Atmospheric
Measurement
Techniques

NDACC harmonized formaldehyde time series from 21 FTIR stations covering a wide range of column abundances

Corinne Vigouroux[1], **Carlos Augusto Bauer Aquino**[2], **Maite Bauwens**[1], **Cornelis Becker**[3], **Thomas Blumenstock**[4], **Martine De Mazière**[1], **Omaira García**[5], **Michel Grutter**[6], **César Guarin**[6], **James Hannigan**[7], **Frank Hase**[4], **Nicholas Jones**[8], **Rigel Kivi**[9], **Dmitry Koshelev**[10], **Bavo Langerock**[1], **Erik Lutsch**[11], **Maria Makarova**[12], **Jean-Marc Metzger**[13], **Jean-François Müller**[1], **Justus Notholt**[14], **Ivan Ortega**[7], **Mathias Palm**[15], **Clare Paton-Walsh**[8], **Anatoly Poberovskii**[12], **Markus Rettinger**[15], **John Robinson**[16], **Dan Smale**[16], **Trissevgeni Stavrakou**[1], **Wolfgang Stremme**[6], **Kim Strong**[11], **Ralf Sussmann**[15], **Yao Té**[10], and **Geoffrey Toon**[17]

[1]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[2]Instituto Federal de Educaçao (IFRO), Ciência e Tecnologia de Rondônia, Porto Velho, Brazil
[3]Stichting Atmospherische en Hydrologische Ontwikkeling (SAHO), Paramaribo, Suriname
[4]Karlsruhe Institute of Technology (KIT), IMK-ASF, Karlsruhe, Germany
[5]Izaña Atmospheric Research Centre (IARC), Agencia Estatal de Meteorología (AEMET), Santa Cruz de Tenerife, Spain
[6]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México (UNAM), 04510 Mexico City, México
[7]Atmospheric Chemistry, Observations & Modeling, National Center for Atmospheric Research (NCAR), Boulder, CO, USA
[8]Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, Australia
[9]Finnish Meteorological Institute (FMI), Sodankylä, Finland
[10]LERMA-IPSL, Sorbonne Université, CNRS, PSL Research University, Observatoire de Paris, 75005 Paris, France
[11]Department of Physics, University of Toronto, Toronto, Canada
[12]Saint Petersburg State University, Atmospheric Physics Department, St Petersburg, Russia
[13]Observatoire des Sciences de l'Univers Réunion (OSU-R), UMS 3365, Université de la Réunion, Saint-Denis, France
[14]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[15]Karlsruhe Institute of Technology (KIT), IMK-IFU, Garmisch-Partenkirchen, Germany
[16]National Institute of Water and Atmospheric Research Ltd (NIWA), Lauder, New Zealand
[17]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Correspondence: Corinne Vigouroux (corinne.vigouroux@aeronomie.be)

Abstract. Among the more than 20 ground-based FTIR (Fourier transform infrared) stations currently operating around the globe, only a few have provided formaldehyde (HCHO) total column time series until now. Although several independent studies have shown that the FTIR measurements can provide formaldehyde total columns with good precision, the spatial coverage has not been optimal for providing good diagnostics for satellite or model validation. Furthermore, these past studies used different retrieval settings, and biases as large as 50 % can be observed in the HCHO total columns depending on these retrieval choices, which is also a weakness for validation studies combining data from different ground-based stations.

For the present work, the HCHO retrieval settings have been optimized based on experience gained from past studies and have been applied consistently at the 21 participating stations. Most of them are either part of the Network for the Detection of Atmospheric Composition Change (NDACC) or under consideration for membership. We provide the harmonized settings and a characterization of the HCHO FTIR products. Depending on the station, the total systematic and random uncertainties of an individual HCHO total column measurement lie between 12 % and 27 % and between 1

and 11×10^{14} molec cm^{-2}, respectively. The median values among all stations are 13 % and 2.9×10^{14} molec cm^{-2} for the total systematic and random uncertainties.

This unprecedented harmonized formaldehyde data set from 21 ground-based FTIR stations is presented and its comparison with a global chemistry transport model shows consistency in absolute values as well as in seasonal cycles. The network covers very different concentration levels of formaldehyde, from very clean levels at the limit of detection (few 10^{13} molec cm^{-2}) to highly polluted levels (7×10^{16} molec cm^{-2}). Because the measurements can be made at any time during daylight, the diurnal cycle can be observed and is found to be significant at many stations. These HCHO time series, some of them starting in the 1990s, are crucial for past and present satellite validation and will be extended in the coming years for the next generation of satellite missions.

1 Introduction

Through reactions with hydroxyl radical (OH) and NO$_x$ (NO + NO$_2$), the volatile organic compounds (VOCs) exert a strong influence on the oxidizing capacity of the atmosphere. These reactions produce ozone and secondary organic aerosols, which affect air quality and global climate. Given their short lifetimes (from a few minutes to a few hours for the most reactive ones, Kesselmeier and Staudt, 1999) and their different sources depending on geographical locations, it is very difficult to derive a global atmospheric burden for most of the VOCs from current measurements. The observation of formaldehyde (HCHO), which is an intermediate product of the degradation of many nonmethane VOCs (NMVOCs) and has a lifetime of only a few hours, allows us to constrain the NMVOCs emissions and to test our understanding of the complex and still uncertain degradation mechanisms of these NMVOCs (Stavrakou et al., 2009). The use of satellite HCHO measurements in combination with tropospheric chemistry transport models to derive NMVOCs emissions has been the subject of several past studies (e.g. Palmer et al., 2003; Millet et al., 2008; Stavrakou et al., 2009; Fortems-Cheiney et al., 2012; Barkley et al., 2013; Marais et al., 2014). The past and present HCHO satellite data sets include those from GOME (1996–2003), SCIAMACHY (2003–2012), OMI (2004–present), GOME-2A (2006–present), OMPS (2011–present), GOME-2B (2012–present), and very recently TROPOMI (2017–present). The NMVOCs emissions derived from top-down approaches using these satellite data sets rely on the accuracy of the measurements. An indirect way to test these accuracies is to compare the emission budgets obtained using two different satellite data sets as in Barkley et al. (2013) for SCIAMACHY and OMI or in Stavrakou et al. (2015) for OMI and GOME-2. While the global emission budgets are

in general consistent (Stavrakou et al., 2015), there are large differences in the top-down estimates on a regional scale, e.g. differences up to nearly 50 % are observed over Amazonia between SCIAMACHY and OMI (Barkley et al., 2013) and up to nearly 25 % between GOME-2 and OMI (Stavrakou et al., 2015). Unambiguously concluding whether these differences are due to biases in the satellite products (due to retrieval settings, vertical sensitivities, horizontal resolution, etc.) or to the diurnal cycle of formaldehyde (SCIAMACHY and GOME-2 measuring in the morning and OMI in the afternoon) requires validation with independent and accurate ground-based measurements (Barkley et al., 2013; De Smedt et al., 2015; Stavrakou et al., 2015).

At present, validation studies of HCHO satellite products have taken place at a few locations only, mainly using aircraft data (Martin et al., 2004; Barkley et al., 2013; Zhu et al., 2016), the MAX-DOAS (Multi-Axis Differential Optical Absorption Spectroscopy) technique (Wittrock et al., 2006; De Smedt et al., 2015) and the FTIR (Fourier transform infrared) technique (Jones et al., 2009; Vigouroux et al., 2009; De Smedt et al., 2015). This is not sufficient to provide a good picture of the accuracy of the satellites, especially given the high geographical variability of formaldehyde. A lot of effort is therefore currently underway to increase the number of ground-based stations providing HCHO data, using the DOAS or the FTIR technique, initiated in view of the TROPOMI Cal/Val activities. This paper presents the work accomplished in this direction using FTIR measurements at most of the NDACC (Network for the Detection of Atmospheric Composition Change) stations, and including some more recent observing stations that will also be part of the NDACC in the near future.

Up to now, time series of HCHO total columns have been studied at only six FTIR stations out of more than 20 FTIR sites currently in operation: Ny-Ålesund (Notholt et al., 1997), Wollongong (Paton-Walsh et al., 2005), Lauder (Jones et al., 2009), Reunion Island (Vigouroux et al., 2009), Eureka (Viatte et al., 2014), and Jungfraujoch (Franco et al., 2015). We note that HCHO has also been measured by the JPL MkIV instrument (Toon, 1991) at various ground-based sites since 1985 (see http://mark4sun.jpl.nasa.gov/ground.html, last access: 5 September 2018), although these data are not used in this work due to their very different acquisition and analysis procedures. The main reasons for having so few FTIR HCHO data available are that (1) it is challenging to find robust retrieval settings for this species that has weak absorption signatures in the infrared, which are, in addition, surrounded by strong lines from interfering gases; (2) HCHO is not part of the NDACC FTIR target species (which are O$_3$, HNO$_3$, HF, HCl, CO, CH$_4$, N$_2$O, ClONO$_2$, HCN, and C$_2$H$_6$, publicly available at http://www.ndsc.ncep.noaa.gov/clickmap/, last access: 5 September 2018). In the above-cited studies, different retrieval settings are used, although the retrieved HCHO total columns can be very sensitive to some of them: e.g. a positive bias of 25 % or even 50 % is found

at Reunion Island if the spectral micro-windows of Franco et al. (2015) or Jones et al. (2009) are used, respectively, instead of those from Vigouroux et al. (2009). Although these high biases are consistent with the uncertainty budgets, it is important to facilitate the interpretation of a satellite or model validation to harmonize the settings among the stations. Therefore, in the present work, we have set up common retrieval settings that can be used at any ground-based site, even under very humid conditions or low HCHO concentrations. These settings will be described in Sect. 2 together with a characterization of the retrieved HCHO products, i.e. their averaging kernels and uncertainty budget. The complete time series of HCHO total columns obtained at the 21 participating stations are shown in Sect. 3, as well as the diurnal cycles and a short assessment of the long-term trends. We then use the chemistry transport model IMAGES (Stavrakou et al., 2015), which provides data for the 2003–2016 period, to show consistency in our harmonized FTIR data sets: comparisons between FTIR and IMAGES monthly mean time series and seasonal cycle at the 21 stations are presented in Sect. 4.

2 Ground-based FTIR HCHO data: description and characterization

2.1 FTIR HCHO monitoring

Table 1 lists the ground-based FTIR stations included in this study, while Fig. 1 shows their geographical distribution. These stations take regular solar absorption measurements under clear-sky conditions, using either the high-resolution spectrometers Bruker 120 M, 125 M, 120 HR, and/or 125 HR, which can achieve a spectral resolution of $0.0035\,\mathrm{cm}^{-1}$ or better, or the Bomem DA8, which can achieve a spectral resolution of $0.004\,\mathrm{cm}^{-1}$. The only lower spectral resolution spectrometer ($0.06\,\mathrm{cm}^{-1}$) used in this study is the Bruker Vertex at Mexico City. This instrument is not accepted by the NDACC FTIR standards at present; therefore Mexico City is the only site in this study that will not be part of NDACC.

The formaldehyde spectral signatures used in ground-based infrared measurements lie in the 3.6 μm region and belong to the ν_1 and ν_5 bands. This implies that, for HCHO, a CaF_2 or KBr beamsplitter and a nitrogen-cooled InSb detector are used together with an optical filter which usually covers the 2400–3310 cm^{-1} region (called an NDSC-3 filter; see e.g. Senten et al., 2008). At St Petersburg a broader filter is used (1700–3400 cm^{-1}). The spectral resolution can be reduced in order to increase the signal-to-noise ratio (SNR). In practice, the spectra used in the present study have a resolution between 0.0035 and 0.009 cm^{-1}, except for Mexico city ($0.075\,\mathrm{cm}^{-1}$).

HBr or N_2O cell measurements are regularly taken to verify the alignment of the instruments. The instrument line shape (ILS) can be obtained by analysing these cell measurements using the LINEFIT programme (Hase et al., 1999). This ILS can impact the shape of gas absorption lines, and its determination by LINEFIT can be used as an input parameter in the forward model of the retrieval codes (Sect. 2.2).
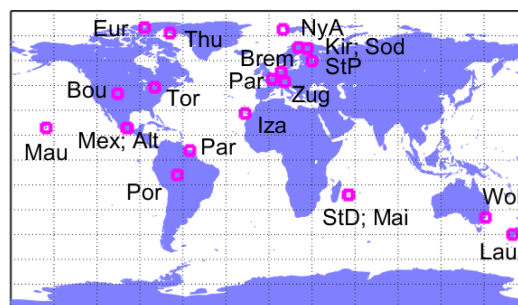
2.2 Harmonized retrieval strategy

We refer to Pougatchev et al. (1995) and/or Hase et al. (2004) for more details on the FTIR retrieval principles. Total columns of atmospheric gases, but also volume mixing ratio vertical profiles, are obtained from their pressure- and temperature-dependent absorption lines. As seen in Table 1, two retrieval algorithms are used in the NDACC FTIR community: PROFITT9 (Hase et al., 2006), and SFIT2 (Pougatchev et al., 1995), which has been updated to SFIT4 09.4.4. It has been demonstrated in Hase et al. (2004) that the profiles and total column amounts retrieved from these two different algorithms under identical conditions are in excellent agreement.

We summarize the forward-model and retrieval parameters that have been harmonized in Table 2. The forward model uses pressure and temperature profiles from NCEP (National Centers for Environmental Prediction) for each site, except that the temporal resolution can vary depending on the retrieval team from daily means, from 6-hourly ones to even hourly interpolated ones.

The dominant source of systematic uncertainty being the spectroscopic parameters, it is crucial that all stations use the same spectroscopic database. We use the compilation from Geoffrey Toon (JPL), the atm16 line list, which is available at http://mark4sun.jpl.nasa.gov/toon/linelist/linelist.html (last access: 5 September 2018). In this atm16 line list, the HCHO and N_2O lines correspond to the HITRAN 2012 database (Rothman et al., 2013). This HITRAN 2012 database includes the latest improved HCHO parameters (broadening coefficients, Jacquemart et al., 2010), which complement the release in HITRAN 2008 (Rothman et al., 2009) of new HCHO line intensities from the same group (Perrin et al., 2009). The spectroscopic parameters for the lines of water vapour and its isotopologues in

Figure 1. Locations of the FTIR stations providing HCHO total columns.

Table 1. Characteristics of the FTIR stations contributing to the present work: location and altitude (in km a.s.l.), time period used in the present study, instrument type, retrieval code, and team.

