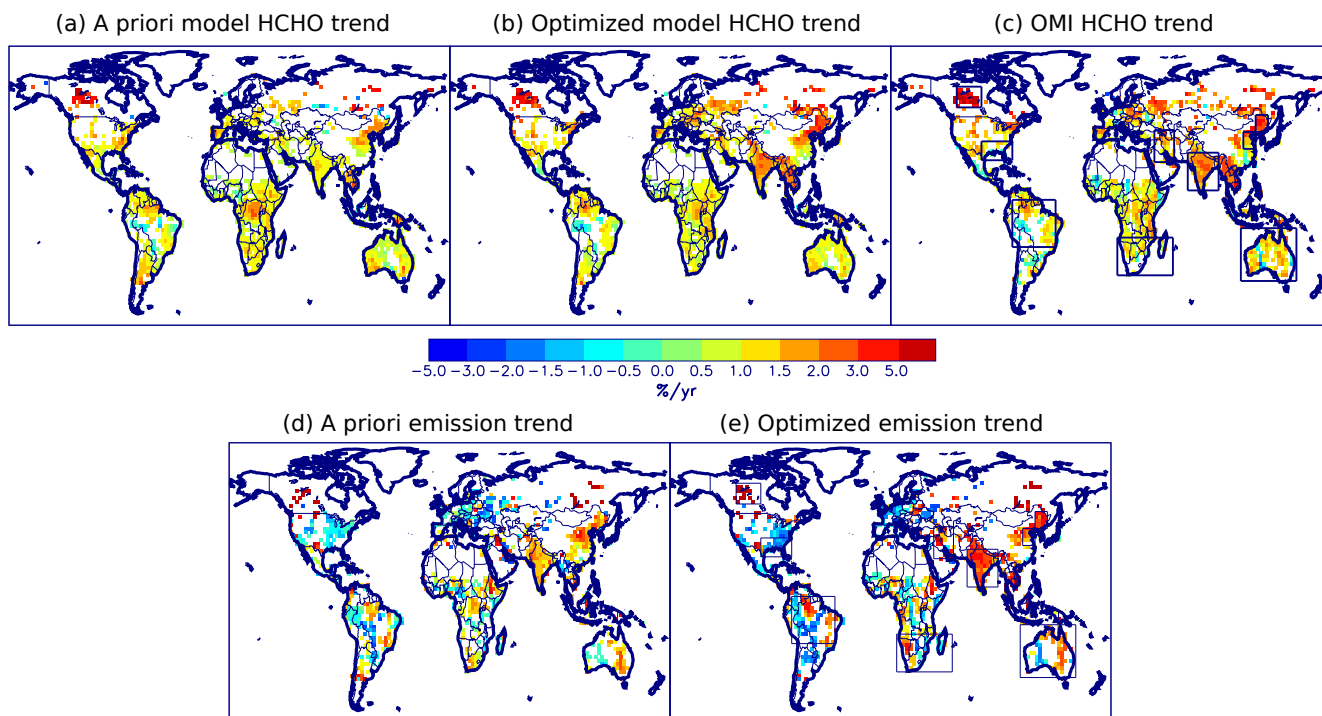



Figure 13. Trends (%/yr) over 2005–2016 of yearly-averaged HCHO columns from (a) the a priori model simulation, (b) the optimized (OPT2) simulation, (c) the bias-corrected OMI retrieval, and trends of total NMVOC emissions from (d) the a priori simulation, and (e) the OPT2 simulation. Pixels with low a priori emission (below $5 \times 10^{10}\,\mathrm{molec.\,cm^{-2}\,s^{-1}}$) or for which the trend uncertainty exceeds 150% of the trend are left as blank. The boxes in panels (c) and (e) show the regions used for calculating the temporal evolution of averaged columns (Fig. 12) and total emissions (Fig. 14).

2016 (Fig. 14). Over the last years, however, the emissions have decreased over the North China Plain (since 2013), likely due to emission regulations (Bauwens et al., 2022). Over India, the apparent stabilization of top-down emissions after 2012 seems contradicted by reports that regulatory measures were not effective in India until the last years (after 2018) (Vohra et al., 2021). More work will be needed to examine the patterns of HCHO changes and the possible causes of the discrepancy.

Over the Eastern U.S., both the a priori and optimized emission trends are generally negative, in spite of an increasing trend in yearly HCHO columns (Fig. 13). This apparent paradox is due to a strong seasonality in trends, as the summertime (May–September) HCHO trends are much lower (more negative) than the annual trends (Fig. S9–S10). For example, the summertime HCHO OMI trend over the Southeastern U.S. is $-0.4\%\,\mathrm{yr^{-1}}$, i.e. $0.7\%\,\mathrm{yr^{-1}}$ below the annual trend; a similar difference is found in the model results. The significant decreasing trend of summertime OMI column over this region, already noted in previous studies (De Smedt et al., 2015; Zhu et al., 2017; Opacka et al., 2021), is difficult to explain. Zhu et al. (2017) proposed that this trend is partly due to the decline in anthropogenic NO_x emissions caused by air quality regulations. Indeed, the yield of HCHO from VOC oxidation is believed to decrease when NO_x levels decrease. However, this effect was found to account

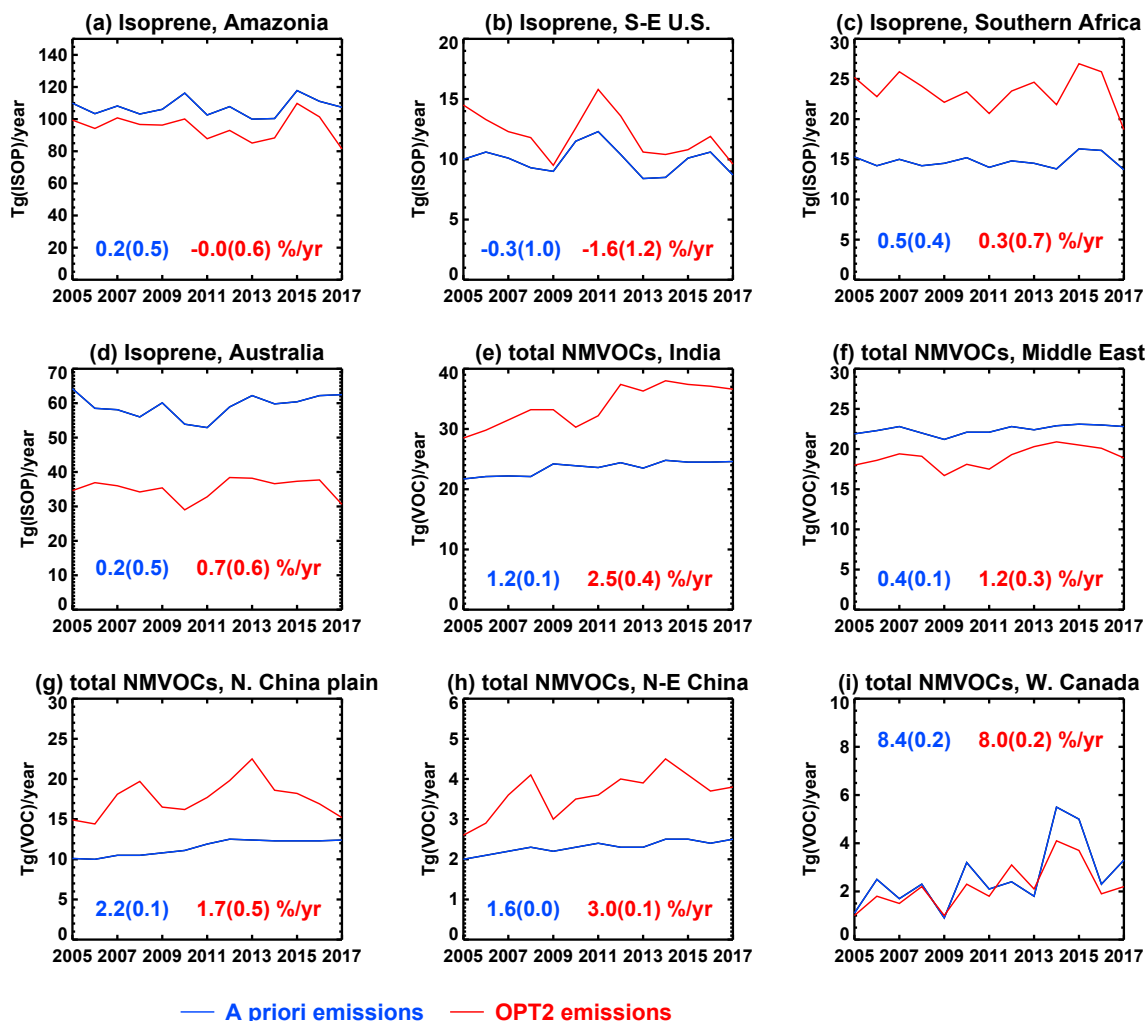



Figure 14. (a-d) Total isoprene emissions, (e-h) total VOC annual emissions ($\mathrm{Tg\,yr^{-1}}$) from the same large regions as in Fig. 12. A priori emissions in blue, optimized emissions (OPT2) in red. Linear regression trends (2005–2016) are given inset, with their 1-σ uncertainty into brackets.

