Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data

C. Vigouroux\textsuperscript{1}, F. Hendrick\textsuperscript{1}, T. Stavrakou\textsuperscript{1}, B. Dils\textsuperscript{1}, I. De Smedt\textsuperscript{1}, C. Hermans\textsuperscript{1}, A. Merlaud\textsuperscript{1}, F. Scolas\textsuperscript{1}, C. Senten\textsuperscript{1}, G. Vanhaelewyn\textsuperscript{1}, S. Fally\textsuperscript{2,\ast}, M. Carleer\textsuperscript{2}, J.-M. Metzger\textsuperscript{3}, J.-F. Müller\textsuperscript{1}, M. Van Roozendael\textsuperscript{1}, and M. De Mazière\textsuperscript{1}

\textsuperscript{1}Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
\textsuperscript{2}Service de Chimie Quantique et de Photophysique (SCQP), Université Libre de Bruxelles (ULB), Brussels, Belgium
\textsuperscript{3}Laboratoire de l’Atmosphère et des Cyclones (LACy), Université de La Réunion, France

\textsuperscript{\ast}now at: Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

Received: 17 June 2009 – Accepted: 17 July 2009 – Published: 27 July 2009

Correspondence to: C. Vigouroux (corinne.vigouroux@aeronomie.be)

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

Formaldehyde (HCHO) columns have been retrieved from ground-based Fourier transform infrared (FTIR) campaign measurements in 2004 and 2007 and from UV-Visible MAX-DOAS measurements in 2004–2005 at the NDACC site of Réunion Island (21°S, 55°E). The FTIR and MAX-DOAS daily mean formaldehyde total columns are inter-compared in their common measurement period, from August to October 2004. The ground-based data are also compared to correlative SCIAMACHY data. The comparisons account for the vertical sensitivity differences of the data sets, by including their respective averaging kernels. Complete error budgets are also presented.

The FTIR and MAX-DOAS daily mean total columns agree very well: no significant bias is observed and the standard deviation of the comparisons is only 8%. Both FTIR and MAX-DOAS HCHO total columns are in good agreement with SCIAMACHY values in the 2004–2005 period, with standard deviations of 21% and 31%, respectively. The same seasonal cycle is observed by the different instruments, with a minimum in austral winter and a maximum in February–March.

The FTIR and MAX-DOAS data are confronted with HCHO columns calculated by a global CTM, the IMAGES model. The model underestimates the HCHO columns by 23–29% in comparison with FTIR, and by 15% in comparison with DOAS. This bias might have multiple causes, including an underestimation of OH concentrations in the model (as indicated by a sensitivity study using prescribed OH fields) and/or an underestimated contribution of large-scale transport of HCHO precursors from Madagascar. The latter hypothesis is comforted by the large observed day-to-day variability of HCHO columns, and by the observation that the peak values of FTIR columns can often be associated with free tropospheric transport patterns from source regions over Madagascar to Réunion Island, according to simulations performed with the Lagrangian particle dispersion model FLEXPART.
1 Introduction

The main sources of formaldehyde (HCHO) in the atmosphere are the photochemical oxidation of methane and non-methane volatile organic compounds (NMVOCs). The sinks of formaldehyde are photolysis, oxidation by OH and dry and wet deposition (Stavrakou et al., 2009a). Due to its short lifetime of only a few hours, its global distribution closely resembles the distribution of its sources. Therefore, over land, observations of formaldehyde provide new constraints on the emissions of reactive NMVOCs (in particular isoprene), as demonstrated by several inverse modeling studies using satellite retrievals of HCHO (e.g. Abbot et al., 2003; Stavrakou et al., 2009b). Far away from the emission regions, e.g. over oceans, formaldehyde observations might provide an opportunity to test our current knowledge regarding methane oxidation, and possibly also to quantify the effect of long-range transport of NMVOCs from source regions.

Formaldehyde total columns have been observed from space by GOME and SCIAMACHY (Abbot et al., 2003; De Smedt et al., 2008; Dufour et al., 2009). Formaldehyde measurements by ground-based FTIR instruments have been reported at mid-latitude (Demoulin et al., 1999; Jones et al., 2007) and polar stations (Notholt et al., 2000), and during a ship cruise in the central Atlantic Ocean (Notholt et al., 2000). MAX-DOAS formaldehyde campaign measurements have been performed in the Italian Po-Valley in summer 2002 (Heckel et al., 2005).