Station	Latitude	Longitude	Altitude	Time period	Instrument	Code	Team
Eureka	80.05° N	86.42° W	0.61	2006–2016	Bruker 125 HR	SFIT4	University of Toronto
Ny-Ålesund	78.92° N	11.92° E	0.02	1993–2017	Bruker 120/5 HR	SFIT4	University of Bremen
Thule	76.52° N	68.77° W	0.22	1999–2016	Bruker 120 M	SFIT4	NCAR
Kiruna	67.84° N	20.40° E	0.42	2005–2016	Bruker 120/5 HR	PROFFIT	KIT / IMK–ASF
Sodankyla	67.37° N	26.63° E	0.19	2012–2017	Bruker 125 HR	SFIT4	FMI & BIRA
St Petersburg	59.88° N	29.83° E	0.02	2009–2017	Bruker 125 HR	SFIT4	St Petersburg State University
Bremen	53.10° N	8.85° E	0.03	2004–2017	Bruker 125 HR	SFIT4	University of Bremen
Paris	48.97° N	2.37° E	0.06	2011–2016	Bruker 125 HR	PROFFIT	LERMA
Zugspitze	47.42° N	10.98° E	2.96	1995–2017	Bruker 120/5 HR	PROFFIT	KIT / IMK–IFU
Toronto	43.60° N	79.36° W	0.17	2002-2016	Bomem DA8	SFIT4	University of Toronto
Boulder	40.04° N	105.24° W	1.61	2010-2016	Bruker 120 HR	SFIT4	NCAR
Izaña	28.30° N	16.48° W	2.37	2005–2016	Bruker 125 HR	PROFFIT	
Mauna Loa	19.54° N	155.57° W	3.40	1995–2016	Bruker 120/5 M	SFIT4	NCAR
Mexico City	19.33° N	99.18° W	2.26	2013–2016	Bruker Vertex 80	PROFFIT	UNAM
Altzomoni	19.12° N	98.66° W	3.98	2012–2016	Bruker 120/5 HR	PROFFIT	UNAM
Paramaribo	5.81° N	55.21° W	0.03	2004–2016	Bruker 120/5 M	SFIT4	University of Bremen
Porto Velho	8.77° S	63.87° W	0.09	2016–2017	Bruker 125 M	SFIT4	BIRA
Saint-Denis	20.90° S	55.48° E	0.08	2004–2011	Bruker 120 M	SFIT4	BIRA
				2011–2013	Bruker 125HR		
Maïdo	21.08° S	55.38° E	2.16	2013–2017	Bruker 125 HR	SFIT4	BIRA
Wollongong	34.41° S	150.88° E	0.03	1996–2007	Bomem DA8	SFIT4	University of Wollongong
				2007–2016	Bruker 125 HR		
Lauder	45.04° S	169.68° E	0.37	2001–2016	Bruker 120 HR	SFIT4	NIWA

Table 2. Summary of the HCHO harmonized forward-model and retrieval parameters. The micro-window limits are given in cm^{-1}.

Pressure and temperature profiles	NCEP
Spectroscopic database	atm16 (= HITRAN 2012 for HCHO)
Solar lines	SFIT4.09.4.4
Micro-windows	MW 1: 2763.42–2764.17 MW 2: 2765.65–2766.01 MW 3: 2778.15–2779.1 MW 4: 2780.65–2782.0
De-weighted spectral sections	2780.967–2780.993 (O_3) 2781.42–2781.48 (CH_4)
Retrieved species	HCHO, HDO, CH_4, O_3, N_2O, solar lines optional: CO_2, H_2O
A priori profiles (except HDO and H_2O)	WACCM v4
Regularization	Tikhonov L1

atm16 are from Toth 2003[1]; some lines from the other strong absorbing gases in the vicinity of HCHO (O_3 and CH_4) have been empirically adjusted or replaced with older HITRAN versions in atm16 when obvious problems were found in the HITRAN 2012 database. For the CO solar lines, we use the line list updated from Hase et al. (2010) that is distributed in the NDACC community (SFIT4 package v09.4.4).

To avoid any bias between the stations due to different spectroscopic parameters, it is also mandatory to harmonize the spectral micro-windows (MWs) containing the HCHO signatures. The challenge of the HCHO retrievals is that this species has very weak absorption signatures in the infrared (below 1 %), and it is therefore very important to minimize the impact of the interfering gases with more intense signatures, either by avoiding MWs with strong interfering lines when feasible or by including them only if they are very well fitted (e.g. no large residuals remain due to bad spectroscopic or incorrect ILS parameters). In past studies, while the micro-window spectral widths differ, some common HCHO signatures were used: the two more intense signatures at about 2778.5 and $2781.0 \, cm^{-1}$ were used in all previous studies (Notholt et al., 1997; Paton-Walsh et al., 2005; Viatte et al., 2014; Jones et al., 2009; Vigouroux et al., 2009), except in Franco et al. (2015), who discarded the

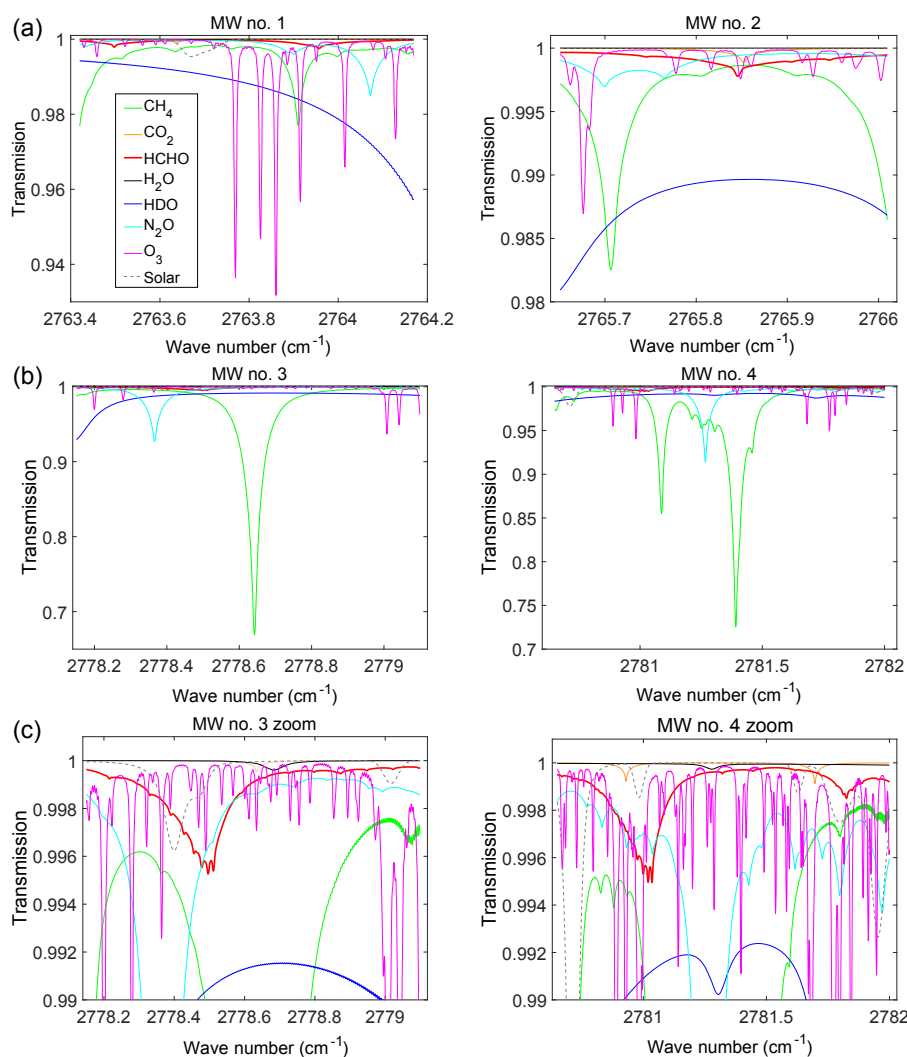[1] http://mark4sun.jpl.nasa.gov/data/spec/H2O/RAToth_H2O.tar (last access: 5 September 2018)

Figure 2. Retrieved contributions of all fitted species in the four MWs **(a, b)** used in the analysis for a spectrum recorded on 12 February 2014 at Maïdo and corresponding to a retrieved HCHO total column of 2.48×10^{15} molec cm^{-2}. Panels **(c)** are magnifications of the MWs nos. 3 and 4.

2781.0 cm^{-1} signature because of the bad residuals due to poorly fitted CH$_4$ lines (from HITRAN 2008, Rothman et al., 2009). In Vigouroux et al. (2009), in which HITRAN 2004 (Rothman et al., 2005) was used, the MWs containing these two stronger signatures were quite narrow (2778.20–2778.59; 2780.80–2781.15 cm^{-1}) in order to minimize residuals due to neighbouring CH$_4$ lines. With the empirically improved CH$_4$ spectroscopy in atm16, we can use larger windows (see Table 2 and Fig. 2), with the advantage of fixing the background and the interfering species more, leading to improved precision and accuracy in the HCHO total columns. We keep the two narrow MWs used in Vigouroux et al. (2009) and Franco et al. (2015) at about 2763.5 and 2765.8 cm^{-1}, which contain less absorption from interfering gases, but the gain in information in degrees of freedom for signal (DOFS; see Rodgers, 2000) is relatively small (0.1–

0.2, compared to about 1.0 to 1.5 from the two main windows).

In Fig. 2 we give an example of a spectrum calculated from the retrieval using a spectrum recorded on the 12 February 2014 at Maïdo and corresponding to a retrieved HCHO total column of 2.48×10^{15} molec cm^{-2}, a DOFS of 1.1, and a root mean square (rms) of 0.11, which compares well to the mean obtained for all measurements at Maïdo of 2.00×10^{15} molec cm^{-2}, 1.2, and 0.12 for columns, DOFS and rms. The corresponding residuals (calculated − observed spectra) are shown in Fig. 3, when the spectroscopic parameters are taken from HITRAN 2012 and with the atm16 empirical line list. We can see the improvement in MW 1 obtained simply by changing the line position of an O$_3$ line (2763.8598 cm^{-1} instead of 2763.8588 cm^{-1}). The spectroscopic parameters in MW 2 are the same in both cases, the
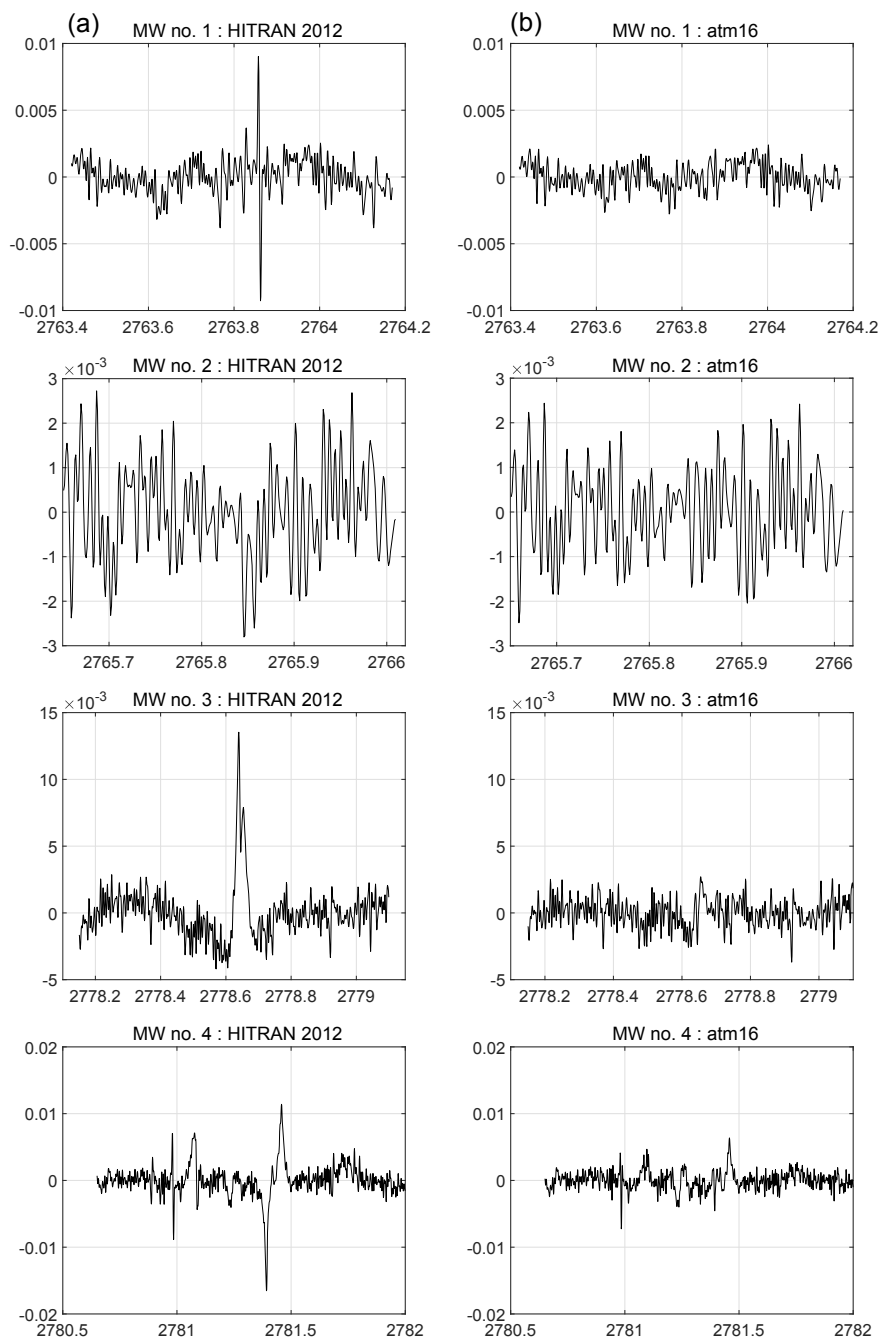
Figure 3. Residuals (calculated – observed spectrum) in each of the four MWs for the retrieval of a spectrum recorded on 12 February 2014 at Maïdo and corresponding to a retrieved HCHO total column of 2.48×10^{15} molec cm^{-2}. The x axis represents the wave number in cm^{-1}. Panels **(a)** are obtained when the HITRAN 2012 spectroscopy is used, and panels **(b)** show the improvement made by using the atm16 line list.

little improvement seen in this MW is due to the better fitting of the other MWs, which allows better-calculated profiles for all gases. The CH$_4$ line in MW 3 is poorly fitted using the HITRAN 2012 line list, and the improvement in the atm16 is due to a change in several spectroscopic parameters (line position, line intensity, etc.). The two more intense CH$_4$

lines in MW 4 have also been improved by using the atm16 line list. However, to further improve the fits, one CH$_4$ line and one O$_3$ line were empirically de-weighted (see Table 2). The comparison of these two line lists shows the crucial need for good spectroscopic parameters in order to obtain precise amounts of atmospheric gases. As seen in Fig. 3 (right panel),

the residuals are not perfect and there is still room for further improvement in the forward-model parameters. The atm line list created by Geoffrey Toon (JPL) is updated every 4 years, when HITRAN provides a new release, so that when the HITRAN line list is improved and provides either similar or better residuals than the atm line list, the empirical parameters of atm are changed by the preferred official database.