for only $\sim 20\%$ of the HCHO column change seen by OMI, based on GEOS-Chem model calculations. The OMI HCHO decline is therefore more likely due to changes in emissions. Zhu et al. (2017) noted that the strong HCHO decline over the Houston-Galveston-Brazoria area likely reflects a fast decrease in anthropogenic point source VOC emissions, but the spatial extent of this effect should be limited, whereas negative trends in OMI columns (and in top-down VOC emissions) are derived

615 over a much vaster area. This suggests a trend in the emissions of isoprene, as its oxidation is the dominant source of HCHO over the Southeastern U.S. (Palmer et al., 2006). Possible drivers of such trend (besides temperature and visible radiation fluxes, already considered in the MEGAN model) include soil moisture stress, CO_2 inhibition and land-use change. The effect

of soil moisture stress is ignored in this study, given the difficulties associated to its parameterization (Bauwens et al., 2016; Opacka et al., 2021), but it could be significant in this area. The effect of CO_2 is considered in our a priori emissions using the parameterization of Possell and Hewitt (2011), which induces a emission decline of about $\sim 0.45\% \, \mathrm{yr}^{-1}$ over 2005–2016, but this effect is highly uncertain.

Using high-resolution tree cover information from the Global Forest Watch database (Hansen et al., 2013), Opacka et al. (2021) showed that a widespread decline in tree cover occurred in the Eastern U.S. over 2001–2016. The estimated trends in isoprene emissions due to land-use change between 2005 and 2016 at the model resolution ($2° \times 2.5°$), calculated according to Opacka et al. (2021), are shown on Fig. S11. The tree cover decline induces negative trends in isoprene fluxes that are generally small ($< 0.5\% \, \mathrm{yr}^{-1}$) but which might contribute to the large negative emission trend inferred from OMI data. Note that those changes due to land-use may be underestimated as they do not consider changes in species composition, e.g. the replacement of high isoprene emitters by comparatively lower isoprene emitters.

Land use change might impact other regions, especially South America. Extensive tree cover loss over a large region extending from Paraguay, Northern Argentina and Eastern Bolivia to the Brazilian States of Rondônia, Mato Grosso and Pará (Opacka et al., 2021) has caused negative trends in isoprene emissions of up to $-2\% \, \mathrm{yr}^{-1}$ (Fig. S11), which likely contribute to the patterns of negative trends inferred from OMI over this continent (Fig. 13e). A more thorough analysis at higher spatial resolution would be needed to separate those effects from other possible drivers of changes in OMI columns (e.g. biomass burning), in South America as well as on other continents.

6 Conclusions

The HCHO vertical columns of the QA4ECV OMI dataset are evaluated using two qualitatively very different measurement types: ground-based vertical columns of formaldehyde measured at 26 FTIR stations worldwide, and in situ HCHO measurements from aircraft campaigns conducted over the U.S. in 2012–2013. A regional atmospheric chemistry-transport model is used to assimilate the airborne measurements and to derive HCHO distributions that can be compared to the satellite data while closely approximating the vertical and horizontal distribution of the measurements. The two datasets are complementary: whereas the FTIR dataset covers a wide range of conditions and is geographically more diverse, the aircraft campaigns used in this work cover a known major VOC source region and are particularly relevant to the evaluation of BVOC emissions using satellite data.

The evaluation of OMI columns against FTIR column data and airborne in situ measurements leads to similar conclusions, in spite of the large qualitative difference between the datasets. The regression parameters are remarkably similar, with nearly identical slopes (0.66) and similar intercepts. On one hand, the large OMI columns (especially above $\sim 12 \times 10^{15} \, \mathrm{molec. \, cm}^{-2}$) are generally underestimated, by a factor of 1.29 (against aircraft data) to 1.36 (against FTIR) for an OMI column of $20 \times 10^{15} \, \mathrm{molec. \, cm}^{-2}$. On the other hand, low OMI columns (especially below $\sim 6 \times 10^{15} \, \mathrm{molec. \, cm}^{-2}$) are often underestimated, by -10% and -37% against FTIR- and aircraft-based regressions, respectively, for an OMI column of $5 \times 10^{15} \, \mathrm{molec. \, cm}^{-2}$.

650   The correlation of FTIR vs. co-located OMI data is relatively low (Pearson's coefficient of 0.67), but the regression is qualitatively consistent with the (much better-correlated) evaluation of TROPOMI HCHO column measurements against FTIR data (Vigouroux et al., 2020). To summarize, both aircraft and FTIR data suggest a significant underestimation (overestimation) of satellite data for high (low) columns, but the discrepancy is slightly lower according to aircraft data. Clearly, more work will be needed to refine those estimates, and especially to elucidate and potentially eliminate the possible causes for disagreement.

655   Ideally, a co-located cross-evaluation of FTIR columns and aircraft profiles (e.g. spiral flights) would be needed to assess their mutual consistency.

The OMI bias against FTIR and aircraft data is tentatively corrected through a linear relationship, and the bias-corrected HCHO columns are used as constraints to optimize the emissions of VOCs in the MAGRITTE global model over 2005–2017. Evaluation of the modelled distributions of HCHO against 11 aircraft datasets spanning 2006–2016 show that the bias cor-

660   rection leads to a considerable reduction of the average bias against aircraft observations, from -27% in the prior simulation and -28% in an optimization without bias correction to -9% when using bias-corrected data. The root mean square deviation (RMSD) is also decreased. Agreement with FTIR data is also improved, especially at South American sites, but the improvement is marginal at mid-latitudes, where the relative uncertainties of the OMI columns are high and the bias correction has a very small effect on the OMI columns.

665   The optimized VOC emissions of the optimization using bias-corrected data (OPT2) are only slightly higher ($+5\%$ globally) than the a priori emissions, but are much higher than previous top-down estimates based on OMI data. In particular, the global top-down isoprene emissions of the OPT2 run ($445\,\mathrm{Tgyr}^{-1}$) are much higher than those derived by a previous inverse modelling study based on OMI HCHO columns. Regionally (e.g. over Africa), the difference exceeds a factor of 2. The major reason for the increased top-down emissions is the much higher HCHO columns used in this work, in large part due to the bias

670   correction. This result demonstrates the importance of validation for any reliable quantitative use of satellite HCHO data.