In the present work, we provide the first time series of ground-based FTIR and MAX-DOAS formaldehyde observations in the tropics. The observation site is Réunion Island (21° S, 55° E) in the Indian Ocean. This station is part of the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndacc.org), a network dedicated to performing high-quality long-term observations of atmospheric trace gases at globally distributed sites. The Réunion station is one of the very few NDACC stations located at southern tropical or subtropical latitudes. In preparation to a permanent installation planned for 2011, three campaigns of FTIR measurements have been performed, in October 2002, from August to October 2004, and from May to November
A UV-visible Multi-Axis DOAS (MAX-DOAS) instrument was operated at the same site from August 2004 to July 2005. The inversion algorithms for both the MAX-DOAS and FTIR spectral data analyses use the Optimal Estimation Method (Rodgers, 2000) to derive information about the vertical distribution of the target gases. Since the number of independent pieces of vertical information for HCHO is low for both instruments, only total column results are presented and discussed in this paper. For the first time, formaldehyde columns retrieved by these two different remote sensing techniques are mutually compared. Whereas Jones et al. (2007) presented comparisons of ground-based FTIR HCHO columns with GOME data, the present paper gives the first comparisons of ground-based FTIR measurements with SCIAMACHY data. In addition, the retrievals are compared with the results of a global CTM, the IMAGESv2 model which has been recently used in an inverse modeling study of NMVOC emissions based on SCIAMACHY data (Stavrakou et al., 2009b).

Sections 2 and 3 describe the FTIR and MAX-DOAS measurements and retrievals of formaldehyde, respectively, including a detailed error budget in each case. Sections 4 and 5 briefly summarize the formaldehyde measurements by SCIAMACHY and the simulations by the CTM IMAGESv2, respectively. In Sect. 6, we compare these formaldehyde datasets, accounting for the vertical sensitivity differences among the instruments by the appropriate use of averaging kernels. In Sect. 7, we discuss the main features of formaldehyde time series at Réunion Island, including the origin of the observed variabilities.

2 Formaldehyde from FTIR observations

2.1 Measurements campaigns

A Bruker 120M Fourier transform infrared (FTIR) spectrometer has been deployed during three campaigns at Réunion Island, in October 2002, from August to October 2004, and from May to November 2007. For the campaigns of 2004 and 2007, the FTIR in-
instrument was located at Saint-Denis (20.9° S, 55.5° E, 50 m a.s.l.). It was operated in an automatic and remotely controlled way by use of BARCOS (Bruker Automation and Remote COntrol System) developed at BIRA-IASB (Neefs et al., 2007). More detailed specifications of the 2002 and 2004 experiments are given in Senten et al. (2008); the 2007 experiment was conducted in an almost identical way. In the present work, we will not include results from the first campaign in 2002, because (i) it represents only one month of data, and (ii) no MAX-DOAS or SCIAMACHY data are available for that year.

The FTIR solar absorption measurements are performed in a wide spectral range (around 600–4500 cm\(^{-1}\)), which allow the retrieval of many species (Senten et al., 2008). The formaldehyde spectra are recorded in the 2400–3310 cm\(^{-1}\) domain. At low solar zenith angles (SZA), the spectral resolution is 0.00513 cm\(^{-1}\) and 5 scans are co-added. At higher SZA, the spectral resolution is reduced to 0.00893 cm\(^{-1}\), and the number of co-added scans to 3. The switch from high to low resolution occurs when the airmass factor changes by more than 7% during the time needed to record the spectra; this occurs around 60° SZA. By reducing the resolution and the number of scans, we effectively shorten the recording time of the spectra.

The volume mixing ratio profiles of target gases are retrieved from the shapes of their absorption lines, which are pressure and temperature dependent. Daily pressure and temperature profiles have been taken from the National Centers for Environmental Prediction (NCEP). The observed absorption line shapes also depend on the instrument line shape (ILS) which is therefore included in the forward model of the retrieval code. In order to know the ILS and at the same time to monitor the alignment of the instrument, a reference low-pressure (2 hPa) HBr cell spectrum was recorded at local noon with the sun as light source, whenever the meteorological conditions permitted so. The software LINEFIT is used for the analysis of the cell spectra, as described in Hase et al. (1999). In this approach, the complex modulation efficiencies are described by 40 parameters (20 for amplitude and 20 for phase orientation) at equidistant optical path differences.
2.2 FTIR retrieval strategy

The FTIR retrievals are performed using the algorithm SFIT2 (Rinsland et al., 1998), version 3.92, jointly developed at the NASA Langley Research Center, the National Center for Atmospheric Research (NCAR) and the National Institute of Water and Atmosphere Research (NIWA). The spectral inversion is based on a semi-empirical implementation of the Optimal Estimation Method (OEM) of Rodgers (2000), in order to solve the ill-posed problem.