In SFIT4 and PROFFIT retrieval codes, based on optimal estimation, a priori information (profile and regularization matrix) needs to be provided. In this work, the a priori HCHO profile, as well as all interfering species except water vapour and its isotopologues, were provided for each station from the v4 of the model WACCM (Garcia et al., 2007). A single profile for each species is used in the time series retrievals and corresponds to the mean of the model profiles calculated at each station from 1980 to 2020. For H_2O and HDO, which have a high atmospheric variability, it is usually preferred (except at the stations Lauder, Mexico City, and Altzomoni) not to use a single a priori profile: for each individual spectrum, the water vapour a priori profiles are taken either from the 6-hourly vertical profiles provided by NCEP or from independent preliminary profile retrievals. The H_2O absorption being very weak in the chosen MWs and the HDO profile being retrieved simultaneously with HCHO, the impact of using a single a priori profile at the three cited stations is assumed to be small. For the regularization matrix $\mathbf{R}$, we followed Vigouroux et al. (2009) and Sussmann et al. (2011) and used ad hoc Tikhonov (Tikhonov, 1963) L1 regularization as described, for example, in Steck (2002), for the reason that we do not have a realistic a priori covariance matrix $\mathbf{S}_{var}$ from other measurements sources, especially with good vertical resolution. The regularization matrix $\mathbf{R} = \alpha \mathbf{L}_1^T \mathbf{L}_1$ is used in most cases for the determination of HCHO low vertical resolution profiles but also for profile retrievals of the interfering species when improvement is observed compared to the fit of a single scaling factor (which is applied to the a priori profiles). This is the case for HDO and CH_4, for which profile retrievals are made, and at some stations for O_3. For the stations Kiruna, Izaña, Zugspitze, and Paris, a scaling of HCHO a priori profiles is preferred to a Tikhonov regularization, but due to the low DOFS available for this species (see Sect. 2.3), this has little influence on the retrieved total columns (below 2 % when tested at Maïdo). For the other stations, the α values are site dependent, since they can depend, for example, on the HCHO amounts or the SNR of the spectra. Note that the SNR value may be the "real" one from the inherent noise in each spectrum but can also be chosen to be an "effective" SNR that is used as well as a regularization parameter. This effective SNR is smaller than the real one, since the residuals in a spectral fit do not only come from pure measurement noise but also from uncertainties in the forward-model parameters. The regularization choice (α and SNR if an effective one is used) is made at each station in order to obtain stable retrievals (no overfitting) with a sig-

nificant decrease in the residuals (no underfitting), as in the well-known L-curve method (Hansen , 1992).

It is worth noting that another important forward-model parameter is the instrumental line shape (ILS) since it impacts the gases absorption line shapes. The treatment of ILS in the retrievals has not been harmonized yet among the stations because the stability and quality of the alignment is site dependent and/or the instrument's PIs have their own preferences. This is, however, another step toward full harmonization that should be done in the future within NDACC. At present, there are three options for considering the ILS, and we refer to Vigouroux et al. (2015) for more details. In the present work, the NIWA, NCAR and University of Bremen stations use a constant and ideal ILS (both modulation efficiency and phase error); i.e. the spectrometers are perfectly aligned. This is a valid approximation based on a LINEFIT ILS analysis of HBr cell spectra measurements (Sect. 2.1). The IMK-ASF, LERMA and UNAM stations use fixed ILS parameters that are previously retrieved using the cell measurements and the LINEFIT code (Hase et al., 1999). For the other stations, the effective apodization parameter is retrieved simultaneously with the target species profiles, while the phase error parameter is assumed to be ideal.

2.3 Characterization: averaging kernels and uncertainty budget

The vertical resolution and sensitivity of the retrieved HCHO products can be characterized by the averaging kernel matrix $\mathbf{A}$ (Rodgers, 2000):

$$\mathbf{A} = \left(\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{R} \right)^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K}, \tag{1}$$

where $\mathbf{K}$ is the weighting function matrix that links the measurement vector $\boldsymbol{y}$ to the state vector $\boldsymbol{x}$: $\boldsymbol{y} = \mathbf{K}\boldsymbol{x} + \boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon}$ representing the measurement error. In our retrievals, we assume $\mathbf{S}_\epsilon$ to be diagonal, with the diagonal elements being the inverse square of the SNR. $\mathbf{R}$ is the regularization matrix, which, in this work, has been chosen as the Tikhonov L1 matrix (see Sect. 2.2).

We give the trace of this averaging kernel matrix $\mathbf{A}$ for the elements corresponding to the HCHO profiles, called the DOFS, in Table 3 for each station. The DOFS range from 1.0 to 1.6, meaning that we can not provide more than one piece of information on the vertical profile. This is the reason that only total columns of HCHO are discussed in this paper and not vertical profiles. In Fig. 4 we show (upper panels) the averaging kernels (AKs, rows of $\mathbf{A}$) for four different stations, with DOFS ranging from 1 (only scaling) to 1.4. Similar averaging kernels are obtained for the other stations with similar DOFS (not shown). We can observe that, in each case, the AKs peak at about the same altitude (8 km) with full width at half maximum of about 16–18 km, showing that we have limited vertical resolution, and that we are mainly sensitive to the whole troposphere, and to a lesser extent to the lowermost

Table 3. Mean of the HCHO total columns (TC) in 10^{14} molec cm^{-2} and degrees of freedom for signal (DOFS) obtained at each FTIR station. The stations with strictly 1 DOFS (Kiruna, Izaña, Zugspitze, and Paris) only make a scaling of the HCHO a priori profile; i.e. no change in the vertical shape of the a priori profile is allowed. We give, in 10^{14} molec cm^{-2}, the mean of (1) the random uncertainties (Rand) that were calculated for each individual HCHO total column (excluding the smoothing part); (2) the smoothing random error (Smoo Rand); (3) the total random error (Total Rand $= \sqrt{\text{Rand}^2 + \text{Smoo Rand}^2}$). We also provide the total random error in % for completeness. We give the mean of the systematic uncertainties in %: first without the smoothing part (Syst), then the smoothing systematic error (Smoo Syst), and the total systematic error (Total Syst $= \sqrt{\text{Syst}^2 + \text{Smoo Syst}^2}$). If the Rodgers and Connor (2003) methodology is used in model–instrument comparisons, only the Rand and Syst uncertainties need to be taken into account (not the total errors). In addition, we provide the mean differences between two subsequent FTIR measurements taken within 30 min (Diff30) in both absolute (10^{14} molec cm^{-2}) and percent units relative to mean TC. The PROFFIT stations are indicated with (*).

Station	DOFS	Mean TC	Rand	Smoo Rand	Total Rand	Syst	Smoo Syst	Total Syst	Diff30
Eureka	1.3	12.7	1.0	0.6	1.2 (9.3 %)	12.2 %	3.5 %	12.8 %	1.5 (11.7 %)
Ny-Ålesund	1.6	15.8	1.8	0.5	1.9 (11.7 %)	13.3 %	3.4 %	13.8 %	3.9 (24.9 %)
Thule	1.1	15.7	1.3	0.9	1.5 (9.8 %)	14.3 %	3.8 %	14.8 %	1.8 (11.7 %)
Kiruna*	1	17.5	3.5	0.8	3.6 (20.8 %)	25.6 %	8.6 %	27.1 %	0.7 (3.8 %)
Sodankyla	1.1	25.4	1.5	1.7	2.3 (9.0 %)	13.4 %	3.8 %	14.1 %	2.4 (9.3 %)
St Petersburg	1.4	59.4	2.6	2.1	3.3 (5.6 %)	13.9 %	2.4 %	14.2 %	2.8 (4.6 %)
Bremen	1.2	59.6	2.3	1.7	2.9 (4.8 %)	12.9 %	2.9 %	13.3 %	3.1 (5.2 %)
Paris*	1	73.0	5.3	1.4	5.5 (7.6 %)	16.3 %	4.6 %	17.0 %	3.3 (4.8 %)
Zugspitze*	1	12.3	2.2	0.5	2.3 (18.6 %)	20.7 %	5.8 %	21.7 %	1.0 (8.0 %)
Toronto	1.3	95.1	5.1	4.1	6.7 (7.1 %)	12.6 %	2.7 %	13.0 %	19.3 (20.4 %)
Boulder	1.1	57.6	2.6	3.9	4.7 (8.2 %)	12.7 %	2.1 %	13.0 %	5.3 (9.2 %)
Izaña*	1	20.4	3.3	0.2	3.3 (16.0 %)	20.9 %	4.4 %	21.4 %	0.8 (4.0 %)
Mauna Loa	1.1	10.1	1.4	1.0	1.8 (17.3 %)	12.5 %	3.8 %	13.1 %	1.4 (14.0 %)
Mexico City*	1.0	220.9	11.1	2.5	11.4 (5.2 %)	12.0 %	1.2 %	12.1 %	24.0 (10.9 %)
Altzomoni*	1.1	21.8	2.3	1.2	2.6 (11.7 %)	16.0 %	3.2 %	16.3 %	2.3 (10.5 %)
Paramaribo	1.5	64.3	3.4	1.3	3.6 (5.6 %)	12.2 %	3.1 %	12.7 %	11.9 (18.5 %)
Porto Velho	1.1	190.0	3.5	8.3	9.1 (4.8 %)	12.8 %	4.1 %	13.5 %	5.9 (3.1 %)
Saint-Denis	1.2	38.8	2.2	0.8	2.4 (6.1 %)	13.4 %	4.3 %	14.1 %	2.8 (7.2 %)
Maïdo	1.2	20.0	1.4	0.4	1.4 (7.3 %)	12.9 %	2.3 %	13.1 %	1.1 (5.6 %)
Wollongong	1.5	78.9	3.0	2.2	3.7 (4.7 %)	10.9 %	3.0 %	11.6 %	11.6 (15.0 %)
Lauder	1.4	25.6	1.5	0.4	1.6 (6.3 %)	12.4 %	2.6 %	12.8 %	3.6 (14.0 %)
Median	1.1	25.6	2.3	1.2	2.9 (7.6 %)	12.9 %	3.4 %	13.5 %	2.8 (9.3 %)

stratosphere. The total column averaging kernel (TotAK), associated with the FTIR-retrieved total columns, is plotted as well. The associated a priori profiles are also shown in Fig. 4 (lower panels) for completeness, together with the mean and standard deviation of the retrieved profiles. As expected by the low DOFS, the shape of the retrieved profiles is very similar to the shape of the a priori profiles.

The uncertainty budget is calculated following the formalism of Rodgers (2000) and can be divided into three different sources: the measurement noise uncertainty (purely random), the forward-model parameter uncertainties (random and systematic), and the smoothing error expressing the uncertainty due to the limited vertical resolution of the retrieval (random and systematic). At each station, the random uncertainty (square root of sum of squares of the measurement noise error and of all the random forward-model errors) and the systematic uncertainty (square root sum of the squares of all systematic errors) are calculated for each single measurement. Except for a few cases (NCAR stations and Wollongong), for

which a typical smoothing error is given, and St Petersburg, for which the mean value for 2013 is given, the smoothing uncertainty is also calculated for each individual measurement. In Table 3 we give the mean of the random and systematic uncertainties, the smoothing uncertainties (both random and systematic parts), and the total random or systematic uncertainties (square root sum of the squares of the random or systematic error and the smoothing random or systematic error), obtained using the FTIR complete time series at each station.

The random uncertainty given in Table 3 is dominated at all sites by the measurement noise with an error covariance matrix S_n calculated as follows:

$$S_n = G_y S_\epsilon G_y^T, \qquad (2)$$

where S_ϵ is assumed to be diagonal, with the square of the inverse of the SNR as diagonal elements, and G_y denotes the contribution matrix $A = G_y K$. In this calculation of the measurement noise error, the SNR must be the real one from
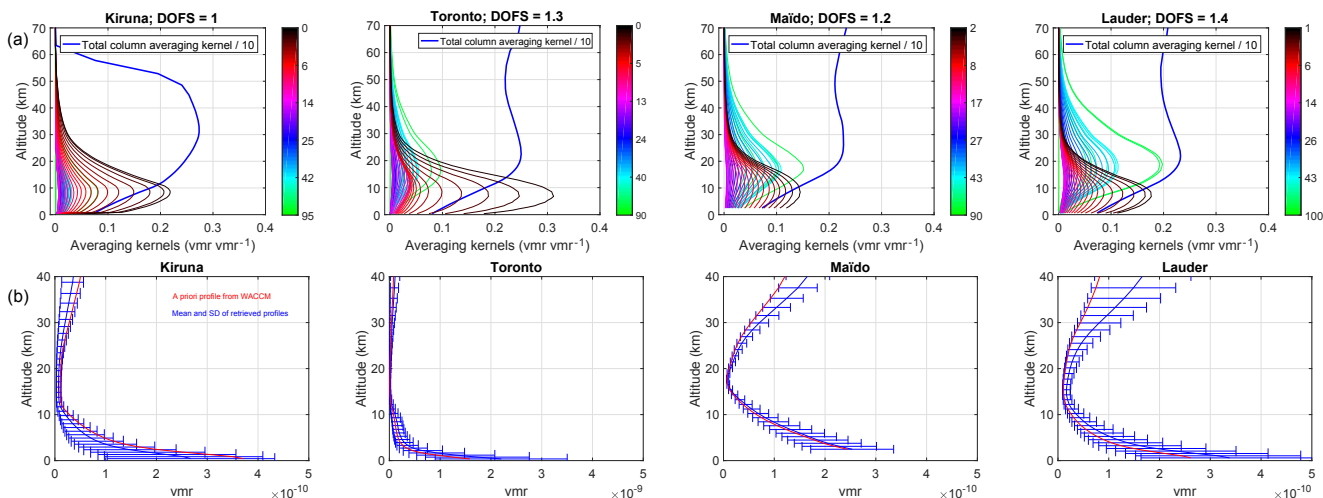
Figure 4. (a) Averaging kernels (rows of **A**) and total column averaging kernel for four of the FTIR stations, with DOFS ranging from 1.0 to 1.4. The total column averaging kernel is also shown in a thick blue line (divided by 10 for visibility). The colour code for the different averaging kernels depending on their altitude is given in the colour bar in kilometres. **(b)** A priori profiles from the WACCM v4 model (red), and the mean and standard deviation of the retrieved profiles for the same four stations.

the noise in the spectra and not a regularization one as can be chosen in the retrieval process (as in Eq. 1; see also Sect. 2.2). For the HCHO spectra used in this study, this SNR can vary between 100 for the worst cases and 3000, with a mean of about 700–1000 for the Bruker 120/5 HR instruments and 500 for the Bomem DA8.

The forward-model parameters error covariance matrices $\mathbf{S}_f$ are calculated according to the following:

$$\mathbf{S}_f = (\mathbf{G}_y \mathbf{K}_b) \mathbf{S}_b (\mathbf{G}_y \mathbf{K}_b)^T, \tag{3}$$

where $\mathbf{S}_b$ is the covariance matrix of $\boldsymbol{b}$, the vector of forward-model parameters. For each individual forward-model parameter, the $\mathbf{K}_b$ sensitivity matrix is mostly calculated by using analytic derivatives, while the covariance matrix $\mathbf{S}_b$ is an estimate of the uncertainty in the model parameter itself.