The biogenic emission updates of the OPT2 inversion bring the model closer to top-down isoprene emissions derived from the spaceborne (CrIS) isoprene columns (Wells et al., 2020) in many regions, including Central and Southern Africa, the Eastern U.S. and parts of South America. The comparison also highlights important differences, which are partly attributed to known weaknesses of the HCHO-based inversion, such as the high uncertainties of low HCHO columns (e.g. over Europe and

675   over Amazonia during the wet season) and the co-occurrence of very extensive vegetation fires (e.g. over the Central African Republic in the dry season). Furthermore, spaceborne isoprene data have their own uncertainties, and the top-down emissions from CrIS are very dependent on OH radical levels, which remain uncertain in remote areas e.g. due to their strong sensitivity to NO_x levels and to other factors affecting the lifetime of isoprene. The interannual variability of HCHO columns is generally well explained over regions where biogenic emissions and biomass burning are the dominant source of VOCs. Large positive

680   emission trends of up to $2-3\%\,\mathrm{yr}^{-1}$ are suggested by OMI data over India and Northern China over 2005–2016, most likely due to rising anthropogenic emissions. Over the last years, however, emissions appear to stabilize over India, and even decline over the North China Plain, as result of emission regulations. Over Southeastern U.S., the observed decline of summertime OMI columns might have multiple causes, including the decline of NO_x levels (and therefore of the HCHO yield from VOC

oxidation) and a decline of isoprene emissions due to CO_2 increase and land use change, both of which potentially accounting

685   for $\sim -0.4\,\% \,\mathrm{yr}^{-1}$ to the isoprene emission trend according to current estimates.

Data availability. The QA4ECV dataset can be found at http://doi.org/10.18758/71021031 (De Smedt et al., 2017). The NASA aircraft campaign datasets are available from the Langley Research Center at https://www-air.larc.nasa.gov/missions/merges. The HCHO FTIR data can be requested from the PIs of each station. The MEGAN-MOHYCAN isoprene inventory is available at https://emissions.aeronomie.be.

Author contributions. JFM coordinated the study and wrote the manuscript. JFM and TS designed and prepared the inversions. JFM and TS

690   carried out the analysis, with help from GMO, BO and AG on specific aspects. IDS described the TROPOMI HCHO data. CV provided the description of the FTIR dataset. CV and BL conducted the OMI validation and bias-correction using FTIR data. BL, CV, CA, MG, JH, FH, RK, EL, EM, MM, JMM, IMo, IMu, TN, JN, IO, MP, AR, WS, KS, RS, and YT provided the FTIR measurements. AF provided the CAMS, TDLAS and DFGAS in situ measurements. All authors read and commented on the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

695   *Acknowledgements.* HCHO satellite data from OMI were produced in the scope of the European FP7 project QA4ECV (grant no. 6007405). We thank Dan Smale and the National Institute of Water and Atmospheric Research (NIWA) for the provision of the Lauder FTIR data, and Nicholas Jones of the Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, Australia, for provision of Wollongong FTIR data. Thanks to Thomas Hanisco and the NASA Goddard Space Flight Center (GSFC) for provision of ISAF data (SENEX). B. O. was supported by the EQUATOR (Emission Quantification of Atmospheric tracers in the Tropics using ObseRvations from satellites, contract

700   no. B2/202/P1/EQUATOR, 2021-2025) project of the BRAIN-be 2.0 programme from the Belgian Science Policy Office (Belspo). G.-M. O. was supported by the SEEDS project funded by the European Commission under the H2020 programme (grant agreement no. 101004318). C. V. was supported by Belspo through the ProDEx project TROVA-E2 (TROPOMI Validation - Phase E2) of the European Space Agency (2020-2023). E. M. is a senior research associate with the F.R.S. - FNRS.

     The Paris site has received funding from Sorbonne Université, the French research center CNRS, the French space agency CNES, and

705   Région Île-de-France. FTIR operations of Rikubetsu and Tsukuba are supported in part by the GOSAT series project. The Rikubetsu NDACC site is funded by the joint research program of the Institute for Space-Earth Environmental Research (ISEE), Nagoya University. This work has been supported by the BMBF (German Ministry of Research and Education) in the project ROMIC-II, subproject TroStra (01LG1904A). We thank the AWI Bremerhaven, Germany and the Meteorological Service Suriname for logistical and the senate of Bremen for logistical and financial support. We acknowledge the support of the station personnel at the AWIPEV research base in Ny-Ålesund, Spitsbergen and

710   C. Becker for support in Paramaribo, Suriname.

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, https://doi.org/10.1029/2003GL017336, 2003.

715 Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D. and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, Atmos. Chem. Phys., 11, 4039–4072, https://doi.org/10.5194/acp-11-4039-2011, 2011.

Andela, N., and van der Werf, G.: Recent trends in African fires driven by cropland expansion and El Niño to La Niña transition, Nat. Clim. Chang., 4, 791–795, https://doi.org/10.1038/nclimate2313, 2014.

720 Andreae, M. O.: Emission of trace gases and aerosols from biomass burning – An updated assessment, Atmos. Chem. Phys., 19, 8523–8546, https://doi.org/10.5194/acp-19-8523-2019, 2019.

Archibald, A. T., Neu, J. L., Elshorbany, Y. F., Copper, O. R., Young, P. J., Akiyoshi, H., Cox, R. A., Coyle, M., Derwent, R. G., Deushi, M., Finco, A., Frost, G. J., Galbally, I. E., Gerosa, G., Granier, C., Griffths, P. T., Hossaini, R., Hu, L., Jöckel, P., Josse, B., Lin, M. Y., Mertens, M., Morgenstern, O., Naja, M., Naik, V., Oltmans, S., Plummer, D. A., Revell, L. E., Saiz-Lopez, A., Saxena, P., Shin, Y.

725 M., Shahid, I., Shallcross, D., Tilmes, S., Trickl, T., Wallington, T. J., Wang, T., Worden, H. M., and Zeng, G.: Tropospheric Ozone Assessment Report: A critical review of changes in the tropospheric ozone burden and budget from 1850 to 2100, Elem. Sci. Anth., 8:1, 1–53, https://doi.org/10.1525/elementa.2020.034, 2020.

Arneth, A., Schurgers, G., Lathière, J., Duhl, T., Beerling, D. J., Hewitt, C. N., Martin, M., and Guenther, A.: Global terrestrial isoprene emission models: sensitivity to variability in climate and vegetation, Atmos. Chem. Phys., 11, 8037–8052, https://doi.org/10.5194/acp-11-8037-2011, 2011.

730

Atkinson, R.: Atmospheric chemistry of VOCs and NOx, Atmos. Environ., 34, 2063–2101, https://doi.org/10.1016/S1352-2310(99)00460-, 2000.

Barkley, M. P., Palmer, P. I., Ganzeveld, L., Arneth, A., Hagberg, D., Karl, T., Guenther, A., Paulot, F., Wennberg, P. O., Mao, J., Kurosu, T. P., Chance, K., Müller, J.-F., De Smedt, I., Van Roozendael, M., Chen, D., Wang, Y. and Yantosca, R.: Can a "state of the art" chemistry

735 transport model simulate Amazonian tropospheric chemistry? J. Geophys. Res., 116, D16302, https://doi.org/10.1029/2011JD015893, 2011.

Barth, M. C., Cantrell, C. A., Brune, W. H., Rutledge, S. A. Crawford, J. H., Huntrieser, H., Carey, L. D., McGorman, D., Weisman, M., Pickering, K. E., Bruning, E.,,erson, B., Apel, E., Biggerstaff, M., Campos, T., Campuzano-Jost, P., Cohen, R., Crounse, J., Day, D. A., Diskin, G., Flocke, F., Fried, A., Garland, C., Heikes, B., Honomichl, S., Hornbrook, R., Huey, L. G., Jimenez, J. L., Lang, T., Lichtenstein,

740 M., Mikoviny, T., Nault, B., O'Sullivan, D., Pan, L. L., Peischl, J., Pollack, I., Richter, D., Riemer, D., Ryerson, T., Schlager, H., St. Clair, J., Walega, J., Weibring, P., Weinheimer, A., Wennberg, P., WIsthaler, A., Wooldridge, P. J. and Ziegler, C.: The Deep Convective Clouds and Chemistry (DC3) Field Campaign, Bull. Amer. Meteor. Soc., 96, 1281–1309, https://doi.org/10.1175/BAMS-D-13-00290.1, 2015.

Bates, K. H., and Jacob, D. J.: A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol, Atmos. Chem. Phys., 19, 9613–9640, https://doi.org/10.5194/acp-19-9613-2019, 2019.

745 Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Werf, G. R. van der, Wiedinmyer, C., Kaiser, J. W., Sindelarova, K., and Guenther, A.: Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations, Atmos. Chem. Phys., 16(15), 10133–10158, https://doi.org/10.5194/acp-16-10133-2016, 2016.