The retrieved vertical profiles are obtained by fitting one or more narrow spectral intervals (microwindows). The vertical information content of the retrieved profiles depends strongly on the choice of microwindows and a priori information (the a priori profile \(x_a\) and the regularization matrix \(R\)). It can be quantified by the number of degrees of freedom for signal (DOFS), which is the trace of the so-called averaging kernel matrix \(A\), defined in Rodgers (2000) by:

\[
A = \frac{\partial \hat{x}}{\partial x} = (K^T S_e^{-1} K + R)^{-1} K^T S_e^{-1} K,
\]

where \(\hat{x}\) and \(x\) are the retrieved and the true state vectors, respectively, \(K\) is the weighting function matrix (\(K = \partial y / \partial x\), with \(y\) the measurement vector), \(S_e\) is the measurement noise covariance matrix, and \(R = S_a^{-1}\), with \(S_a^{-1}\) the a priori covariance matrix in the OEM (Rodgers, 2000).

2.2.1 Choice of microwindows and spectroscopic databases

The choice of microwindows is critical for a species like formaldehyde which has very weak absorptions in the infrared. In particular, it is important to minimize the impact of interfering species absorbing in the same spectral region. Total columns of formaldehyde have been retrieved already from different ground-based FTIR studies: Demoulin et al. (1999) have used microwindows centered at 2806.6 cm\(^{-1}\), 2831.5 cm\(^{-1}\) and 2833.2 cm\(^{-1}\) for the Jungfraujoch station. Formaldehyde total columns were also
measured at the ground-based stations of Mc-Murdo and Ny-Ålesund, located in the Antarctic and the Arctic respectively (Notholt et al., 1997), in seven microwindows with central wavenumbers at 2760.9, 2774.8, 2778.4, 2781.2, 2798.0, 2869.9, 2914.5 cm$^{-1}$. Notholt et al. (2000) have given latitudinal variations of HCHO total columns obtained during a ship cruise, using the 2869.44–2870.33 cm$^{-1}$ microwindow. More recently, Jones et al. (2007) measured long-term formaldehyde columns at Lauder (45° S) using a set of seven microwindows centered around 2713.85 (HDO line), 2778.5, 2780.8, 2856.2 (solar CO line), 2869.85, 2912.2, and 2914.65 (NO$_2$ line) cm$^{-1}$.

Table 1 gives the list of microwindows used in this work. The selection of microwindows is not trivial at the location of Saint-Denis because there is a considerable amount of water vapour. We therefore have to avoid as much as possible interferences with water vapour lines. The microwindows with the strongest HCHO signal are the ones at 2778 and 2780 cm$^{-1}$, already used in Notholt et al. (1997) and Jones et al. (2007). The two first microwindows at 2763 and 2765 cm$^{-1}$ were chosen for their relatively small absorptions from interfering species. These four microwindows (2763, 2765, 2778 and 2780 cm$^{-1}$) are part of the microwindows set for HCHO advised by Meier et al. (2004). The 2810 cm$^{-1}$ microwindow was indicated by the LINEFINDING tool of Notholt et al. (2006), together with the 2778 and 2780 cm$^{-1}$ ones. Other microwindows indicated by LINEFINDING with even larger priority than the 2810 cm$^{-1}$ one have been discarded because, either they contain water vapour lines or too strong HDO lines, or they contain some spectroscopic features that were badly fitted in the retrieval. The last microwindow at 2855 cm$^{-1}$ is added to get additional information on the HDO and solar absorptions, and so to reduce the correlation between the HDO and solar signals and the HCHO one in the other microwindows.