Effort has been made in this study to harmonize the uncertainty budget at all sites. This is done by calculating the errors from the same forward-model parameters (solar zenith angle, temperature, spectroscopic line parameters, baseline, etc.) across the network and by choosing the same $\mathbf{S}_b$ matrix for relevant parameters (i.e. when they are not site or instrument dependent, e.g. for the spectroscopic line parameters). However, some differences remain between the SFIT4 and PROFFIT codes that result in small differences that still occur between the two groups of users, despite the use of harmonized parameters. For the SFIT4 users, the random uncertainty given in Table 3 is dominated by the measurement noise (Eq. 2). We see from Table 3 that the random error is between 1.0 and 3.6×10^{14} molec cm^{-2} for the SFIT4 stations equipped with the high-resolution Bruker spectrometers 120/5 HR or M (the higher values coming from the 120/125 M instruments at Paramaribo and

PortoVelho), while it can reach 5.1×10^{14} molec cm^{-2} with the Bomem DA8 in Toronto. For the PROFFIT users, the random uncertainty is calculated to be a little bit larger (from 3.5 to 5.3×10^{14} molec cm^{-2}) for the sites with high-resolution spectrometers, and 11.1×10^{14} molec cm^{-2} with the low-resolution spectrometer Bruker Vertex 80 at Mexico City. The main difference between SFIT4 and PROFFIT is the additional error calculated at the PROFFIT stations due to the channelling of the spectra. However, in Table 3 we also give the mean differences between two subsequent FTIR measurements taken within 30 min (Diff30) as an upper limit of the total random uncertainty: this difference can be larger than the error budget if HCHO has faster variability than 30 min, but with enough statistics, the mean differences should not be lower than the total random errors. We see that this empirical upper estimation of total random uncertainty has a median value (2.8×10^{14} molec cm^{-2}) very close to the median total random uncertainty obtained by error propagation theory (2.9×10^{14} molec cm^{-2}), which gives confidence in the overall FTIR error estimation. At all the PROFFIT sites, except the highly polluted one (Mexico city), the total random uncertainty is larger than the Diff30, which could be an indicator that the uncertainty calculated in PROFFIT is slightly too conservative, probably due to this additional channelling error that would be estimated to be too large. For SFIT4 users, the Diff30 values are usually close, within 0.5×10^{14} molec cm^{-2} of the calculated total random uncertainty, with the exceptions of Ny-Ålesund and Lauder, where the small calculated errors of 1.9 and 1.6×10^{14} molec cm^{-2} might be a little bit optimistic, and with the exceptions of Toronto, Wollongong, and Paramaribo, where differences of

7 to 13×10^{14} molec cm^{-2} are observed between the Diff30 values and the total random errors.

After the measurement noise error (and the channelling for PROFFIT users), the largest contributions to the random uncertainty due to the forward-model parameters come from the temperature, the interfering species, and the off-set baseline. For temperature, the $\mathbf{S}_b$ matrix has been estimated using the differences between an ensemble of NCEP and sonde temperature profiles at Reunion Island, leading to 2 to 4 K in the troposphere and 3 to 6 K in the stratosphere. This matrix is currently used by all SFIT4 users, while for the PROFITT users, these chosen values are smaller (1 K in the troposphere, 2 K up to the middle–upper stratosphere, and 5 K for the highest levels). For each interfering species, the associated $\mathbf{S}_b$ matrix is the covariance matrix obtained with the WACCM v4 climatology. At some stations, the ILS also has a high contribution to the random error budget.

If one uses the FTIR HCHO measurements to validate a model or a satellite with a fine vertical resolution, considering the random and systematic uncertainties (without smoothing) in Table 3 (4th and 7th columns) is sufficient to make correct comparisons, because the smoothing error due to the low vertical resolution of the FTIR measurements vanishes if one takes into account the FTIR averaging kernels and a priori profile in the comparisons (Rodgers and Connor, 2003). However, if one wants to have a better knowledge of the real precision of the FTIR data themselves, this smoothing uncertainty can be estimated for the random part using the smoothing error covariance $\mathbf{S}_s^{\text{rand}}$ (Rodgers, 2000):

$$\mathbf{S}_s^{\text{rand}} = (\mathbf{I} - \mathbf{A})\mathbf{S}_{\text{var}}(\mathbf{I} - \mathbf{A})^T, \qquad (4)$$

where $\mathbf{S}_{\text{var}}$ should represent the natural variability of the target molecule. For HCHO, this $\mathbf{S}_{\text{var}}$ variability matrix is unfortunately not well known due to the poor number of vertically resolved measurements. In Table 3, the smoothing errors have been calculated taking the covariance matrices obtained using the WACCM v4 profiles at each station as an approximation of the $\mathbf{S}_{\text{var}}$ matrices. However, models usually underestimate the variability, and we expect that the smoothing errors provided here may be underestimated, especially in locations where HCHO is expected to have stronger vertical gradient variability than in the model. As an example, in the study by Vigouroux et al. (2009), the $\mathbf{S}_{\text{var}}$ was taken from aircraft measurements PEM-Tropics-B, and led to a smoothing error estimation of 14 % at Saint-Denis, while the present estimation based on the WACCM model gives about only 2 % for this station. However, the $\mathbf{S}_{\text{var}}$ matrix constructed from PEM-Tropics-B showed from 33 % to 70 % of HCHO variability which seems too high compared to what is observed at Reunion Island from the FTIR measurements (about 20 %). This illustrates that, ideally, the $\mathbf{S}_{\text{var}}$ matrix should be re-evaluated at the sites whenever better model data or correlative measurements become available. Since the FTIR data sets always include their associated averaging kernel matri-

ces, this re-evaluation can be done a posteriori by future users using Eq. (4).

The smoothing systematic uncertainty, reflecting the bias that would occur on the retrieved profile if the a priori $\mathbf{x}_a$ is biased compared to the real expected profile $< \mathbf{x} >$, is calculated following von Clarmann (2014):

$$\mathbf{S}_s^{\text{syst}} = (\mathbf{I} - \mathbf{A})(\mathbf{x}_a - < \mathbf{x} >)(\mathbf{x}_a - < \mathbf{x} >)^T(\mathbf{I} - \mathbf{A})^T. \qquad (5)$$

The $\mathbf{x}_a - < \mathbf{x} >$ is obviously not known (otherwise, $< \mathbf{x} >$ would be chosen as the correct a priori in the retrievals). Therefore, we have chosen to use $\mathbf{x}_a - < \mathbf{x} > = -50\%$, -20%, -10%, $+10\%$, $+8\%$ and $+5\%$ for the ground–4, 4–8, 8–13, 13–25, 25–40 and 40–120 km layers, respectively. The values have to vary with altitude to induce a different a priori profile shape: if 50 % is used at all altitudes, the a priori profile is then different from $< \mathbf{x} >$ by a simple scaling factor, and the systematic smoothing error is close to zero. Using the above values, we obtain smoothing systematic errors from 1 to 9 % (median value of 3.4 %), which is small compared to the other systematic error sources (Table 3). These values assume that the model WACCM profile shapes are not too far from reality, which should be the case: due to the known short lifetime of HCHO and its production at or near the surface, we expect that the mean profile peaks at the ground. This is, as for the random smoothing part, only an estimate of the smoothing systematic error. As discussed in von Clarmann (2014), one would even prefer to not give these smoothing errors at all. We prefer to give them to provide the reader with at least an idea of the impact of the smoothing in the precision and accuracy of our FTIR HCHO measurements. However, when making model or instrument comparisons, the appropriate use of the averaging kernel and a priori profile information, following Rodgers and Connor (2003), allows the user to implicitly take the smoothing uncertainty into account. This means that, for satellite or model comparison, if the methodology of Rodgers and Connor (2003) is used, there cannot be some different systematic biases at different stations due to different $\mathbf{x}_a - < \mathbf{x} >$.

The dominating systematic uncertainty sources are the spectroscopic parameters: the line intensities and the pressure broadening coefficients of the absorption lines present in our MWs. For the HCHO spectroscopic parameters, the line list in atm16 follows HITRAN 2012 (Rothman et al., 2013), which used the work of Jacquemart et al. (2010), and we use 10 % for the three parameters (line intensity, air-, and self-broadening coefficients). The larger error source is then the HCHO line intensity parameter and to a lesser extent the HCHO air-broadening coefficient. In addition, the uncertainties in HCHO columns due to the interfering species spectroscopic parameters are calculated. The dominant ones were found to be due to the pressure broadening coefficients of CH_4, HDO, and N_2O, with an order of magnitude of about 20 % of the error due to the HCHO line intensity.

The other systematic error sources due to forward-model parameters are lower or within a few percent (ILS, tempera-

ture), except for the PROFFIT channelling source (from 7 % to 17 %), which also has a systematic component. We see from Table 3 that the total systematic uncertainty is between 12 % and 15 % at the SFIT4 stations. For the PROFFIT stations, it lies between 12 % and 27 %.

3 Complete FTIR individual HCHO column data sets

3.1 A network sampling very low to highly polluted levels of HCHO

In Fig. 5 we show the individual HCHO total columns obtained at each station for a single year only (2016, except for Saint-Denis: 2011), in order to better see the day-to-day variability. The complete time series at each station are shown in the Supplement (Fig. S1). The error bars in Figs. 5 and S1 are the total random uncertainties; i.e. we do not include the systematic errors in order to better visualize the precision of the FTIR measurements compared to the observed day-to-day variability. The FTIR network samples a wide range of concentrations. Indeed, we can first distinguish the "clean" sites (shown with the same vertical axis with maximum 15×10^{15} molec cm^{-2}) such as the Arctic stations (Eureka, Ny-Ålesund, Thule, Kiruna, Sodankyla), the marine stations (Izaña, Mauna Loa, Maïdo, Saint-Denis, and Lauder, the former three also being high-altitude stations), and the high-mountain sites (Zugspitze and Altzomoni). These clean sites can have HCHO concentrations at the limit of detection (few 10^{13} molec cm^{-2}) with mean values of 10–25 × 10^{14} molec cm^{-2} (Table 3), except for Saint-Denis, which reaches a mean of 39×10^{14} molec cm^{-2}.

Second, we show the intermediate concentration sites (with the same vertical axis with maximum 30×10^{15} molec cm^{-2}) such as the tropical coastal site Paramaribo and the midlatitude polluted sites in or close to cities and/or vegetation (Peterhof close to St Petersburg, Bremen, Paris, Boulder). These intermediate sites have mean HCHO total columns of 58–73 × 10^{14} molec cm^{-2}. The sites with the highest levels of HCHO (vertical axis with maximum 70×10^{15} molec cm^{-2}) are Toronto and Mexico City, where large anthropogenic emissions are indeed expected (means of 95 and 221×10^{14} molec cm^{-2}), and places which are also affected by large biogenic emissions such as Wollongong (mean of 79×10^{14} molec cm^{-2}) and the new site of Porto Velho, located at the edge of the Amazon rainforest (mean of 190×10^{14} molec cm^{-2}).

3.2 HCHO diurnal cycle

As explained in the introduction, to reconcile the different results obtained using satellites observing at different times (e.g. SCIAMACHY and GOME-2 measuring in the morning and OMI in the afternoon), it is crucial to have ground-based observations of the HCHO diurnal cycles (Barkley et al., 2013; De Smedt et al., 2015; Stavrakou et al., 2015). The

diurnal cycle is also important for model validation, since emissions, chemistry and other processes depend on the time of the day. Our FTIR data set is now able to provide the diurnal cycles at 21 different locations. To separate the effect of the strong seasonal cycle (shown in the next section), we give the diurnal cycle in four different seasons in Fig. 6 for a selection of the sites, while the other ones are provided in the Supplement (Fig. S2). As seen from Figs. 6 and S2, the diurnal cycles are often site and season dependent. While there is no clear diurnal cycle at the Arctic sites and at some of the midlatitude cities during winter (St Petersburg, Bremen, Toronto), we usually see an increase from the morning, which is often more pronounced in June–July–August (and December–January–February in the Southern Hemisphere) at most of the stations (in the cities but also at marine sites). A maximum is often found around midday (St Petersburg, Mexico City, Izaña, Saint-Denis, and Wollongong) or much later in the afternoon (16:00–18:00 local time, LT), as in Bremen, Paris, Toronto, Lauder, and Altzomoni. Only in a few cases is a minimum found at midday (St Petersburg in SON, Zugspitze in MAM–SON, Sodankyla in MAM). The marine sites at high altitudes (free of local pollution) have a small minimum at about 08:00 LT (Izaña, Maïdo). This diversity in the FTIR diurnal cycles is also observed with MAX-DOAS at other sites (De Smedt et al., 2015): a very weak diurnal cycle at OHP (southern France) in winter and spring; a minimum around midday at Beijing and Xianghe in spring and autumn, and a constant increase in summer (as observed with FTIR for Bremen, Toronto, and Paris). The diurnal cycles observed at the Jungfraujoch station from FTIR measurements (Franco et al., 2015) show, for all months of the year, a midday maximum, which is very different from the ones observed at our closest station Zugspitze. The IMAGES model shows diurnal cycles in phase agreement with our FTIR measurements at Zugspitze, except for the summer, for which the model diurnal cycle is very weak (not shown). However, two sites very close together can indeed observe different diurnal cycles (as seen for Saint-Denis and Maïdo). More investigation is needed to understand the different diurnal cycles at these two mountain sites.