Bauwens, M., Stavrakou, T., Müller, J.-F., Van Scheybroeck, B., De Cruz, L., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Laffineur, Q., Amelynck, C., Schoon, N., Heinesch, B., Holst, T., Arneth, A., Ceulemans, R., Sanchez-Lorenzo, A., and Guenther, A.: Recent past (1979-750 2014) and future (2070-2099) isoprene fluxes over Europe simulated with the MEGAN-MOHYCAN model, Biogeosci., 15, 3673–3690, https://doi.org/10.5194/bg-2017-532, 2018.

Bauwens, M., Verreyken, B., Stavrakou, T., Müller, J.-F., and De Smedt, I.: Spaceborne evidence for significant anthropogenic VOC trends in Asian cities over 2005-2019, Environ. Res. Lett., 17, 015008, https://doi.org/10.1088/1748-9326/ac46eb, 2022.

Berndt, T., Huttinen, N., Herrmann, H., and Hansel, A.: First oxidation products from the reaction of hydroxyl radicals with isoprene for 755 pristine environmental conditions, Commun. Chem., 2:21, https://doi.org/10.1038/s42004-019-0120-9, 2019.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J. and Schultz, M.: Global model of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23073–23095, https://doi.org/10.1029/2010JD014870, 2001.

Boeke, N. L., Marshall, J. D., Alvarez, S., Chance, K. V., Fried, A., Kurosu, T. P., Rappenglück, B., Richter, D., Walega, J., Weibring, P., 760 and Millet, D. B.: Formaldehyde columns from the Ozone Monitoring Instrument: Urban versus background levels and evaluation using aircraft data and a global model, J. Geophys. Res., 116, D05303, https://doi.org/10.1029/2010JD014870, 2011.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res., 109, https://doi.org/10.1029/2003JD003962, 2004.

Cazorla, M., Wolfe, G. M., Bailey, S. A., Swanson, A. K., Arkinson, H. L., and Hanisco, T. F.: A new airborne laser-induced fluorescence 765 instrument for in situ detection of formaldehyde throughout the troposphere and lower stratosphere, Atmos. Meas. Tech., 8, 541–552, https://doi.org/10.5194/amt-8-541-2015, 2015.

Chameides, W. L., Lindsay, R. W., Richardson, J., and Kiang, C. S.: The role of biogenic hydrocarbons in urban photochemical smog: Atlanta as a case study, Science, 241, 1473–1475, https://doi.org/10.1126/science.3420404, 1988.

Chen, X., Millet, D. B., Singh, H. B., Wisthaler, A., Apel, E. C., Atlas, E. L., Blake, D. R., Brown, S. S., Crounse, J. D., de Gouw, J. A., 770 Flocke, F., Fried, A., Heikes, B. G., Hornbrook, R. S., Mikoviny, T., Min, K.-E., Müller, M., Neuman, J. A., O'Sullivan, D. W., Peischl, J., Pfister, G. G., Richter, D., Roberts, J. M., Ryerson, T. B., Shertz, S., Treadaway, V., Veres, P. R., Walega, J., Warneke, C., Washenfelder, R. A., Weibring, P., and Yuan, B.: On the sources and sinks of atmospheric VOCs: An integrated analysis of recent aircraft campaigns over North America, Atmos. Chem. Phys., 19, 9097-9123, https://doi.org/10.5194/acp-19-9097-2019, 2019.

Crawford, J. H., and Pickering, K. E.: DISCOVER-AQ: Advancing strategies for air quality observations in the next decade, Environ. 775 Manage, Air and Waste Management Association's Magazine for Environmental Managers, 4, Pittsburgh, Pa, Air and Waste Management Association, 2014.

De Mazière, M., Thompson, A. M., Kurylo, M. J., Wild, J. D., Bernhard, G., Blumenstock, T., Braathen, G. O., Hannigan, J. W., Lambert, J.-C., Leblanc, T., McGee, T. J., Nedoluha, G., Petropavlovskikh, I., Seckmeyer, G., Simon, P. C., Steinbrecht, W., and Strahan, S. E.: The Network for the Detection of Atmospheric Composition Change (NDACC): history, status and perspectives, Atmos. Chem. Phys., 18, 780 4935-4964, https://doi.org/10.5194/acp-18-4935-2018, 2018.

De Smedt, I., Mûller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formalde-hyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 4947–4963, https://doi.org/10.5194/acp-8-4947-2008, 2008.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, 785 C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendael, M.: Diurnal, seasonal and long-term variations of global formaldehyde

columns inferred from combined OMI and GOME-2 observations, Atmos. Chem. Phys., 15, 12519–12545, https://doi.org/10.5194/acp-15-12519-2015, 2015.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395–2426, https://doi.org/10.5194/amt-11-2395-2018, 2018.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561–12593, https://doi.org/10.5194/acp-21-12561-2021, 2021.

Di Carlo, P., Brune, W. H., Martinez, M., Harder, H., Lesher, R., Ren, X. R., Thornberry, T., Carroll, M. A., Young, V., Shepson, P. B., Riemer, D., Apel, E., and Campbell, C.: Missing OH reactivity in a forest: Evidence for unknown reactive biogenic VOCs, Science, 304, 722–725, https://doi.org/10.1126/Science.1094392, 2004.

Emmerson, K. M., Galbally, I. E., Guenther, A. B., Paton-Walsh, C., Guerette, E.-A., Cope, M. E., Keywood, M. D., Lawson, S.J., Molloy, S. B., Dunne, E., Thatcher, M., Karl, T., and Malekinia, S. D.: Current estimates of biogenic from Eucalypts uncertain for Southeast Australia,

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

Franco, B., Mahieu, E., Emmons, L. K., Tzompa-Sosa, Z. A., Fischer, E. V., Sudo, K., Bovy, B., Conway, S., Griffin, D., Hanningan, J. W., Strong, K., and Walker, K. A.: Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in North America, Environ. Res. Lett., 11, 044010, http://dx.doi.org/10.1088/1748-9326/11/4/044010, 2016.

Fried, A., Walega, J. G., Olson, J. R., Crawford, J. H., Chen , G., Weibring, P., Richter, D., Roller, C., Tittel, F. K., Heikes, B. G., Snow, J. A., Shen, H., O'Sullivan, D. W., Porter, M., Fuelberg, H., Halland, J. and Millet, D. B.: Formaldehyde over North America and the North Atlantic during the summer 2004 INTEX campaign: methods, observed distributions, and measurement-model comparisons, J. Geophys. Res., 113, D10302, https://doi.org/10.1029/2007JD009185, 2008.

Fried, A., Cantrell, C., Olson, J., Crawford, J. H., Weibring, P., Walega, J., Richter, D., Junkermann, W., Volkamer, R., Sinreich, R., Heikes, B. G., O'Sullivan, D., Blake, D. R., Blake, N., Meinardi, S., Apel, E., Weinheimer, A., Knapp, D., Perring, A., Cohen, R. C., Fuelberg, H., Shetter, R. E., Hall, S. R., Ullmann, K., Brune, W. H., Mao, J., Ren, X., Huey, L. G., Singh, H. B., Hair, J. W., Riemer, D., Diskin, G., and Sachse, G.: Detailed comparisons of airborne formaldehyde measurements with box models during the 2006 INTEX-B and MILAGRO campaigns: potential evidence for significant impacts of unmeasured and multi-generation volatile organic carbon compounds, Atmos. Chem. Phys., 11, 11867–11894, https://doi.org/10.5194/acp-11-11867-2011, 2011.

Fried, A., Walega, J., Weibring, P., Richter, D., Simpson, I. J., Blake, D. R., Blake, N. J., Meinardi, S., Barletta, B., Hughes, S. C., Crawford, J. H., Diskin, G., Barrick, J., Hair, J., Fenn, M., Wisthaler, A., Mikoviny, T., Woo, J.-H., Park, M., Kim, J., Min, K.-E., Jeong, S., Wennberg, P. O., Kim, M. J., Crounse, J. D., Teng, A. P., Bennett, R., Yang-Martin, M., Shook, M. A., Huey, G., Tanner, D., Knote, C., Kim, J., Park, R. and Brune, W.: Airborne formaldehyde and volatile organic compound measurements over the Daesan petrochemical complex on Korea2019s northwest coast during the Korea-United States Air Quality study: Estimation of emission fluxes and effects on air quality, Elem. Sci. Anth., 8: 1, https://doi.org/10.1525/elementa.2020.121, 2020.