The profiles of the interfering species H$_2$O, HDO, CH$_4$, and N$_2$O were retrieved beforehand and independently in the microwindows listed also in Table 1. The daily means of the H$_2$O and HDO retrieved profiles were used as daily a priori profiles in the retrievals of CH$_4$, and scaled using a single parameter for each species. Finally, the daily means of the retrieved profiles of the four molecules (H$_2$O, HDO, CH$_4$, and N$_2$O)
are used as daily a priori profiles in the retrievals of HCHO. For the interfering species having a small impact on the HCHO retrievals, a single a priori profile is used for all spectra. So for O$_3$, we have used the 5-years (1985–1989) climatology from UGAMP (http://badc.nerc.ac.uk/data/ugamp-o3-climatology) above Réunion Island. For the CO solar lines, we used the empirical line-by-line model of Hase et al. (2006). In the formaldehyde retrieval process, the a priori profile of each of the interfering species is scaled using a single parameter.

We have used, for all species except HCHO, the HITRAN 2004 spectroscopic line parameters (Rothman et al., 2005), with the additional official updates published on the HITRAN web-site (http://www.cfa.harvard.edu/hitran/). For formaldehyde, we have used the very recent linelist of Perrin et al. (2009).

### 2.2.2 Choice of a priori information

The a priori profile $x_a$, from the ground to 12 km, has been constructed from data composites of the airborne experiment PEM-Tropics-B (Raper et al., 2001). We used the average formaldehyde concentration over the Southern tropical Pacific (0 to 30° S; 160° E to 95° W) based on the data composites available at http://acd.ucar.edu/emmons/DATACOMP/camp_table.htm, which is an update of the database described in Emmons et al. (2000). Between 12 and 20 km, we have used a yearly mean of the formaldehyde profiles obtained from the IMAGESv2 model. For the upper stratosphere, the profile is based on the MIPAS-Envisat measurements (Steck et al., 2008), with a maximum of about 80 pptv around 40 km at 20° S latitude. The a priori profile adopted in the FTIR retrievals is shown in Fig. 1.

In the usual OEM, the constraint matrix $R$ is the inverse of the a priori covariance matrix $S_a$. Ideally, $S_a$ should express the natural variability of the target gas, and thus should be as realistic as possible and evaluated from appropriate climatological data (Rodgers, 2000). However, for formaldehyde at Réunion Island, this information is not available and therefore, we have chosen to use Tikhonov L1 regularization (Tikhonov, 1963), i.e., the constraint matrix is defined as $R=\alpha L_1^T L_1$, with $\alpha$ the regularization...
strength and \( L_1 \) the first derivative operator:

\[
L_1 = \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 \\
0 & -1 & 1 & \ddots & \\
& \ddots & \ddots & \ddots & 0 \\
0 & \ldots & 0 & -1 & 1
\end{pmatrix}
\]

In this approach, the difference between the retrieved profile and the a priori profile is constrained to be a constant profile. This regularization avoids the appearance of spurious oscillations in the retrieved formaldehyde profiles that occurred in the usual OEM implementation. This approach was adopted recently for \( \text{H}_2\text{O} \) (Sussmann et al., 2009) and, within an European FTIR network, for \( \text{CH}_4 \) (Sussmann et al., 2009).

For the determination of \( \alpha \), we have followed the method illustrated in Fig. 1 of Steck (2002) which shows that the obtained DOFS is determined by the value of \( \alpha \). Initial trials with the usual OEM approach indicated that the degrees of freedom for signal (DOFS) that can reasonably be achieved for our HCHO retrievals is about 1.1, implying a value of \( \alpha = 55 \).

### 2.2.3 Vertical information in FTIR retrievals

As mentioned previously, the vertical information contained in the FTIR retrievals can be characterized by the averaging kernel matrix \( A \) and its trace (DOFS). This matrix depends on measurement and retrieval parameters including the solar zenith angle, the spectral resolution and signal to noise ratio, the choice of spectral microwindows, the regularization matrix, . . . We obtain a mean DOFS of 1.1 and its standard deviation (1\( \sigma \)) for all measurements at different solar zenith angles is 0.1. The rows of \( A \) are the so-called averaging kernels and they represent the sensitivity of the retrieved profile to the real profile. We give in Fig. 2 the mean of the averaging kernels obtained in the formaldehyde retrievals. As expected with the DOFS close to one, we can see that the averaging kernels are not vertically resolved. They all peak at about the same altitude.
The retrieved profile is sensitive only to a change in the true profile that occurs between the ground and about 20 km, with a maximum of sensitivity around 10 km. It is useful to note that such a change between 0 and 20 km will also impact the retrieved profile above 20 km where we do not have sensitivity. This is due to the Tikhonov approach where the difference between the retrieved and the a priori profile is constrained to be a constant profile (see Sect. 2.2.2). However, the HCHO column above 20 km represents only about 1.5% of the total column, and this effect is therefore negligible.