We see from Fig. 6 that the FTIR measurements at Porto Velho do not show a clear pattern, in particular if one is interested in the 09:30 and 13:30 LT differences between the overpasses of two different satellites (De Smedt et al., 2015). From the 1-year data available at present at this new site, it seems that the diurnal cycle cannot help reconciling the differences observed over Rondônia between GOME-2 and OMI (De Smedt et al., 2015). In contrast, the diurnal cycles observed over cities confirm that the observation of a positive bias between OMI (13:30 LT) and GOME-2 (09:30 LT) over urban areas can be indeed explained, at least partly, by the diurnal cycle.
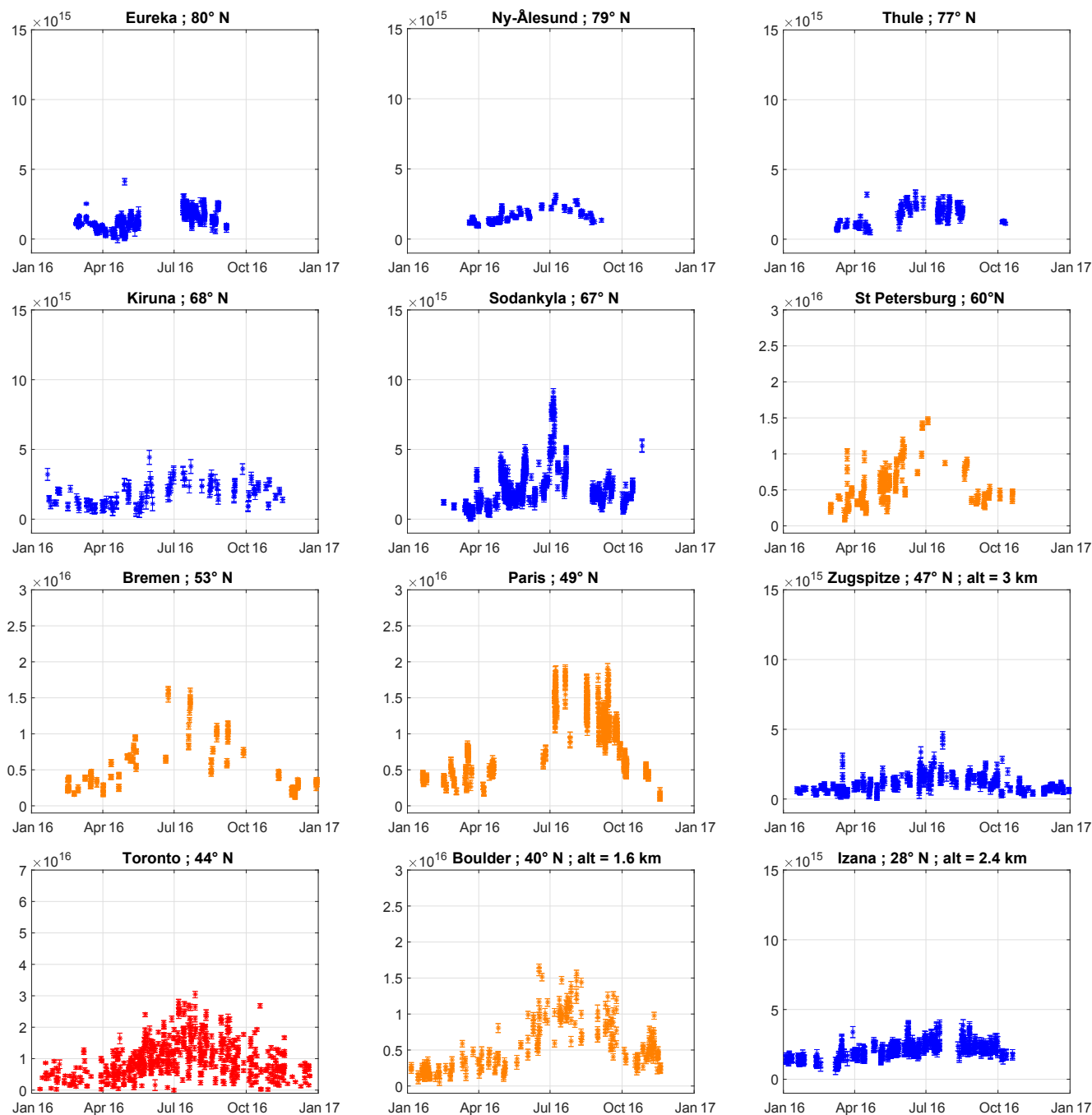
Figure 5.

3.3 Long-term HCHO trends

The length of the HCHO time series allows trends to be derived for some stations. We have calculated the trends at each station using the monthly mean time series $Y_m(t)$ with a simple model, including a fit of the seasonal cycles:

$$\begin{aligned}
Y_m(t) = {} & A_0 + A_1 \cdot \cos(2\pi t/12) + A_2 \cdot \sin(2\pi t/12) \\
& + A_3 \cdot \cos(4\pi t/12) + A_4 \cdot \sin(4\pi t/12) \\
& + A_5 \cdot t,
\end{aligned}$$

with A_5 as the annual trend.

It turned out that, due to the very high variability of HCHO, the uncertainties in the trends are often too large
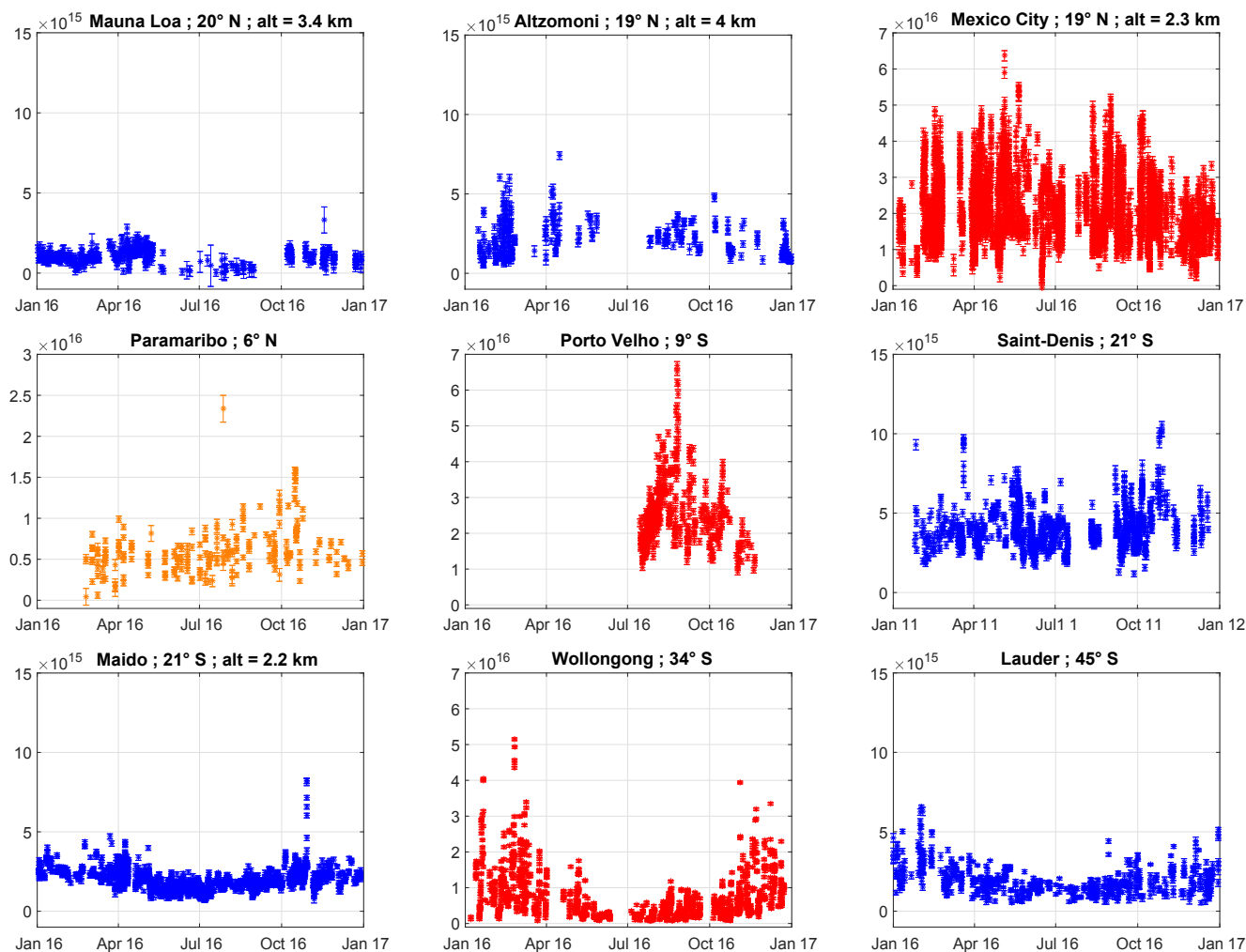
Figure 5. Overview of the individual HCHO total columns (molec cm^{-2}) at each station for a single year (2016, except for Saint-Denis: 2011). The complete time series at each station are shown in the Supplement (Fig. S1). The clean, intermediate, and high-level HCHO sites are shown using blue, orange, and red colours. The error bars are the total random uncertainty. When the altitude of the station is higher than 1.5 km, it is explicitly shown.

to obtain significant values. A more sophisticated multi-regression model might be able to reduce the uncertainties, but this is beyond the scope of this paper. However, for a few stations, significant trends are found. They are mainly negative: at St Petersburg ($-3.9 \pm 3.3\%$ decade^{-1}), Mexico City ($-9.6 \pm 5.1\%$ decade^{-1}), Wollongong ($-18.8 \pm 10.8\%$ decade^{-1}), and close to significance at Zugspitze ($-7.7 \pm 7.7\%$ decade^{-1}). Only the marine sites Izaña and Saint-Denis show positive significant trends ($+17.3 \pm 15.2$ and $+15.8 \pm 5.2\%$ decade^{-1}). Note that at Maïdo, the trend is not significant. A careful combination of the measurements at both Reunion Island sites (Saint-Denis and Maïdo) could be carried out in the future.

For the longest time series, we observe a very good agreement with previous studies in general. The negative trends observed over the European stations St Petersburg

and Zugspitze are in agreement with the negative trends observed by OMI (2004–2014) over St Petersburg and Germany (De Smedt et al., 2015). At the Jungfraujoch station, a negative trend ($-6.1 \pm 2.6\%$ decade^{-1}) was also observed for the 1996–2015 period (Franco et al., 2016). Note that the calculation of the uncertainties in the trends in our study takes into account the autocorrelation in the residuals, following Santer et al. (2000), which increases the uncertainties. For Zugspitze, the uncertainty without correcting for this autocorrelation, as in Franco et al. (2016) or De Smedt et al. (2015), would be 4.9% (instead of 7.7%), showing then a more significant trend, in agreement with these studies. The non-significant trends observed at the northern European station (Kiruna), and the midlatitude American stations (Toronto, Boulder) are in agreement with De Smedt et al. (2015). In the Southern Hemisphere, the negative trend
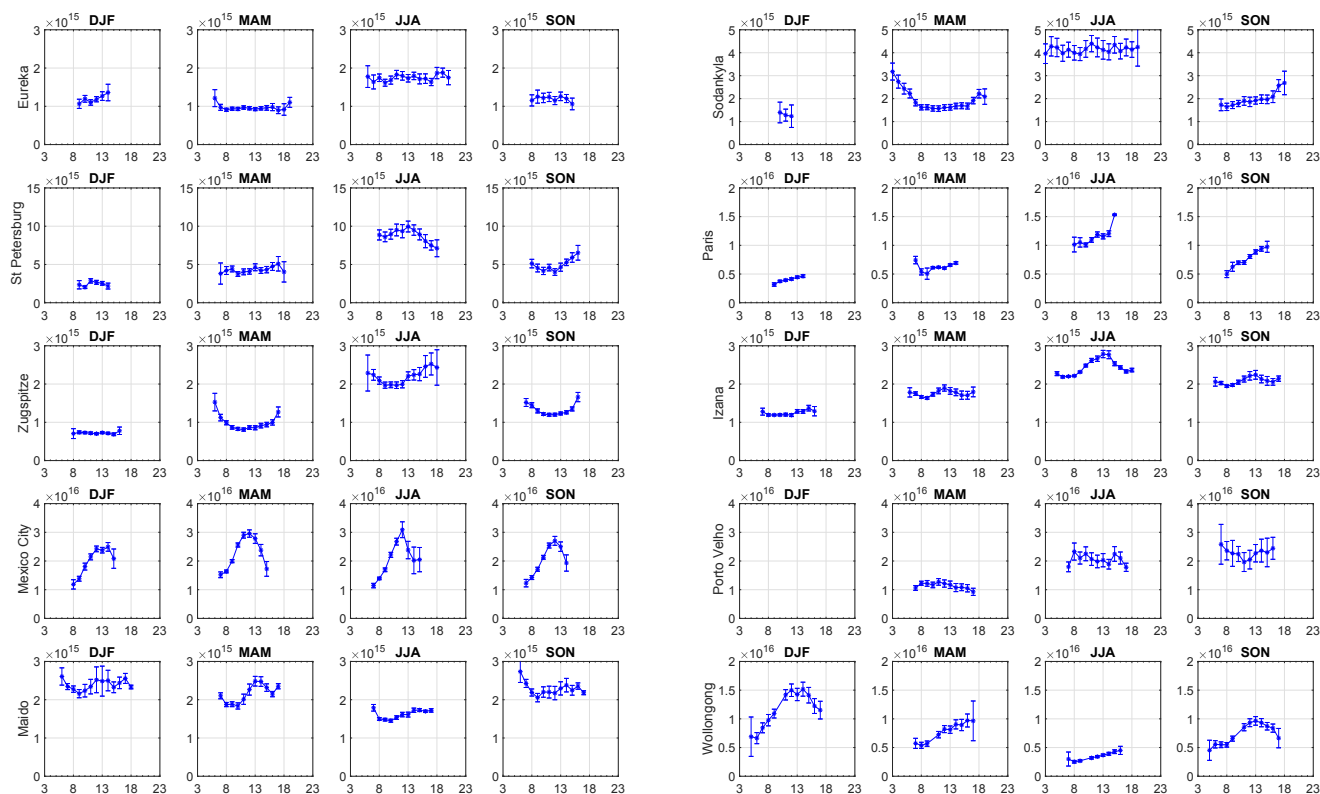
Figure 6. Diurnal cycles of HCHO total columns (molec cm^{-2}) at selected stations for the four seasons. The diurnal cycles for the other stations are shown in the Supplement (Fig. S2). The error bars are the standard errors of the mean: $2 \times \sigma/\sqrt{n}$, with σ the standard deviation and n the number of measurements at a given time. If $n < 8$, the hourly value is not shown. The time is the local standard time meridian (LSTM).

observed at Wollongong was also found in De Smedt et al. (2015), as well as a positive trend at Madagascar, which is near Reunion Island, in agreement with the high positive trend observed at Saint-Denis. At Lauder, OMI also shows a non-significant trend (De Smedt et al., 2015).

4　HCHO FTIR and IMAGES model comparisons

In this study, we do not aim to validate the model input parameters or attribute different emission sources at the different stations. We use the model to assess the internal consistency of the network using harmonized retrieval settings. This means that we expect that, for the same latitude regions and/or type of sites (polluted; clean), the comparisons with the model will give consistent biases. In the Supplement we provide a global map of IMAGES climatological daytime HCHO columns (2005–2015) together with the mean columns observed at the FTIR stations (Fig. S3). This map illustrates the very different levels covered by the FTIR stations and the overall good agreement with the calculated levels of IMAGES. However, Fig. S3 can only provide a qualitative comparison due to the different measurement periods

covered. We give quantitative comparisons in the present section.

4.1　IMAGES model description

The IMAGESv2 global model calculates the distribution of 170 chemical compounds gases with a time step of 6 h at $2° \times 2.5°$ resolution, with 40 hybrid (σ pressure) levels in the verticals between the surface and the lower stratosphere (44 hPa level). The model calculates daily averaged concentrations of chemical compounds. The effect of diurnal variations is accounted for through correction factors on the photolysis and kinetic rates obtained from a full diurnal cycle simulation using a time step of 20 min. The same model simulation also stores the diurnal shapes of formaldehyde columns required for the comparison with FTIR data on files. Meteorological fields (winds, temperature, humidity, 3-dimensional cloud cover, solid and liquid cloud water content, large-scale and convective precipitation rates, visible downward radiation, convective updraught fluxes, boundary layer diffusivities, snow depth, sea ice fraction, surface roughness lengths, surface sensible heat flux, friction velocity, etc.) are obtained from ERA-Interim analyses of

the European Centre for Medium-range Weather Forecasts (ECMWF).