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, https://doi.org/10.1029/2006JD007853, 2007.

Gu, D., Guenther, A. B., Shilling, J. E., Yu, H., Huang, M., Zhao, C., Yang, Q., Martin, S. T., Artaxo, P., Kim, S., Seco, R., Stavrakou, T., Longo, K. M., Tóta, J., de Souza, R. A. F., Vega, O., Liu, Y., Shrivastava, M., Alves, E. G., Santos, F. C., Leng, G., and Hu, Z.: Airborne observations reveal elevational gradient in tropical forest isoprene emissions, Nature Commun, 8, 15541, https://doi.org/10.1038/ncomms15541, 2017.

Gu, S., Guenther, A., and Faiola, C.: Effects of anthropogenic and biogenic volatile organic compounds on Los Angeles air quality, Environ. Sci. Technol., 55, 12191–12201, https://doi.org/10.1021/acs.est.1c01481, 2021.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., McKay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J. and Zimmerman, P.: A global model of natural volatile organic compound emissions

Fried, A., Barth, M. C., Bela, M., Weibring, P., Richter, D., Walega, J., Li, Y., Pickering, K., Apel, E., Hornbrook, R., Hills, A., Riemer, D. D., Blake, N., Blake, D. R., Schroeder, J. R., Luo, Z. J., Crawford, J. H., Olson, J., Rutledge, S., Betten, D., Biggerstaff, M. I., Diskin, G. S., Sachse, G., Campos, T., Flocke, F., Weinheimer, A., Cantrell, C., Pollack, I., Peischl, J., Froyd, K., Wisthaler, A., Mikoviny, T. and Woods, S.: Convective transport of formaldehyde to the upper troposphere and lower stratosphere and associated scavenging in thunderstorms over the central United States during the 2012 DC3 study, J. Geophys. Res. Atmos., 121, 7430–7460, https://doi.org/10.1002/2015JD024477, 2016.

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(11), 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5(6), 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., and Kommareddy, A.: High-resolution global maps of 21st-century forest cover change, Science, 342, 850–853, https://doi.org/10.1126/science.1244693, 2013.

Hase, F., Demoulin, P., Sauval, A. J., Toon, G. C., Bernath, P. F., Goldman, A., Hannigan, J. W., and Rinsland, C. P.: An empirical line-by-line model for the infrared solar transmittance spectrum from 700 to 5000 cm^{-1}, J. Quant. Spectrosc. Radiat. Transf., 102, 450–463, https://doi.org/10.1016/j.jqsrt.2006.02.026, 2006.

Hesbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biatavi, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hickman, J. E., Andela, N., Tsigaridis, K., Galy-Lacaux, C., Ossohou, M., and Bauer, S. E.: Reductions in NO_2 burden over north equatorial Africa from decline in biomass burning in spite of growing fossil fuel use, 2005 to 2017, Proc. Acad. Nat. Soc., 118, e2002579118, https://doi.org/10.1073/pnas.2002579118, 2021.

Houweling, S., Dentener, F., and Lelieveld, J.: The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry, J. Geophys. Res., 103, 10673–10696, https://doi.org/10.1029/97JD03582, 1998.

Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean, M., Schaaf, E., and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970-2012, Atmos. Chem. Phys., 17, 7683–7701, https://doi.org/10.5194/acp-17-7683-2017, 2017.

Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, Atmos. Chem. Phys., 10, 5191–5212, https://doi.org/10.5194/acp-10-5191-2010, 2010.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmos. Chem. Phys., 15, 11411–11432, https://doi.org/10.5194/acp-15-11411-2015, 2015.

Jones, N. B., Riedel, K., Allan, W., Wood, S., Palmer, P. I., Chance, K., and Notholt, J.: Long-term tropospheric formaldehyde concentrations deduced from ground-based fourier transform solar infrared measurements, Atmos. Chem. Phys., 9, 7131–7142, https://doi.org/10.5194/acp-9-7131-2009, 2009.

Karl, T., Striednig, M., Graus, M., Hammerle, A., and Wohlfahrt, G.: Urban flux measurements reveal a large pool of oxygenated volatile organic compound emissions, Proc. Nat. Acad. Sci., 115, 1186–1191, https://doi.org/10.1073/pnas.1714715115, 2018.

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., 113, D18308, https://doi.org/10.1029/2008JD010290, 2008.

Kwon, H.-A., González Abad, G., Nowlan, C. R., Chong, H., Souri, A. H., Vigouroux, C., Röhling, A., Kivi, R., Makarova, M., Notholt, J., Palm, M., Winkler, H., Té, Y., Sussmann, R., Rettinger, M., Mahieu, E., Strong, K., Lutsch, E., Yamanouchi, S., Nagahama, T., Hanninan, J. W., Zhou, M., Murata, I., Grutter, M., Stremme, W., De Mazière, M., Jones, N., Smale, D., Morino, I.: Validation of OMPS Suomi NPP and OMPS NOAA-20 formaldehyde total columns with NDACC FTIR observations, Earth Space Sci., 10, e2022EA002778, https://doi.org/10.1029/2022EA002778, 2023.

Lawson, S. J., Selleck, P. W., Galbally, I. E., Keywood, M. D., Harvey, M. J., Lerot, C., Helmig, D. and Ristovski, Z.: Seasonal in situ observations of glyoxal and methylglyoxal over the temperate oceans of the Southern Hemisphere, Atmos. Chem. Phys., 15, 223–240, https://doi.org/10.5194/acp-15-223-2015, 2015.

Lelieveld, J., T. Butler, T. M., Crowley, J. N., Dillon, T. J., Fischer, H., Ganzeveld, L., Harder , H., Lawrence, M. G., Martinez, M., Taraborrelli, D., and Williams, J.: Atmospheric oxidation capacity sustained by a tropical forest, Nature, 452, 737–740, https://doi.org/10.1038/nature06870, 2008.

Levelt, P. F., van den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O., and Saari, H.: The ozone monitoring instrument, IEEE T. Geosci. Remote. Sens., 44, 1093–1101, 2006.

Madronich, S. and Flocke, S.: The role of solar radiation in atmospheric chemistry, in: Handbook of Environmental Chemistry, edited by: Boule, P., Springer Verlag, Heidelberg, 1–26, 1998.

Marais, E. A., Jacob, D. J., Guenther, A., Chance, K., Kurosu, T. P., Murphy, J. G., Reeves, C. E., and Pye, H. O. T.: Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter, Atmos. Chem. Phys., 14, 7693–7703, https://doi.org/10.5194/acp-14-7693-2014, 2014.

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

Martin, R. V., Parrish, D. D., Ryerson, T. B., Nicks Jr., T. B., Chance, K., Kurosu, T. P., Jacob, D. J., Sturges, E. D., Fried, A. and Wert, B. P.:

Martinez, M., Harder, H., Kovacs, T. A., Simpas, J. B., Bassis, J., Lesher, R., Brune, W. H., Frost, G. J., Williams, E. J., Stroud, C. A., Jobson,
905 B. T., Roberts, J. M., Hall, S. R., Shetter, R. E., Wert, B., Fried, A., Alicke, B., Stutz,J., Young, V. L., White, A. B., and Zamora, R. J.: OH and HO$_2$ concentrations, sources, and loss rates during the Southern Oxidants Study in Nashville, Tennessee, summer 1999, 108, 4617, https://doi.org/10.1029/2003JD003551, 2003.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts,
910 J. M., Tyerson, T. B., and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science, 359, 760–764, https://doi.org/10.1126/science.aaq0524, 2018.

Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross-sections of formaldehyde between 223 and 323 K in the wavelength range 225-375 nm, J. Geophys. Res., 105, 7089–7101, https://doi.org/10.1029/1999JD901074, 2000.