### 2.3 FTIR error budget

As explained in more detail in Senten et al. (2008), we follow the formalism of Rodgers (2000) in which the difference between the retrieved profile \( \hat{x} \) and the real profile \( x \) of the atmosphere can be written as:

\[
\hat{x} - x = (A - I)(x - x_a) + G_y K_b (b - b_r) + G_y (y - y_r),
\]

where \( A \) is the averaging kernel matrix as defined previously, \( I \) the identity matrix, \( G_y \) the gain matrix representing the sensitivity of the retrieved parameters to the measurement, \( K_b \) the sensitivity matrix of the spectrum to the forward model parameters \( b \), \( b_r \) the estimated model parameters, \( y \) the observed spectrum, and \( y_r \) the calculated spectrum. Equation (2) splits the error in the retrieved profile into three different error sources: the smoothing error expressing the uncertainty due to the limited vertical resolution of the retrieval, the forward model parameters error, and the measurement error.

The smoothing error covariance \( S_s \) is calculated as:

\[
S_s = (I - A) S_{\text{var}} (I - A)^T,
\]

where \( S_{\text{var}} \) should represent the natural variability of formaldehyde. Its diagonal elements are estimated from the average observed variability in 5°×5° pixels during PEM-Tropics-B (see Sect. 2.2.2). This variability increases from 33% at the surface up to 15900
70% at 12 km. Above 12 km, we also assume a variability of 70%. For the off-diagonal elements, we have chosen a Gaussian correlation with a correlation length of 4 km, which was obtained from the IMAGESv2 model.

The forward model parameters error covariance matrix $S_f$ is calculated according to:

$$S_f = (G_y K_b) S_b (G_y K_b)^T,$$

in which $S_b$ is the covariance matrix of $b$.

For each individual model parameter, the $K_b$ matrix, which gives the sensitivity of the spectrum to the parameter, is obtained by a perturbation method, while the covariance matrix $S_b$ is an estimation of the uncertainty on the model parameter itself.

The model parameters giving rise to a systematic error on the retrieved formaldehyde are the spectroscopic parameters: the line intensities and the pressure broadening coefficients of the absorption lines present in our microwindows. The uncertainties on the line intensities are between 7 and 10% for the strong and medium lines used in this work (Perrin et al., 2009). We use 10%, to be considered as an upper limit. The values for the air-broadening coefficients have not been updated since the work of Tejwani and Yeung (1977), who did not provide uncertainties. We therefore use a 10% uncertainty.

The largest contributions to the model parameters random error on the formaldehyde total columns are due to the temperature uncertainty and the interfering species, in particular HDO. For temperature, which influences formaldehyde retrievals via the temperature dependence of the absorption lines, the $S_b$ matrix was estimated using the differences between the NCEP and ECMWF temperature profiles calculated for Réunion Island in the period August to October 2004, as explained in more details in Senten et al. (2008). For each of the main interfering species (HDO, $H_2O$, and $CH_4$), the $S_b$ matrix was estimated from the error budget on their prior retrievals.

We also considered the contributions to the random error on the formaldehyde total columns due to uncertainties in the solar zenith angle, the wavenumber shift, the baseline, and the ILS parameters. Only the ILS parameters uncertainties lead to a significant error on the total columns. In their case, the $S_b$ matrix was estimated using...
the differences between the parameters obtained with LINEFIT (see Sect. 2.1) for two adjacent days of cell measurements.

The measurement noise error covariance matrix $S_n$ is calculated by:

$$S_n = G_y S_e G_y^T,$$

(5)

where $S_e$ is assumed to be diagonal, with the square of the spectral noise as diagonal elements. The spectral noise within the selected microwindows is determined as the root mean squared (rms) value of the differences between the observed and calculated spectrum.

The errors on the total column $\Delta TC$ are easily derived from the error covariance matrices $S$ using:

$$\Delta TC = g^T S g,$$

(6)

with $g$ the operator that transforms the volume mixing ratio profile in the corresponding total column amount.

Table 2 summarizes the significant individual contributions to the random and systematic error budget on the formaldehyde total columns.