Anthropogenic emissions of NO_x, CO, SO_2, and NMVOC are provided by the Hemispheric Transport of Air Pollution data set version 2 (HTAPv2) (Janssens-Maenhout et al., 2015), with the NMVOC speciation provided by the emission inventory of the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (Lamarque et al., 2010). Emissions from open vegetation fires are taken from the last version of the Global Fire Emissions Database, GFED4s, which includes the contribution of small fires (Randerson et al., 2012; Giglio et al., 2013). The GFED data are available at daily frequency at $0.25° \times 0.25°$ from 1997 to the present (http://www.globalfiredata.org, last access: 5 September 2018). The vertical distribution of these emissions follows Sofiev et al. (2013). Isoprene and monoterpene emissions are obtained from the MEGAN-MOHYCAN model (Müller et al., 2008; Stavrakou et al., 2014; Guenther et al., 2012) for all years of the study period at a resolution of $0.5° \times 0.5°$ (http://tropo.aeronomie.be/models/isoprene.htm, last access: 5 September 2018). Methanol biogenic emissions are obtained from the inverse modelling study of Stavrakou et al. (2011). Besides the dependence on temperature, visible radiation, leaf area, and leaf age, the model accounts for the inhibition of isoprene emissions under drought conditions through a dimensionless soil moisture activity factor (γ_{SM}). However, the parameterization of γ_{SM} is very uncertain, as discussed in Bauwens et al. (2016), and we assume $\gamma_{SM} = 1$ in this study. The average global annual emissions are 419 Tg yr^{-1} isoprene, 100 Tg yr^{-1} methanol, and 103 Tg yr^{-1} monoterpenes.

The chemical degradation mechanism of pyrogenic NMVOCs is described in Bauwens et al. (2016). The oxidation mechanism for isoprene is also based on Bauwens et al. (2016), with a few updates. It accounts for the revised kinetics of isoprene peroxy radicals according to the Leuven Isoprene Mechanism version 1 (LIM1) (Peeters et al., 2014) and is further modified to account for laboratory findings (Teng et al., 2017; Bates et al., 2016). The formaldehyde yield in isoprene oxidation by OH is close to 2.4 mol mol^{-1} in high NO_x (1 ppbv NO_2, after 2 months of simulation) and 1.9 mol mol^{-1} at low NO_x (0.1 ppbv NO_2). The chemical mechanism for monoterpenes is simplified, with product yields of formaldehyde, acetone, methylglyoxal and glyoxal based on box model calculations using the α- and β-pinene oxidation mechanism from the Master Chemical Mechanism (MCM) (Saunders et al., 2003). The overall formaldehyde yield is 4.2 HCHO per monoterpene oxidized, reducing to 2.3 after subtracting the contributions of acetone and methylglyoxal oxidation. This yield is further reduced by 45 % to account for wet deposition of intermediate and secondary organic aerosol formation. This fraction of 45 % is higher but of the same order as the estimated overall impact of deposition on the average HCHO yield from isoprene oxidation (28 %), based on IMAGES model calculations. The higher fraction

for monoterpenes is intended to account for the impact of the more complex chemistry and larger number of oxygenated intermediates involved in their oxidation compared to isoprene. The large deposited fraction is uncertain but appears justified by the larger number and lower volatility of intermediates involved in formaldehyde formation from monoterpene oxidation.

The calculation of the model columns at the FTIR stations accounts for its location in the horizontal (nearest model pixel) for the FTIR a priori profiles and averaging kernels as prescribed in Rodgers and Connor (2003), as well as for the station altitude above sea level. The model column is calculated from the calculated formaldehyde profile, between the altitude of the station and the uppermost model level (approximately 20 km), and from the a priori FTIR profile, above that level. When the model surface lies higher than the station, the model column is increased by a partial column assuming a constant mixing ratio between the two altitudes, taken equal to the value at the lowermost model level. The monthly averaged formaldehyde columns are calculated by accounting for the temporal sampling of the observations at each site and month. Also, the local time of each observation is taken into account by rescaling the daily averaged concentration using the formaldehyde diurnal shape factors calculated by the model with a time step of 20 min.

4.2 HCHO monthly means and seasonal cycle comparisons

We compare the monthly means of FTIR HCHO total columns at each station with the IMAGES columns calculated for the 2003–2016 period. The time series of both products are shown in Fig. 7. Since the random uncertainty of the FTIR monthly means is divided by the square root of the number of measurements within each month, the dominant contribution to the FTIR error bars in Fig. 7 is the systematic uncertainty (estimated at 11–26 %. The smoothing uncertainty is not included in model comparisons using Rodgers and Connor, 2003). Except for very few cases (Mexico City and Paramaribo), the model is in overall good agreement with the FTIR measurements in terms of absolute levels (Fig. 7) and seasonal cycle (Fig. 8).

For each station the correlation, the bias and the standard deviation (SD) of the statistical comparisons between the monthly means, mean(IMAGES(smoothed) − FTIR)/mean(FTIR), are summarized in Table 4. The median correlation between FTIR and IMAGES for the 21 stations is very high (0.81), with weaker values at the Mexican stations (0.4/0.5) and at Mauna Loa (0.10). The median standard deviation for all comparisons is 25 % (ranging from 11 % to 41 %). This agreement is good considering the FTIR variability (i.e. the SD) of HCHO monthly means (median of 35 %). The standard deviation of the comparisons can be explained partly by the lower variability of the model monthly means (31 %) compared to FTIR, as seen in Fig. 7. In addi-
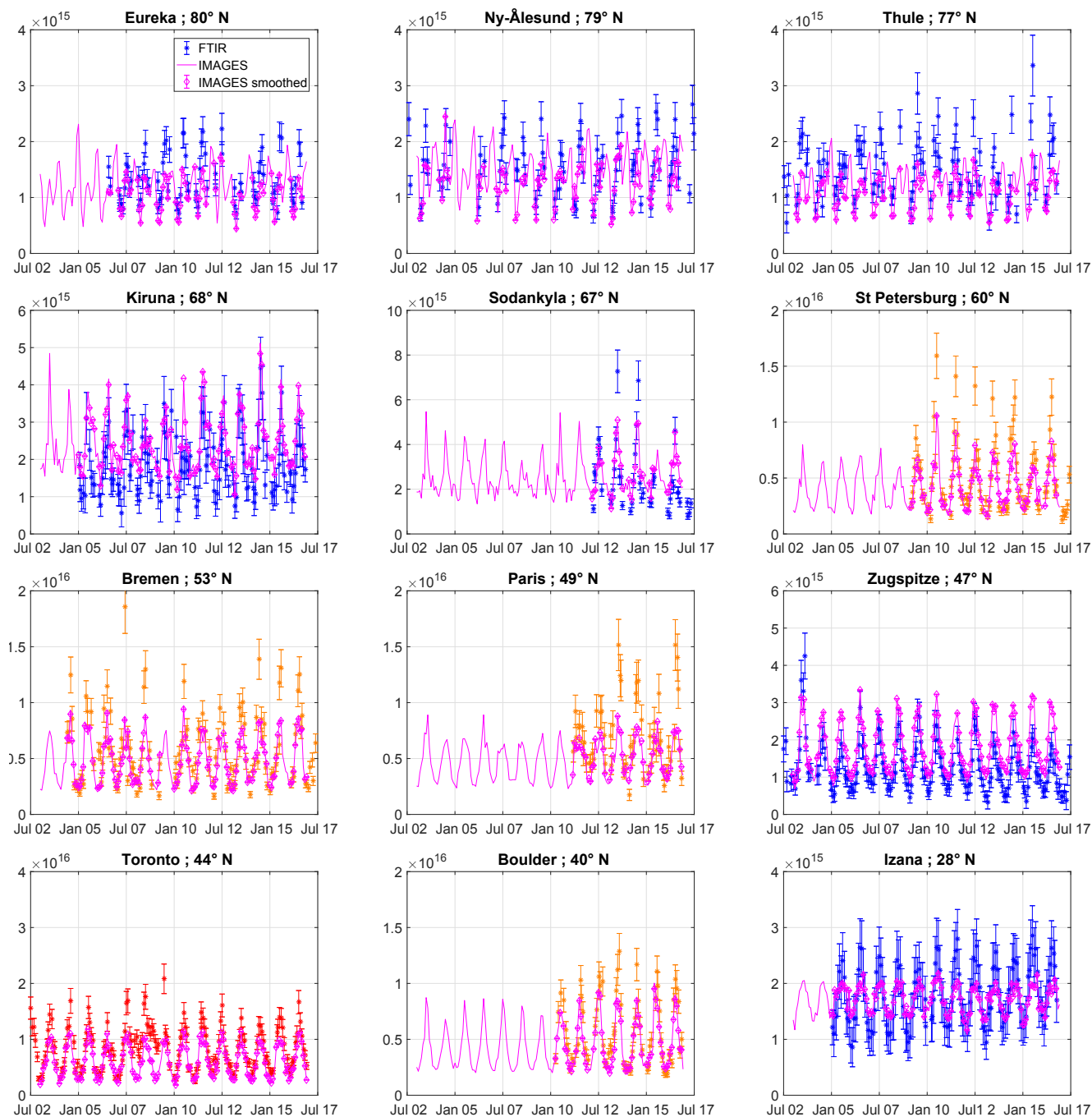
Figure 7.

tion, the variability of the model data within a month is also much smaller (median of about 11 %; this SD within a month is shown as magenta error bars in Fig. 7) than the FTIR one (median of about 28 %).

The median of IMAGES and FTIR differences is small (−15 %) and within the FTIR systematic uncertainty estimated at 11 %–26 %. However, the biases range from −64 % to +51 %, which requires an investigation of their possible

causes. The main source of systematic uncertainty is the spectroscopic parameters, which have been harmonized in this work, with each station using the same line parameters database, and the same spectral MWs. Therefore, it is expected that all FTIR stations should provide consistent HCHO total columns within 5 %–17 % (systematic errors due to other sources than spectroscopic ones). To check this,
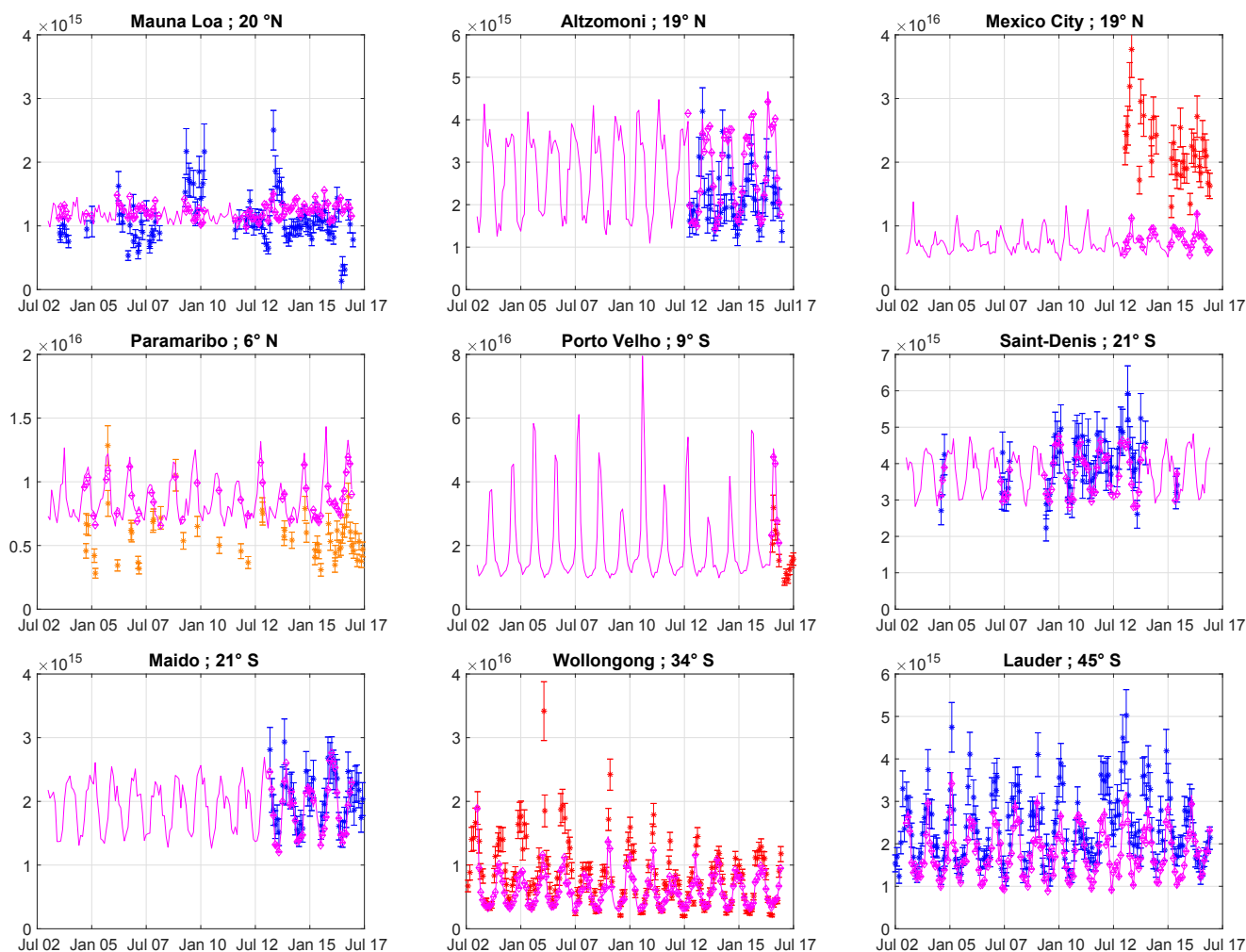
Figure 7. Monthly means of HCHO total columns (molec cm^{-2}) at each station for FTIR measurements are shown with stars (clean, intermediate, and high-level HCHO sites are shown using blue, orange, and red colours) and model data (magenta line for raw model data; magenta diamonds for model data smoothed by the FTIR AK). The FTIR error bars represent the total uncertainties in monthly means which, due to monthly averaging, are mainly the systematic uncertainties. The model error bars represent the standard deviation of the model for each month.

we divide the FTIR stations according to their concentrations levels and latitudes, and use the model for comparisons.

4.2.1 Clean Arctic sites

We distinguish two groups of Arctic sites, Eureka, Ny-Ålesund and Thule, which are very remote (77–80° N), and two European sites, Kiruna and Sodankyla (67–68° N). As seen in Table 4, the former group shows similar negative biases in the model compared to the data (−20 %/−17 %/−28 %), while the latter group shows positive biases (+32 %/+11 %). Except at Kiruna, the biases are not constant through the year, with the model showing less pronounced seasonal cycles (see also Fig. 8). The model underestimates the summer HCHO levels at the three 77–80° N stations (−26 %/−20 %/−28 %), while the winter levels are

in close agreement (+6 %/−3 %). At the 67–68° N stations, the model is positively biased in winter (+27 %/+56 %), as well as in summer in Kiruna (+22 %). Note that the Arctic sites do not record measurements during polar nights, so the winter months basically correspond to February (Fig. 8).