Miao, R., Chen, Q., Shrivastava, M., Chen, Y., Zhang, L., Hu, J., Zheng, Y. and Liao, K.: Process-based and observation-constrained SOA
915 simulations in China: The role of semivolatile and intermediate-volatility organic compounds and OH levels, Atmos. Chem. Phys., 21, 16183–16201, https://doi.org/10.5194/acp-21-16183-2021, 2021.

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B.,,erso, B. E. and Clarke, A. D.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission, J. Geophys. Res., 111, D24S02, https://doi.org/10.1029/2005JD006853, 2006.

920 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., 113, D02307, https://doi.org/10.1029/2007JD008950, 2008.

Millet, D. B., Guenther, A., Siegel, D. A., Nelson, N. B., Singh, H. B., de Gouw, J. A., Warneke, C., Williams, J., Eerdekens, G., Sinha, V., Karl, T., Flocke, F., Apel, E. C., Riemer, D. D., Palmer, P. I., and Barkley, M.: Global atmospheric budget of acetaldehyde: 3-D model
925 analysis and constraints from in-situ and satellite observations, Atmos. Chem. Phys., 10, 3405–3425, https://doi.org/10.5194/acp-10-3405-2010, 2010.

Misztal, P. K., Karl, T., Weber, R., Jonsson, H. H., Guenther, A. B. and Goldstein, A. H.: Airborne flux measurements of biogenic isoprene over California, Atmos. Chem. Phys., 14, 10631–10647, https://doi.org/10.5194/acp-10631-2014, 2014.

Molina, L. T., Madronich, S., Gaffney, J. S., Apel, E., de Foy, B., Fast, J., Ferrare, R., Herndon, S., Jimenez, J. L., Lamb, B., Osornio-Vargas,
930 A. R., Russell, P., Schauer, J. J., Stevens, P. S., Volkamer, R., and Zavala, M.: An overview of the MILAGRO 2006 campaign: Mexico City emissions and their transport and transformation, Atmos. Chem. Phys., 10, 8697–8760, doi:10.5194/acp-10-8697-2010, 2010.

Müller, J.-F., and Brasseur, G.: IMAGES: A three-dimensional chemical transport model of the global troposphere, J. Geophys. Res., 100, 16,445–16,490, https://doi.org/10.1029/94JD03254, 1995.

Müller, J.-F. and Stavrakou, T.: Inversion of CO and NOx emissions using the adjoint of the IMAGES model, Atmos. Chem. Phys., 5,
935 1157-1186, https://doi.org/10.5194/acp-5-1157-2005, 2005.

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

Müller, J.-F., Stavrakou, T., Bauwens, M., Compernolle, S., and Peeters, J.: Chemistry and deposition in the Model of Atmospheric compo-
940   sition at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE). Part B. Dry deposition, Geophys. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-317, 2018a.

Müller, J.-F., Stavrakou, T., Bauwens, M., George, M., Hurtmans, D., Coheur, P-F., Clerbaux, C., and Sweeney, C.: Top-down CO emissions based on IASI observations and hemispheric constraints on OH levels, Geophys. Res. Lett., 45, 1621–1629, https://doi.org/10.1002/2017GL076697, 2018b.

945   Müller, J.-F., Stavrakou, T., and Peeters, J.: Chemistry and deposition in the Model of Atmospheric composition at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE v1.1). Part A. Chemical mechanism, Geosci. Model Dev., 12, 2307–2356, https://doi.org/10.5194/gmd-12-2307-2019, 2019.

Nölscher, A. C., Yanez-Serrano, A. M., Wolff, S., Carioca de Araujo, A., Lavric, J. V., Kesselmeier, J., and Williams, J.: Unexpected sea-
sonality in quantity and composition of Amazon rainforest air reactivity, Nature Comm., 7, 10383, https://doi.org/10.1038/ncomms10383,
950   2016.

Oomen, G.-M., Müller, J.-F., Stavrakou, T., De Smedt, I., Blumenstock, T., Kivi, R., Makarova, M., Palm, M., Röhling, A., Té, Y., Vigouroux, C., Friedrich, M. M., Friess, U., Hendrick, F., Merlaud, A., Piters, A., Richter, A., Van Roozendael, M., and Wagner, T.: Weekly-derived top-down VOC fluxes over Europe from TROPOMI HCHO data in 2018–2021, EGUsphere, https://doi.org/10.5194/egusphere-2023-1972, 2023.

955   Novelli, A., Vereecken, L., Dorn, H.-P., Hofzumahaus, A., Holland, F., Yu, Z., Rosanka, S., Reimer, D., Gkatzelis, G. I., Bohn, B., Taraborrelli, D., Rohrer, F., Tillmann, R., Wegener, R., Kiendler-Scharr, A., Wahner, A., and Fuchs, H.: Importance of isomerization reactions for the OH radical regeneration from the photo-oxidation of isoprene investigated in the atmospheric simulation chamber SAPHIR, Atmos. Chem. Phys., 20, 3333–3355, https://doi.org/10.5194/acp-20-3333-2020, 2020.

Opacka, B., Müller, J.-F., Stavrakou, T., Bauwens, M., Sindelarova, K., Doubalová, J., and Guenther, A.: Global and regional impacts of
960   land cover changes on isoprene emissions derived from spaceborne data and the MEGAN model, Atmos. Chem. Phys., 21, 8413–8436, https://doi.org/10.5194/acp-21-8413-2021, 2021.

Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K. and Kurosu, T. P.: Mapping isoprene emissions over North America using formaldehyde column observations from space, J. Geophys. Res., 108, 4180, https://doi.org/10.1029/2002JD002153, 2003.

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.,
965   Pressley, S. N., Lamb, B. and Sumner, A. L.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, J. Geophys. Res., 111, D12315, https://do.org/:10.1029/2005JD006689, 2006.

Pan, X., Ichoku, C., Chin, M., Bian, H., Darmenov, A., Colarco, P., Ellison, L., Kucesra, T., da Silva, A., Wang, J., Oda, T., and Cui, G.: Six global biomass burning emission datasets: intercomparison and application in one global aerosol model, Atmos. Chem. Phys., 20, 969–994, https://doi.org/10.5194/acp-20-969-2020, 2020.

970   Paulot, F., Crounse, J. D., Kjaergaard, H. G., Kürten, A., Clair, J. M. St. Clair, Seinfeld, J. H., and Wennberg, P. O.: Unexpected epoxide formation in the gas-phase photooxidation of isoprene, Science, 325, 730-733, https://doi.org/10.1126/science.1172910, 2009.

Peeters, J., Müller, J.-F., Stavrakou, T., and Nguyen, S. V.: Hydroxyl radical recycling in isoprene oxidation driven by hydrogen bonding and hydrogen tunneling : the upgraded LIM1 mechanism, J. Phys. Chem. A, 118, 8625–8643, https://doi.org/10.1021/jp5033146, 2014.

Pétron, G., Karion, A., Sweeney, C., Miller, B. R., Montzka, S. A., Frost, G. J., Trainer, M., Tans, P., Andrews, A., Kofler, J., Helmig, D., Guenther, D., Dlugokencky, E., Lang, P., Newberger, T., Wolter, S., Hall, B., Novelli, P., Brewer, A., Conley, S., Hardesty, M., Banta, R., White, A., Noone, D., Wolfe, D. and Schnell, R.: A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin, J. Geophys. Res., 119, 6836–6852, http://dx.doi.org/10.1002/2013JD021272, 2014.

Platt, U., and Stutz, J.: Differential Optical Absorption Spectroscopy (DOAS): Principles and applications, ISBN 978-3-540-21193-8, Springer, Berlin-Heidelberg, 2008.

Possell, M. and Hewitt, C. N.: Isoprene emissions from plants are mediated by atmospheric CO_2 concentrations, Glob. Change Biol., 17, 1595–1610, https://doi.org/10.1111/j.1365-2486.2010.02306.x, 2011.

Pougatchev, N. S. and Connor, B. J. and Rinsland, C. P.: Infrared measurements of the ozone vertical distribution above Kitt Peak, J. Geophys. Res., 100, 16689–16697, https://doi.org/10.1029/95JD01296, 1995.