The comparisons shown in Sect. 6 use the daily means of FTIR and MAX-DOAS total columns. The random errors on these daily means are reduced by a factor $\sqrt{n}$, $n$ being the number of measurements within the day. For FTIR this number $n$ varies from 1 to 20, but with a median value of only 2. We have not divided the smoothing error (Eq. 3) by $\sqrt{n}$ because the natural variability of formaldehyde ($S_{\text{var}}$) within a day was found (from the FTIR measurements) to be of the same order of magnitude as the variability from day to day. Finally, the mean total error on daily mean FTIR total columns is simply the square root of the sum of (1) the square of the total systematic error, (2) the square of the smoothing error, and (3) the sum of the square of the other random error contributions which have been divided by the number $n$ of measurements within the day. We give in Table 2, the mean of the random, systematic, and total errors on daily mean total columns.
3 Formaldehyde from MAX-DOAS observations

3.1 Instrumental set-up and DOAS analysis

The IASB-BIRA MAX-DOAS instrument operated at Réunion Island is described in detail in Theys et al. (2007). It consists of a grating spectrometer installed inside a thermoregulated protection case and connected to an entrance telescope through a fiber optic bundle. The telescope itself is connected to a rotating mirror allowing viewing elevation angles to be scanned from 3° above the horizon up to zenith (90° elevation angle). The spectrometer is mounted with a ruled grating of 600 grooves/mm covering a spectral range from 300 to 450 nm. It is equipped with a 1340×400 back-illuminated CCD detector cooled to −40°C using a triple stage Pelletier system. Spectrometer resolution and sampling ratio are 0.75 nm FWHM and 7 pixels/FWHM, respectively. The MAX-DOAS instrument was operated on the roof of the LACy (Laboratoire de l’Atmosphère et des Cyclones) of the University of Saint-Denis from August 2004 to July 2005. It pointed towards the sea (North direction) at the following elevation angles above the horizon: 3°, 6°, 10°, 18°, and zenith. A complete scan required approximately 15 min.

Measured zenith and off-axis radiance spectra are analyzed using the DOAS technique (Platt and Stutz, 2008). Formaldehyde differential slant column densities (DSCDs), which are the direct product of the DOAS analysis, are retrieved in the 336–358 nm wavelength range, taking into account the spectral signature of O₃, NO₂, BrO, the collision pair of oxygen molecules O₄, and the Ring effect. The cross-sections used in the DOAS fit are taken from the following sources: Meller and Moortgat (2000) for HCHO (293 K), Bogumil et al. (2003) for O₃ (223 K and 243 K), Vandaele et al. (1997) for NO₂ (220 K), Wilmouth et al. (1999) for BrO (228 K), and Greenblatt et al. (1990) for O₄. To correct for the Ring effect (Grainger and Ring, 1962), a pseudo absorption cross section generated from Vountas et al. (1998) using the SCIATRAN radiative transfer model (Rozanov et al., 2001) is included in the fit. Similar settings are used for the retrieval of O₄ slant columns except for the fitting interval, which is slightly shifted
towards longer wavelengths (338.5–364.5 nm) in order to capture the strong O$_4$ absorption band centered at 360 nm. We use a daily noon zenith reference.

### 3.2 Profiling algorithm

The IASB-BIRA profiling algorithm, also based on the OEM of Rodgers (2000) and designed to retrieve vertical distributions of BrO and NO$_2$ from zenith-sky DOAS observations (Hendrick et al., 2004, 2007), has been adapted to the retrieval of HCHO profiles from MAX-DOAS measurements. In the OEM applied to HCHO MAX-DOAS observations, a vertical profile $\hat{x}$ of trace gas concentration is retrieved given an a priori profile $x_a$, the measurements $y$ (here, a set of HCHO DSCDs corresponding to one scan and which are a function of the solar zenith angle (SZA), the relative azimuth angle $\varphi$ between the sun and the viewing direction, and the elevation angle $\theta$), their respective uncertainty covariance matrices ($S_a$ and $S_\epsilon$), and the matrix $K$ of the weighting functions. Since HCHO is an optically thin absorber, the OEM for the linear case can be considered:

$$\hat{x} = x_a + S_a K^T (KS_a K^T + S_\epsilon)^{-1} (y - Kx_a).$$

The weighting functions indicate the sensitivity of the measurements to a change in the vertical profile. Therefore, as for $y$, they also depend on SZA, $\varphi$, and $\theta$. The matrix $K$ is determined by consecutively perturbing each layer of the a priori profile and recalculating the set of measurements using the so-called forward model which describes the physics of the measurements. The forward model consists of the radiative transfer model (RTM) UVspec/DISORT (Mayer and Kylling, 2005). This model is based on the discrete ordinate method and includes a treatment of multiple scattering in the pseudo-spherical geometry approximation. It has been validated through several intercomparison exercises (Hendrick et al., 2006; Wagner et al., 2007). The a priori profile used for all retrievals is the profile also used for the FTIR retrievals (see Fig. 1). Pressure and temperature profiles are taken from NCEP and the ozone profile is from the AFGL reference atmosphere for the tropics (Anderson et al., 1986). The altitude...
grid is as follows: 0, 0.1, 0.2 to 10 km in steps of 0.2 km, 12 to 20 km in steps of 2 km, and 25 to 90 km in steps of 5 km. The wavelength is fixed to 339 nm and the surface albedo (lambertian reflector) to 0.08 (annual mean for 20.9° S and 55.5° E, given by the Koelemeijer et al. (2003) climatology). The measurement covariance matrix $S_\epsilon$ is taken diagonal, with the square of the statistical errors on the HCHO DOAS fit as its diagonal elements. This results from the fact that in most cases, the residuals from the DOAS fit are found to be dominated by the random noise of the detectors. The $S_\epsilon$ matrix being fixed, the a priori covariance matrix $S_a$ can act as a tuning parameter (Schofield et al., 2004) and diagonal elements corresponding to an error of 70% have been empirically determined in order to ensure a good fit of the measurements. $S_a$ also contains off-diagonal elements accounting for correlations between HCHO values at different altitude levels. These terms have been added as Gaussian functions using a correlation length of 400 m (Hendrick et al., 2004).

Since the light path of the different off-axis directions is strongly dependent on aerosols, a good estimate of the aerosol extinction profile is required to calculate accurate HCHO weighting functions for vertical profile retrieval. For this purpose, we have used MAX-DOAS measurements of the oxygen collision complex $O_4$ similarly to Heckel et al. (2005). The principle is the following: since the $O_4$ profile is well-known and nearly constant (it varies with the square of the $O_2$ monomer), the observed $O_4$ absorption depends on the atmospheric distribution of photon paths. Therefore, any change in the light path distribution due to aerosols has an impact on $O_4$ MAX-DOAS measurements (Wagner et al., 2004; Frieß et al., 2006). In order to prevent any complication in the radiative transfer due to the presence of clouds, aerosol extinction profiles (and therefore HCHO profiles) are retrieved only for clear-sky days. The selection of clear-sky days is based on the shape and smoothness of the $O_4$ DSCD diurnal variation (Wagner et al., 2004; Theys et al., 2007). Morning and afternoon are treated separately in order to also include days with only clear-sky morning or afternoon.

Ideally, an aerosol extinction profile should be retrieved at each MAX-DOAS scan since the aerosol loading can vary during the day. However, using the RTM
UVspec/DISORT, this would require unrealistically large computing time. Therefore, we derive a mean aerosol profile for each selected clear-sky morning or afternoon using a look-up table (LUT) approach like in Irie et al. (2008). In our approach, a LUT of O$_4$ airmass factors (AMF$_{O4}$) is created for a large number of different aerosol extinction profiles and viewing geometries employing the RTM UVspec/DISORT. The different aerosol extinction profiles are generated by applying a perturbation to the extinction profile given by the aerosol model of Shettle (1989) included in UVspec/DISORT and corresponding to oceanic conditions with a visibility of 100 km (see Fig. 3). This perturbation depends on three parameters: its altitude (A), the extinction coefficient at altitude A (B) and at the ground (C). A, B, and C were varied in the following ranges: 0, 0.1, 0.2 to 1.2 km in steps of 0.2 km (A), 0.1 to 0.9 km$^{-1}$ in steps of 0.1 km$^{-1}$ (B), and 0.05 to 0.95 km$^{-1}$ in steps of 0.1 km$^{-1}$ (C). This means that 800 aerosol extinction profiles and corresponding AMF$_{O4}$ sets have been generated.