4.2.2 Midlatitude cities

Very similar biases (−16 %/−15 %/−22 %) between IMAGES and FTIR are obtained at the three European cities, St Petersburg (the site is actually at Peterhof, a small coastal city at about 30 km west of St Petersburg), Bremen, and Paris. As for the Arctic sites, the model underestimates the amplitude of the seasonal cycle (Fig. 8), leading to smaller biases in winter (−14 % to −17 %) compared to summer (−19 % to −30 %).
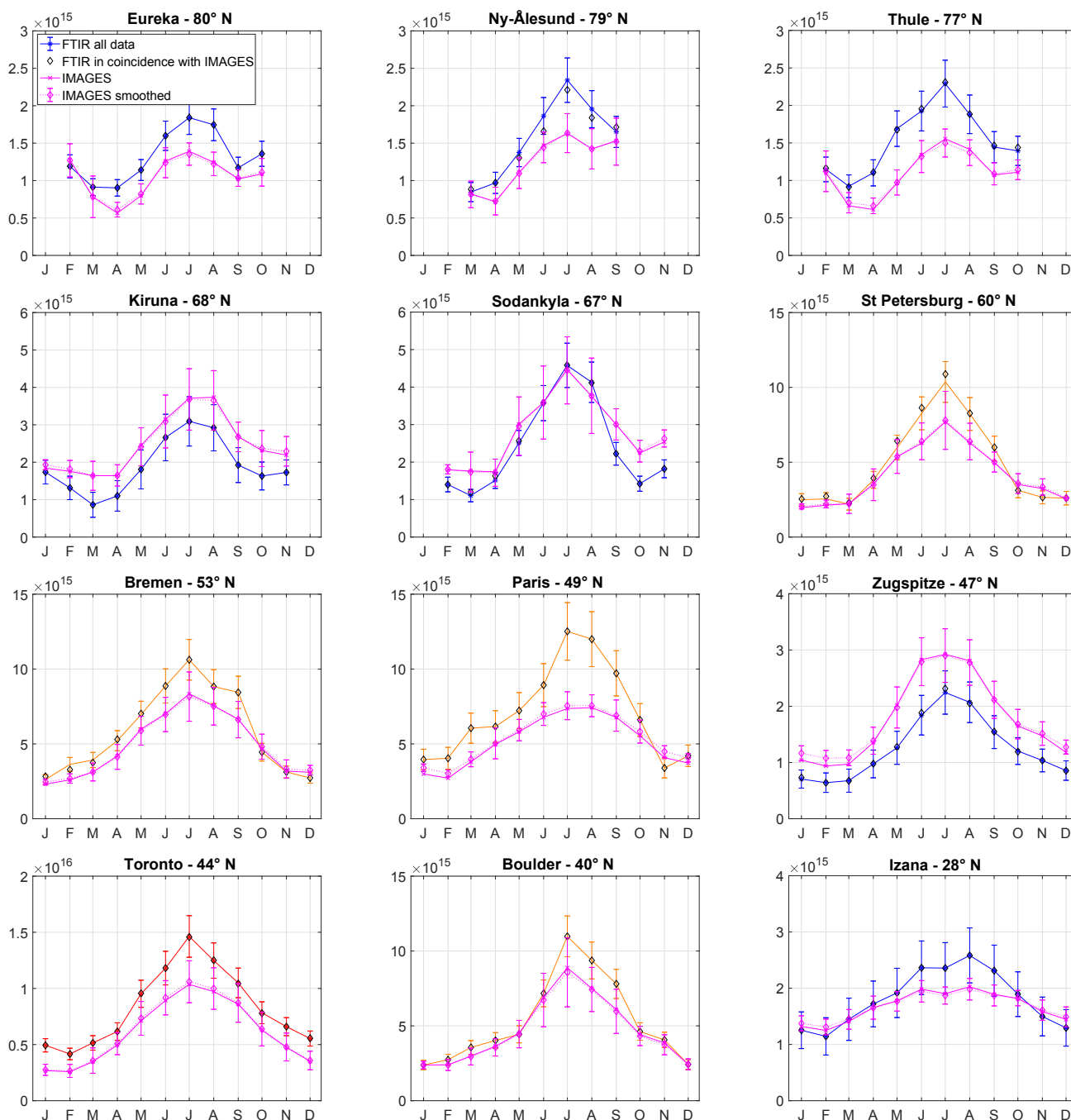
Figure 8.

The North American sites Toronto and Boulder give similar biases ($-26\%/-17\%$), especially in summer ($-25\%/-17\%$). Toronto is the only midlatitude urban site where the model shows a higher underestimation of the HCHO levels in winter (-39%).

4.2.3　High-mountain sites

The mountain sites are more difficult to model, especially when they are close to cities. They are often very clean sites, but the model cannot reproduce this at the current resolution ($2° \times 2.5°$) when they are surrounded by emission sources in the same pixel. This seems to be the case at Altzomoni, which lies in the same model pixel as Mexico City, lead-
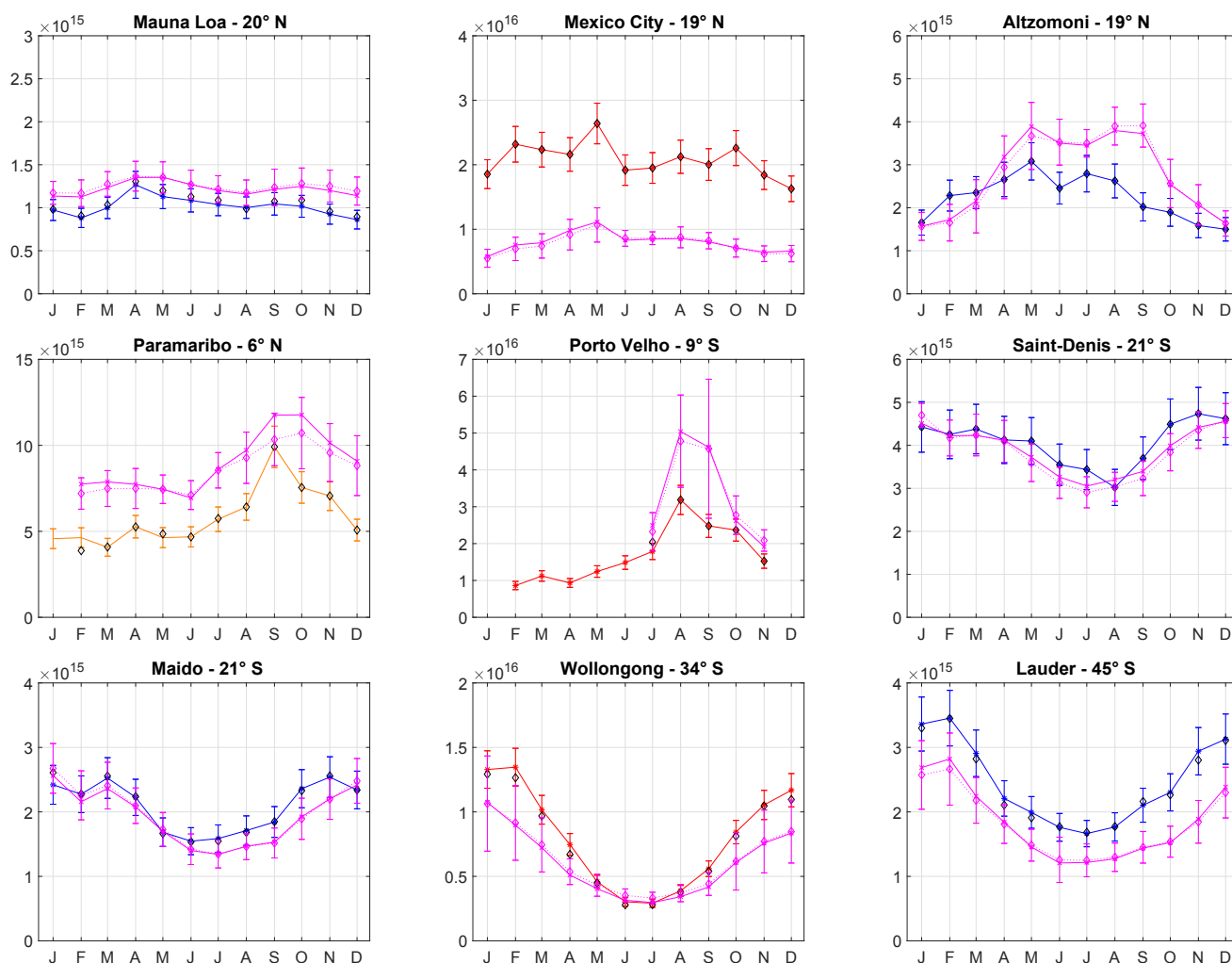
Figure 8. Seasonal cycle of HCHO total columns (molec cm^{-2}) at each station for FTIR measurements (clean, intermediate, and high-level HCHO sites are shown using blue, orange, and red stars when all FTIR data are used; black diamonds correspond to the seasonal cycles when only data in coincidence with the model are used) and model data (magenta line for raw model data; magenta diamonds for model data smoothed by the FTIR AK). The FTIR error bars mainly represent the systematic uncertainties. The model error bars represent the standard deviation of the model for each month. Only the model data in coincidence with FTIR measurements are taken into account in these seasonal cycles.

ing to an overestimation of 26 %, much larger in summer (+49 %), and at the European station Zugspitze, where the model overestimates the HCHO levels by +41 %. Note that, in the study of Franco et al. (2015), a negative bias (−13 %) was observed between FTIR at Jungfraujoch (47° N, 8° E) and IMAGES, but the retrieval settings used were different than in the present study. Only a change in the spectroscopic database, from HITRAN 2008 to HITRAN 2012, led to lower HCHO columns by 49 % at Jungfraujoch (Franco et al., 2015). It is therefore not possible at present to compare the biases obtained at these two close stations.

At the mountain site of Izaña, located in a clean marine area, the model and FTIR are in overall good agreement (−3 %), with a negative bias in summer (−19 %) and a positive one in winter (+22 %), as a result of the weak seasonal amplitude in the model.

A moderate positive model bias is calculated at Mauna Loa (+13 %), which is more pronounced in winter (+24 %), and a good agreement is seen between the model and FTIR mean seasonal cycle (Fig. 8). The observed variability (34 %) is, however, important at this site and similar to the clean Arctic sites (Fig. 7), with values ranging from 0.5 to 2.5 × 10^{15} molec cm^{-2}. This is not reproduced by the model values lying within 1–1.5 × 10^{15} molec cm^{-2}. The causes of the pronounced observed variability are unclear at present.

Table 4. Correlation (Corr), bias $\pm$ standard deviation (SD_{stat}) of the statistical comparisons between the monthly means, mean(IMAGES(smoothed) $-$ FTIR)/mean(FTIR). Also given are the mean of the standard deviations in the IMAGES and the FTIR monthly means, i.e. the variability within a month (SD_m), and the standard deviation of the whole FTIR and IMAGES monthly mean time series (SD_{all}). All numbers, except the correlations, are given in %.

Station	Corr	bias $\pm$ SD$_{stat}$	bias $\pm$ SD$_{stat}$	bias $\pm$ SD$_{stat}$	SD$_m$ IMAGES / FTIR	SD$_{all}$ IMAGES / FTIR
	All	All	JJA	DJF	Within a month	All
Eureka	0.77	-20 ± 21	-26 ± 15	$+6 \pm 22$	10/28	28/32
Ny-Ålesund	0.72	-17 ± 23	-20 ± 18	–	9/25	30/33
Thule	0.74	-28 ± 24	-28 ± 23	-3 ± 18	9/31	28/35
Kiruna	0.80	$+32 \pm 27$	$+22 \pm 20$	$+27 \pm 37$	10/28	31/44
Sodankyla	0.85	$+11 \pm 33$	-4 ± 27	$+56 \pm 35$	12/34	37/60
St Petersburg	0.94	-16 ± 29	-25 ± 20	-14 ± 23	12/32	43/60
Bremen	0.87	-15 ± 30	-19 ± 27	-16 ± 40	8/20	42/56
Paris	0.84	-22 ± 29	-30 ± 25	-17 ± 40	6/19	30/45
Zugspitze	0.87	$+41 \pm 26$	$+32 \pm 24$	$+59 \pm 24$	10/31	37/51
Toronto	0.88	-26 ± 23	-25 ± 16	-39 ± 44	15/40	46/47
Boulder	0.93	-17 ± 22	-17 ± 15	-13 ± 32	12/24	47/52
Izaña	0.81	-3 ± 20	-19 ± 9	$+22 \pm 15$	8/18	14/29
Mauna Loa	0.10	$+13 \pm 35$	$+14 \pm 45$	$+24 \pm 35$	9/28	9/34
Mexico City	0.45	-64 ± 21	-59 ± 17	-66 ± 26	17/37	18/23
Altzomoni	0.43	$+26 \pm 41$	$+49 \pm 22$	-6 ± 22	16/42	35/29
Paramaribo	0.67	$+51 \pm 25$	$+59 \pm 15$	$+85 \pm 17$	12/35	17/33
			DJF	JJA		
Porto Velho	0.87	$+41 \pm 35$	–	$+35 \pm 35$	24/27	39/26
Saint-Denis	0.71	-7 ± 13	-3 ± 12	-9 ± 15	9/27	16/18
Maïdo	0.87	-7 ± 11	$+3 \pm 7$	-14 ± 7	13/20	23/20
Wollongong	0.83	-26 ± 37	-29 ± 34	-3 ± 35	18/50	43/59
Lauder	0.77	-25 ± 22	-24 ± 17	-26 ± 26	11/31	31/35
			Summer	Winter		
Median	0.81	-15 ± 25	-19 ± 19	-5 ± 26	11/28	31/35

4.2.4 Central and South American sites

The model falls short in reproducing the enhanced HCHO levels observed at Mexico City (ca.2×10^{16} molec cm^{-2}), mainly due to the coarse-model resolution ($2° \times 2.5°$), as suggested by the strong negative bias (-64 %), which is almost constant across the year.

Comparison at two sites in South America, the coastal site of Paramaribo and the Porto Velho site at the edge of the Amazon rainforest, indicates a consistent model overestimation ($+51$ %/$+41$ %). At Porto Velho, this overestimation is more significant during the dry season (August–September, Fig. 8), which corresponds to the maximum of fire intensity in Amazonia. An overestimation of biogenic (isoprene) and biomass-burning emissions in Amazonia was already found in IMAGES in the study of Bauwens et al. (2016).

4.2.5 Southern Hemisphere 21–45° S sites

The two marine sites at Reunion Island (Saint-Denis at sea level, and Maïdo at 2.2 km altitude) show a small model bias (-7 %) and standard deviation, especially at Maïdo (11 %). At these sites, HCHO shows the lowest variability in the monthly means (18–20 %), and the model reproduces the seasonal cycle quite well. As shown in Fig. 8, the largest seasonal bias is not found in austral summer (DJF) as seen in the Northern Hemisphere sites, but during September–November months, which correspond to the maximum of the biomass-burning period in southern Africa and Madagascar, close to Reunion Island. The biomass-burning source at this location might be underestimated, whereas it was overestimated in South America.