Read, K. A., Carpenter, L. J., Arnold, S. R., Beale, R., Nightingale, P. D., Hopkins, J. R., Lewis, A. C., Lee, J. D., Mendes, L.,and Pickering, S. J.: Multiannual observations of acetone, methanol, and acetaldehyde in remote Tropical Atlantic air: Implications for atmospheric OVOC budgets and oxidative capacity, Envrion. Sci. Technol., 46, 11028–11039, https://doi:10.1021/es302082p, 2012.

Richter, D., Weibring, P., Walega, J. G., Fried, A., Spuler, S. M., and Taubman, M. S.: Compact highly sensitive multi-species airborne mid-IR spectrometer, Appl. Phys. B, 119, 119—131, https://doi.org/10.1007/s00340-015-6038-8, 2015.

Rodgers, C. D.: Inverse methods for atmospheric sounding : theory and practice, World Scientific Publishing, Singapore-New Jersey-London-Hong Kong, ISBN 13: 9789810227401, https://api.semanticscholar.org/CorpusID:60696486, 2000.

Rodgers, C. D., and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, https://doi.org/10.1029/2002JD002299, 2003.

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

Sanchez, D., Seco, R., Gu, D., Guenther, A., Mak, J., Lee, Y., Kim, D., Ahn, J., Blake, D., Herndon, S., Jeong, D., SUllivan, J. T., Mcgee T., Park, R. and Kim, S.: Contributions to OH reactivity from unexplored volatile organic compounds measured by PTR-ToF-MS – a case study in a suburban forest of the Seoul metropolitan area during the Korea–United States Air Quality Study (KORUS-AQ) 2016, Atmos. Chem. Phys., 21, 6331–6345, https://doi.org/10.5194/acp-21-6331-2021, 2021.

Seinfeld, J., and Pandis, S.: Atmospheric chemistry and physics: From air pollution to climate change, 1326 pp., John Wiley and Sons (eds.), https://doi.org/10.1080/00139157.1999.10544295, New York, 1988.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379–1389, https://doi.org/10.2307/2285891, 1968.

Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S., and Chance, K.: Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, J. Geophys. Res., 110, D24301, doi:10.1029/2004JD005629, 2005.

Shu, Y. S., and Atkinson, R.: Rate constants for the gas-phase reactions of O_3 with a series of terpenes and OH radical formation from the O_3 reactions with sesquiterpenes at 296 $\pm$2 K, Int. J. Chem. Kin., 26, 1193–1205, https://doi.org/10.1002/kin.550261207, 1994.

Sinha, V., Williams, J., Lelieveld, J., Ruuskanen, T., Kajos, M. K., Patokoski, J., Hellen, H., Hakola, H., Mogensen, D., Boy, M., Rinne, J., and Kulmala, M.: OH reactivity measurements within a boreal forest: Evidence for unknown reactive emissions, 44, 6614–6620, https://doi.org/10.1021/es101780b, 2010.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P. and Knorr, W.: Global dataset of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317–9341, https://doi.org/10.5194/acp-14-9317-2014, 2014.

Singh, H. B., Brune, W. H., Crawford, J. H., Flocke, F., and Jacob, D. J.: Chemistry and transport of pollution over the Gulf of Mexico and the Pacific: spring 2006 INTEX-B campaign overview and first results, Atmos. Chem. Phys., 9, 2301–2318, https://doi.org/10.5194/acp-9-2301-2009, 2009.

Sofiev, M., Vankevich, R., Ermakova, T., and Hakkarainen, J.: Global mapping of maximum emission heights and resulting vertical profiles of wildfire emissions, Atmos. Chem. Phys., 13, 7039–7052, https://doi.org/10.5194/acp-13-7039-2013, 2013.

Spinei, E., Whitehill, A., Fried, A., Tiefengraber, M., Knepp, T. N., Herndon, S., Herman, J. R., Müller, M., Abuhassan, N., Cede, A., Weibring, P., Richter, D., Walega, J., Crawford, J., Szykman, J., Valin, L., Williams, D. J., Long, R., Swap, R. J., Lee, Y., Nowak, N., and Poche, B.: The First Evaluation of formaldehyde column observations by Pandora spectrometers during the KORUS-AQ field study, Atmos. Meas. Tech., 11, 4943–4961, https://doi.org/10.5194/amt-11-4943-2018, 2018.

Spracklen, D. V., Jimenez, J. L., Carslaw, K. S., Worsnop, D. R., Evans, M. J., Mann, G. W., Zhang, Q., Canagaratna, M. R., Allan, J., Coe, H., McFiggans, G., Rap, A., and Forster, P.: Aerosol mass spectrometer constraint on the global secondary organic aerosol budget, Atmos. Chem. Phys., 11, 12109–12136, https://doi.org/10.5194/acp-11-12109-2011, 2011.

Spurr, R. J. D.: LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems, in: Light Scattering Reviews, edited by: Kokhanovsky, A., 229–271, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-48546-9, 2008.

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

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

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P.-F., and Guenther, A.: How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?, Atmos. Chem. Phys., 15, 11861-11884, https://doi.org/10.5194/acp-15-11861-2015, 2015.

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., and Guenther, A.: Impact of climate variability on volatile organic compounds emissions assessed using OMI formaldehyde observations, Geophys. Res. Lett., 45, 8681–8689, https://doi.org/10.1029/2018GL078676, 2018.

Tikhonov, A.: Solution of incorrectly formulated problems and the regularization method, Sov. Dok., 4, 1035–1038, 1963.

Toon, O. B., Maring, H., Dibb, J., Ferrare, R., Jacob, D. J., Jensen, E. J., Luo, Z. J., Mace, G. G., Pan, L. L., Pfister, L., Rosenlof, K. H., Redemann, J., Reid, J. S., Singh, H. B., Yokelson, R., Chen, G., Jucks, K. W., and Pszenny, A.: Planning, implementation and scientific

1050   goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC[4]RS) field mission, J. Geophys. Res.-Atmos., 121, 4967, https://doi.org/10.1002/2015JD024297, 2016.

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

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., Yu, K., Miller, C. C., Yantosca, R. M., Sulprizio, M. P., Thompson,
1055   A. M., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Cohen, R. C., Laughner, J. L., Dibb, J. E., Hall, S. R., Ullmann, K., Wolfe, G. M., Pollack, I. B., Peischl, J., Neumann, J. A., and Zhou, X.: Why do models overestimate surface ozone in the Southeast United States? Atmos. Chem. Phys., 16, 13561–13577, https://doi.org/10.5194/acp-2016-110, 2016.

Travis, K. R., Heald, C. L., Allen, H. M., Apel, E. C., Arnold, S. R., Blake, D. R., Brune, W. H., Chen, X., Commane, R., Crounse, J. D., Daube, B. C., Diskin, G. S., Elkins, J. W., Evans, M. J., Hall, S. R., Hintsa, E. J., Hornbrook, R. S., Kasibhatla, P. S., Kim, L. J., Luo,
1060   G., McKain, K., Millet, D. B., Moore, F. L., Peischl, J., Ryerson, T. B., Sherwen, T., Thames, A. B., Ullmann, K., Wang, X., Wennberg, P. O., Wolfe, G. M. and Yu, F.: Constraining remote oxidation capacity with ATom observations, Atmos. Chem. Phys., 20, 7753–7781, https://doi.org/10.5194/acp-20-7753-2020, 2020.

Tzompa-Sosa, Z. A., Henderson, B. H., Keller, C. A., Travis, K., Mahieu, E., Franco, B., Estes, M., Helmig, D., Fried, A., Richter, D., Weibring, P., Walega, J., Blake, D. R.and Hannigan, J. W., Ortega, I., Conway, S., Strong, K. and Fischer, E. V.: Atmo-
1065   spheric implications of large C_2-C_5 alkane emissions from the U.S. oil and gas industry, J. Geophys. Res., 124, 1148–1169, https://doi.org/10.1029/2018JD028955, 2019.

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

1070   Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements to the OMI O_2–O_2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9, 6035–6049, https://doi.org/10.5194/amt-9-6035-2016, 2016.

Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., De Smedt, I., Hermans, C., Merlaud, A., Scolas, F., Senten, C., Vanhaelewyn, G., Fally, S., Carleer, M., Metzger, J.-M., Müller, J.-F., Van Roozendael, M., and De Mazière, M.: Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data, Atmos. Chem. Phys., 9, 9523–9544,
1075   https://doi.org/10.5194/acp-9-9523-2009, 2009.

Vigouroux, C., Bauer Aquino, C. A., Bauwens, M., Becker, C., Blumenstock, T., De Mazi'ere, M., Garcia, O., Grutter, M., Guarin, C., Hannigan, J., Hase, F., Jones, N., Kivi, R., Koshelev, D., Langerock, B., Lutsch, E., Makarova, M., Metzger, J.-M., Müller, J.-F., Notholt, J., Ortega, I., Palm, M., Paton-Walsh, C., Poberovskii, A., Rettinger, M., Robinson, J., Smale, D., Stavrakou, T., Stremme, W., Strong, K., Sussmann, R., Té, Y., and Toon, G.: NDACC harmonized formaldehyde time-series from 21 FTIR stations covering a wide range of
1080   column abundances, Atmos. Meas. Tech., 11, 5049–5073, https://doi.org/10.5194/amt-11-5049-2018, 2018.

Vigouroux, C., Langerock, B., Aquino, C. A. B., Blumenstock, T., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J., Jones, N., Kivi, R., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morin, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., Van Roozendael, M., Wang, P., and Winkler, H.: TROPOMI/S5P formaldehyde validation using an extensive network of ground-based FTIR stations, Atmos. Meas. Tech., 3751–3767,
1085   https://doi.org/10.5194/amt-13-3751-2020, 2020.

Vohra, K., Marais, E. A., Suckra, S., Kramer, L., Bloss, W. J., Sahu, R., Gaur, A., Tripathi, S. N., Van Damme, M., Clarisse, L. and Coheur, P.-F.: Long-term trends in air quality in major cities in the UK and India: a view from space, Atmos. Chem. Phys., 21, 6275–6296, https://doi.org/10.5194/acp-21-6275-2021, 2021.

Wang, S., Apel, E., Hornbrook, R., Campos, T., Gao, R.-S., Rogers, D., Pierce, B., Hall, S., and Volkamer, R.: Sources and sinks of OVOCs in the tropical free troposphere during TORERO, Poster A51E-0115, AGU Fall meeting, San Francisco, 2012.

Warneke, C., Trainer, M., de Gouw, J. A., Parrish, D. D., Fahey, D. W., Ravishankara, A. R., Middlebrook, A. M., Brock, C. A., Roberts, J. M., Brown, S. S., Neuman, J. A., Lerner, B. M., Lack, D., Law, D., Hübler, G., Pollack, I., Sjostedt, S., Ryerson, T. B., Gilman, J. B., Liao, J., Holloway, J., Peischl, J., Nowak, J. B., Aikin, K. C., Min, K.-E., Washenfelder, R. A., Graus, M. G., Richardson, M., Markovic, M. Z., Wagner, N. L., Welti, A., Veres, P. R., Edwards, P., Schwarz, J. P., Gordon, T., Dube, W. P., McKeen, S. A., Brioude, J., Ahmadov, R., Bougiatioti, A., Lin, J. J., Nenes, A., Wolfe, G. M., Hanisco, T. F., Lee, B. H., Lopez-Hilfiker, F. D., Thornton, J. A., Keutsch, F. N., Kaiser, J., Mao, J., and Hatch, C. D.: Instrumentation and measurement strategy for the NOAA SENEX aircraft campaign as part of the Southeast Atmosphere Study 2013, Atmos. Meas. Tech., 9, 3063–3093, https://doi.org/10.5194/amt-9-3063-2016, 2016.

Weibring, P., Richter, D., Walega, J. G., and Fried, A.: First demonstration of a high performance difference frequency spectrometer on airborne platforms, Opt. Express, 15 (21), 13,476–13,495, https://doi.org/10.1364/OE.15.013476, 2007.

Wells, K. C., Millet, D. B., Payne, V. H., Deventer, M. J., Bates, K. H., de Gouw, J. A., Graus, M., Warneke, C., Wisthaler, A., and Fuentes, J. D.: Global measurements of isoprene from space: Constraints on emissions and atmospheric oxidation, Nature, 585, 225–233, https://doi.org/10.1038/s41586-020-2664-3, 2020.

Wells, K. C., Millet, D. B., Payne, V. H., Vigouroux, C., Aquino, C. A. B., De Mazière, M., de Gouw, J. A., Graus, M., Kurosu, T., Warneke, C., Wisthaler, A.: Next-generation isoprene measurements from space: Quantifying daily variability at high resolution, J. Geophys. Res., 127, e2021JD036181, https://doi.org/10.1029/2021JD036181, 2022.

Wennberg, P. O., Bates, K. H., Crounse, J. D., Dodson, L. G., McVay, R., Mertens, L. A., Nguyen, T. B., Praske, E., Schwantes, R. H., Smarte, M. D., St Clair, J. M., Teng, A. P., Zhang, X., and Seinfeld, J. H.: Gas-phase reactions of isoprene and its major oxidation products, Chem. Rev., 118, 3337–3390, https://doi.org/10.1021/acs.chemrev.7b00439, 2018.

Williams, J. E., Boersma, K. F., Le Sager, P., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation, Geosci. Model Dev., 10, 721–750, https://doi.org/10.5194/gmd-10-721-2017, 2017.

Yang, Y. D., Shao, M., Wang, X. M., Nolscher, A. C., Kessel, S., Guenther, A., and Williams, J.: Towards a quantitative understanding of total OH reactivity: A review, Atmos. Environ., 134, 147–161, https://doi.org/10.1016/j.atmosenv.2016.03.010, 2016.

Yu, H., Guenther, A., Gu, D., Warneke, C., Geron, C., Goldstein, A., Graus, M., Karl, T., Kaser, L., Misztal, P., and Yuan, B.: Airborne measurements of isoprene and monoterpene emissions from southeastern U.S. forests, Sci. Total Environ., 595, 149–158, http://doi.org/10.1016/j.scitotenv.2017.03.262, 2017.

Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W.: Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling, Remote Sens. Environ., 115, 1171–1187, https://doi.org/10.1016/j.rse.2011.01.001, 2011.

Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., van Geffen, J. H. G. M., Beirle, S., Wagner, T., Van Roozendael, M., Marchenko, S., Lamsal, L. N., and Eskes, H. J.: Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty characterization, and trends, Atmos. Meas. Tech., 11, 4033–4058, https://doi.org/10.5194/amt-11-4033-2018, 2018.

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M., Sulprizio, M. P., De Smedt, I., Gonzalez Abad, G., Chance, K., Li, C., Ferrare, R., Fried, A., Hair, J. W., Hanisco, T. F., Richter, D., Jo Scarino, A., Walega, J., Weibring, P.,

and Wolfe, G. M.: Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC[4]RS aircraft observations over the southeast US, Atmos. Chem. Phys., 16, 13477–13490, https://doi.org/10.5194/acp-16-13477-2016, 2016.

Zhu, L., Mickley, L. J., Jacob, D. J., Marais, E. A., Sheng, J., Hu, L., Gonzalez Abad, G., and Chance, K.: Long-term (2005-2014) trends in formaldehyde columns across North America as seen by the OMI satellite instrument; Evidence of changing emissions of volatile organic compounds, J. Geophys. Res., 44, 7079–7086, https://doi.org/10.1002/2017GL073859, 2017.

Zhu, L., González Abad, G., Nowlan, C. R., Chan Miller, C., Chance, K., Apel, E. C., DiGangi, J. P., Fried, A., Hanisco, T. F., Hornbrook, R. S., Hu, L., Kaiser, J., Keutsch, F. N., Permar, W., St. Clair, J. M. and Wolfe, G. M.: Validation of satellite formaldehyde (HCHO) retrievals using observations from 12 aircraft campaigns, Atmos. Chem. Phys., 20, 12329–12345, https://doi.org/10.5194/acp-20-12329-2020, 2020.