The Wollongong site shows the same behaviour as most of the Northern Hemisphere sites: an overall underestimation of the model (-26 %), which is larger in austral summer (-29 %). A first look at the Lauder comparison gives a similar annual bias (-25 %), which remains constant over the year, as seen in Table 4 and Fig. 8. However, Fig. 7 shows that, during the austral winters (JJA) 2012 to 2015, the FTIR time series presents unusually high columns. By limiting the comparison to the first years of the period, a better agreement

with the model in winter is obtained at Lauder, as for many other sites.

Since the time series at Saint-Denis, Wollongong and Lauder have been published in the past using different retrieval strategies (Vigouroux et al., 2009; Jones et al., 2009; Zeng et al., 2015), here we report the bias observed at these stations between the previous and present data sets. The bias is only 1.4 % at Saint-Denis, the present HCHO columns being smaller than the previous data set in Vigouroux et al. (2009), in which the a priori profile and the spectroscopy were different (mostly for interfering species, the HCHO spectroscopic intensity parameters being from the same work of Perrin et al., 2009), and the MWs were smaller than in the present work. Therefore, the comparisons with MAX-DOAS shown in Vigouroux et al. (2009) would still provide good agreement between the two techniques. Concerning Lauder and Wollongong, where the previous retrieval strategy was from Jones et al. (2009), the present HCHO columns are 49 % smaller than the previous data sets. Therefore, the new data set is in much closer agreement with the simulation of four different models that were all found 50 % lower than the old Lauder and Wollongong data sets (Zeng et al., 2015). From sensitivity tests, this high bias between the two strategies is very likely mostly due to the 2869.65–2870.0 cm^{-1} window used in Jones et al. (2009).

5 Conclusions

Only five NDACC FTIR sites have delivered HCHO time series until now (Paton-Walsh et al., 2005; Jones et al., 2009; Vigouroux et al., 2009; Viatte et al., 2014; Franco et al., 2015), using different retrieval settings. The small number of stations and the differences in bias associated with the different retrieval strategies made it difficult to use the FTIR network as a coherent tool for satellite or model validation. In this study, we have designed a harmonized HCHO retrieval strategy to derive total columns at 21 stations, at locations characterized by very different concentrations, from very clean Arctic sites where HCHO is at the limit of detection (a few 10^{13} molec cm^{-2}) to highly polluted sites such Mexico City or Porto Velho, near the Amazon rainforest, where columns up to 7×10^{16} molec cm^{-2} have been observed. This network includes well-established NDACC stations, as well as several new sites (Sodankyla, Boulder, Paris, Porto Velho) that aim to be affiliated with NDACC. The FTIR network is also growing, with new sites such as Hefei in China, which will again expand its spatial coverage.

We have presented the retrieval settings that have been optimized for this challenging species, and the FTIR HCHO products have been characterized by their averaging kernels and their uncertainty budget. The systematic uncertainty of an individual HCHO total column measurement lies between 12 % and 27 %, with some differences remaining between the SFIT4 code users (12 %–15 %) and the PROFFIT users

(12 %–27 %), which needs to be investigated in the future within the NDACC InfraRed Working Group. The random uncertainty lies between 1 and 11×10^{14} molec cm^{-2}, with a median value of 2.9×10^{14} molec cm^{-2}. The high maximum value is due to the lower quality of the Bruker Vertex compared to the high-resolution ones (Bruker 120/5 M or 120/5 HR).

In addition to the well-defined seasonal cycles, the diurnal cycles were presented at each site. These observations are crucial for interpreting the differences observed between satellites measuring at different local times. For example, the diurnal cycle at Porto Velho which shows insignificant variations suggests that the negative bias observed over Rondônia between OMI (13:30 LT) and GOME-2 (09:30 LT) (De Smedt et al., 2015) is unlikely due to the diurnal cycle. In contrast, the FTIR diurnal cycles in the cities confirm that the positive bias between OMI and GOME-2 over urban areas is likely due, at least partly, to the diurnal cycle.

The monthly mean time series as well as the seasonal cycles have been compared to the IMAGES model. We did not aim to evaluate the model but show that the FTIR network provides coherent absolute values and seasonal cycles. We observed an overall good agreement with IMAGES, which usually (but not always) underestimated the HCHO total columns (median bias ± standard deviation of $-15\% \pm 25\%$), with a more pronounced bias during summer ($-19\% \pm 19\%$). The similar biases obtained at stations under similar conditions (clean Arctic sites, urban sites, marine sites) strengthen our confidence in the harmonization of the HCHO products within the network. When the model showed different behaviours for some of the stations, we could explain it by either the oversized model pixel ($2.0° \times 2.5°$), especially for high-altitude sites such as Zugspitze, Altzomoni, and Mexico City, or an overestimation of the biogenic and biomass-burning sources in South America (Paramaribo, Porto Velho), which was already pointed out in Bauwens et al. (2016). However, for a few sites, the behaviour of the model remained unexplained (positive biases at Kiruna and Sodankyla, the too-low model variability at Mauna Loa).

These HCHO time series, harmonized and well characterized, provide an important data set for past and present satellites, and model validation. They are continuously extended by new measurements and will be used in the coming years for the validation of new satellites, such as Sentinel 5P and Sentinel 4.

Competing interests. The authors declare that they have no conflict of interest.

Edited by: Michel Van Roozendael
Reviewed by: three anonymous referees

References

Barkley, M. P., De Smedt, I., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E., and Mao, J.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, J. Geophys. Res.-Atmos., 118, 6849–6868, https://doi.org/10.1002/jgrd.50552, 2013.

Bates, K. H., Nguyen, T. B., Teng, A. P., Crounse, J. D., Kjaergaard, H. G., Stoltz, B. M., Seinfeld, J. H., and Wennberg, P. O.: Production and fate of C4 dihydroxycarbonyl compounds from isoprene oxidation, J. Phys. Chem. A, 120, 106–117, 2016.

Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., 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, 10133–10158, https://doi.org/10.5194/acp-16-10133-2016, 2016.

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

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., Hendrick, F., Van Roozendael, M., Müller, J.-F., Stavrakou, T., Marais, E. A., Bovy, B., Bader, W., Fayt, C., Hermans, C., Lejeune, B., Pinardi, G., Servais, C., and Mahieu, E.: Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations, Atmos. Meas. Tech., 8, 1733–1756, https://doi.org/10.5194/amt-8-1733-2015, 2015.

Franco, B., Marais, E. A., Bovy, B., Bader, W., Lejeune, B., Roland, G., Servais, C., and Mahieu, E.: Diurnal cycle and multi-decadal trend of formaldehyde in the remote atmosphere near 46° N, Atmos. Chem. Phys., 16, 4171–4189, https://doi.org/10.5194/acp-16-4171-2016, 2016.

Garcia, R. R., Marsh, D. R., Kinnison, D. E., Boville, B. A., and Sassi, F.: Simulation of secular trends in the middle atmosphere, 1950–2003, J. Geophys. Res., 112, D09301, https://doi.org/10.1029/2006JD007485, 2007.

Giglio, L., Randerson, J. T., and Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), J. Geophys. Res., 118, 317–328, https://doi.org/10.1002/jgrg.20042, 2013.

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, 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Hansen, P. C.: Analysis of discrete ill-posed problems by means of the L-curve, SIAM Review, 34, 561–580, 1992.

Hase, F., Blumenstock, T., and Paton-Walsh, C.: Analysis of the instrumental line shape of high-resolution Fourier transform IR spectrometers with gas cell measurements and new retrieval software, Appl. Optics, 38, 3417–3422, 1999.

Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Höpfner, M., Jones, N. B., Rinsland, C. P., and Wood, S. W.: Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, J. Quant. Spectrosc. Ra., 87, 25–52, 2004.

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. Ra., 102, 450–463, 2006.

Hase, F., Wallace, L., McLeod, S. D., Harrison, J. J., and Bernath, P. F.: The ACE-FTS atlas of the infrared solar spectrum, J. Quant. Spectrosc. Ra., 111, 521–528, 2010.

Jacquemart, D., Laraia, A., Kwabia Tchana, F., Gamache, R. R., Perrin, A., and Lacome, N: Formaldehyde around 3.5 and 5.7 µm: measurement and calculation of broadening coefficients, J. Quant. Spectrosc. Ra., 111, 1209–1222, 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.

Kesselmeier, J. and Staudt, M.: Biogenic Volatile Organic Compounds (VOC): An Overview on Emission, Physiology and Ecology, J. Atmos. Chem., 33, 23–88, 1999.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmos. Chem. Phys., 10, 7017–7039, https://doi.org/10.5194/acp-10-7017-2010, 2010.

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.

Martin, R. V., Parrish, D. D., Ryerson, T. B., Nicks, D. K. Jr., Chance, K., Kurosu, T. P., Jacob, D. J., Sturges, E. D., Fried, A., and Wert, B. P.: Evaluation of GOME satellite measurements of tropospheric NO_2 and HCHO using regional data from aircraft campaigns in the southeastern United States, J. Geophys. Res., 109, D24307, https://doi.org/10.1029/2004JD004869, 2004.

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.

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

Notholt, J., Toon, G., Stordal, F., Solberg, S., Schmidbauer, N., Becker, E., Meier, A., and Sen, B.: Seasonal variations of atmospheric trace gases in the high Arctic at 79° N, J. Geophys. Res., 102, 12855–12861, https://doi.org/10.1029/97JD00337, 1997.

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

Paton-Walsh, C., Jones, N. B., Wilson, S. R., Haverd, V., Meier, A., Griffith, D. W. T. and, Rinsland, C. P.: Measurements of trace gas emissions from Australian forest fires and correlations with coincident measurements of aerosol optical depth, J. Geophys. Res., 110, D24305, https://doi.org/10.1029/2005JD006202, 2005.

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.

Perrin, A., Jacquemart, D., Kwabia Tchana, F., and Lacome, N.: Absolute line intensities measurements and calculations for the 5.7 and 3.6 µm bands of formaldehyde, J. Quant. Spectrosc. Ra., 110, 700–716, 2009.

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

Randerson, J. T., Chen, Y., van der Werf, G. R., Rogers, B. M., and Morton, D. C.: Global burned area and biomass burning emissions from small fires, J. Geophys. Res., 117, G04012, https://doi.org/10.1029/2012JG002128, 2012.

Rodgers, C. D.: Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics – vol. 2, World Scientific Publishing Co., Singapore, 2000.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res. 108, 4116–4129, 2003.

Rothman, L. S., Jacquemart, D., Barbe, A., Benner, D. C., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, C., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Flaud, J.-M., Gamache, R. R., Goldman, A., Hartmann, J. M., Jucks, K. W., Maki, A. G., Mandin, J. Y., Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J., Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P., and Wagner, G.: The HITRAN 2004 molecular spectroscopic database, J. Quant. Spectrosc. Ra., 96, 139–204, 2005.

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

Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coudert, L. H., Devi, V. M., Drouin, B. J., Fayt, A., Flaud, J.-M., Gamache, R. R., Harrison, J. J., Hartmann, J.-M., Hill, C., Hodges, J. T., Jacquemart, D., Jolly, A., Lamouroux, J., Roy, R. J. L., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T., Mikhailenk, S., Müller, H. S. P., Naumenko, O. V., and Nikitin, A. V.: The HITRAN2012 molecular spectroscopic database, J. Quant. Spectrosc. Ra., 130, 4–50, https://doi.org/10.1016/j.jqsrt.2013.07.002, 2013.

Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E., and Taylor, K. E.: Statistical significance of trends and trend differences in layer-average atmospheric temperature series, J. Geophys. Res., 105, 7337–7356, 2000.

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

Senten, C., De Mazière, M., Dils, B., Hermans, C., Kruglanski, M., Neefs, E., Scolas, F., Vandaele, A. C., Vanhaelewyn, G., Vigouroux, C., Carleer, M., Coheur, P. F., Fally, S., Barret, B., Baray, J. L., Delmas, R., Leveau, J., Metzger, J. M., Mahieu, E., Boone, C., Walker, K. A., Bernath, P. F., and Strong, K.: Technical Note: New ground-based FTIR measurements at Ile de La Réunion: observations, error analysis, and comparisons with independent data, Atmos. Chem. Phys., 8, 3483–3508, https://doi.org/10.5194/acp-8-3483-2008, 2008.

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.

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., Guenther, A., Wild, M., and Xia, X.: Isoprene emissions over Asia 1979–2012: impact of climate and land-use changes, Atmos. Chem. Phys., 14, 4587–4605, https://doi.org/10.5194/acp-14-4587-2014, 2014.

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.

Steck, T.: Methods for determining regularization for atmospheric retrieval problems, Appl. Optics, 41, 1788–1797, 2002.

Sussmann, R., Forster, F., Rettinger, M., and Jones, N.: Strategy for high-accuracy-and-precision retrieval of atmospheric methane from the mid-infrared FTIR network, Atmos. Meas. Tech., 4, 1943–1964, https://doi.org/10.5194/amt-4-1943-2011, 2011.

Teng, A. P., Crounse, J. D., and Wennberg, P.: Isoprene peroxy radical dynamics, J. Am. Chem. Soc., 139, 5367–5677, 2017.

Tikhonov, A.: On the solution of incorrectly stated problems and a method of regularization, Dokl. Acad. Nauk SSSR+, 151, 501–504, 1963.

Toon, G. C.: The JPL MkIV interferometer, Opt. Photonics News, 2, 19–21, 1991.

Viatte, C., Strong, K., Walker, K. A., and Drummond, J. R.: Five years of CO, HCN, C_2H_6, C_2H_2, CH_3OH, HCOOH and H_2CO total columns measured in the Canadian high Arctic, Atmos. Meas. Tech., 7, 1547–1570, https://doi.org/10.5194/amt-7-1547-2014, 2014.

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, https://doi.org/10.5194/acp-9-9523-2009, 2009.

Vigouroux, C., Blumenstock, T., Coffey, M., Errera, Q., García, O., Jones, N. B., Hannigan, J. W., Hase, F., Liley, B., Mahieu, E., Mellqvist, J., Notholt, J., Palm, M., Persson, G., Schneider, M., Servais, C., Smale, D., Thölix, L., and De Mazière, M.: Trends of ozone total columns and vertical distribution from FTIR observations at eight NDACC stations around the globe, Atmos. Chem. Phys., 15, 2915–2933, https://doi.org/10.5194/acp-15-2915-2015, 2015.

von Clarmann, T.: Smoothing error pitfalls, Atmos. Meas. Tech., 7, 3023–3034, https://doi.org/10.5194/amt-7-3023-2014, 2014.

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

Zeng, G., Williams, J. E., Fisher, J. A., Emmons, L. K., Jones, N. B., Morgenstern, O., Robinson, J., Smale, D., Paton-Walsh, C., and Griffith, D. W. T.: Multi-model simulation of CO and HCHO in the Southern Hemisphere: comparison with observations and impact of biogenic emissions, Atmos. Chem. Phys., 15, 7217–7245, https://doi.org/10.5194/acp-15-7217-2015, 2015.

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., González 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.