In practice, for a given morning or afternoon, we proceed as follows to determine the “mean” aerosol extinction profile: for each aerosol extinction profile, the AMF$_{O4}$ are interpolated to the SZA, $\varphi$, and $\theta$ corresponding to the different off-axis directions of the measured O$_4$ DSCDs. The O$_4$ DSCDs are then plotted as a function of the interpolated AMF$_{O4}$ (the so-called Langley-plot) and the aerosol extinction profile corresponding to the set of AMF$_{O4}$ that gives the most linear Langley-plot is considered as a good proxy of the true extinction profile (a Langley-plot with good linearity means that the same O$_4$ vertical column is obtained for all viewing directions). The objective criterion used for determining the linearity of the Langley-plot is the RMS of the linear fit applied to the data. For each clear-sky morning or afternoon, the selected aerosol extinction profile is the one corresponding to the set of AMF$_{O4}$ giving the lowest RMS below a threshold value of $300 \times 10^{40}$ mol$^2$/cm$^5$. If this threshold value is not reached, the corresponding clear-sky morning or afternoon is not taken into account in the study. An example of an aerosol extinction profile selection is presented in Fig. 4.

Using this RMS-based criterion, aerosol extinction profiles are selected for 52 mornings and 42 afternoons, and all together, only 6 different profiles are found (see Fig. 5).
and used to calculate appropriate sets of HCHO weighting functions.

3.3 Characterization of the retrieval

3.3.1 Retrieval fit results

The HCHO profile retrievals are quality-checked by examining the retrieval fit results, which are the comparison between the measured HCHO DSCDs and those calculated using the retrieved profiles. Figure 6 shows examples of retrieval fit results obtained on 25/11/2004 morning for some MAX-DOAS scans between 58° and 68° SZA. A similarly good agreement is found between measured and calculated HCHO DSCDs for all selected mornings and evenings, meaning that the settings used for the HCHO weighting function calculations, especially the aerosol extinction profiles, are good estimates of the real atmospheric conditions.

3.3.2 Information content

The information content is characterized through the calculation of the averaging kernel matrix $A$, defined in Eq. (1), for each measurement scan. As seen in Sect. 2.2.3, the averaging kernels (rows of the $A$ matrix) express the sensitivity of the retrieved profile $\hat{x}$ to the true atmospheric profile $x$. A typical example of HCHO averaging kernels is presented in Fig. 7.

We see that, according to the MAX-DOAS geometry, most of the information on the vertical distribution of HCHO contained in the measurements is located below 2.5 km of altitude. The trace of $A$ (DOFS) is $0.7 \pm 0.1$ on average for a scan. This rather low DOFS value is mainly due to the relatively poor signal to noise ratio of the measurements performed in this campaign and to the small number of elevation angles (5 with zenith) available for each scan. A sensitivity analysis indicates that a DOFS of up to 1.6 could be achieved by reducing the noise on HCHO SCDs ($0.45 \times 10^{16} \text{ mol/cm}^2$ for this campaign) down to $0.15 \times 10^{16} \text{ mol/cm}^2$ as for our most recent instrument operating in
Beijing, and by considering four additional elevation angles (2, 8, 15, and 30°).

### 3.3.3 Error budget

As already discussed in Sect. 2.3, the total retrieval error consists of the sum of the smoothing error, the forward model parameters error, and the measurement noise error. They are calculated according to Eqs. (3–5), respectively. The matrix $S_{\text{var}}$ used in Eq. (3) for the MAX-DOAS smoothing error is the same as for the FTIR smoothing error, since it represents our best estimate of the HCHO variability at Réunion Island. The forward model parameters $b$ for which we calculate $K_b$ for the MAX-DOAS retrievals, and their associated uncertainties $S_b$ are listed in Table 3.

The smoothing, measurement noise, and forward model parameters errors on the retrieved HCHO total columns are presented in Table 4 as well as the total retrieval error.

The total error on the columns is calculated by adding in quadrature the systematic and total retrieval errors. The main source of systematic error is the uncertainty on the HCHO cross-sections. This uncertainty is estimated to be 9%, which is the mean difference between the Cantrell et al. (1990) and Meller and Moortgat (2000) HCHO cross sections. The total error on individual HCHO total columns is about 27%. As for FTIR total columns in Sect. 2.3, we give also in Table 4 the mean of the error on daily mean total columns for MAX-DOAS. The random error is significantly decreased because the number of MAX-DOAS measurements within a day varies from 4 to 84, with a median value of 36.

### 4 Formaldehyde from SCIAMACHY observations

Global tropospheric HCHO columns have been retrieved from the SCIAMACHY UV-Visible nadir spectrometer onboard ENVISAT (launched in 2001), using the differential optical absorption spectroscopy (DOAS) technique (Platt and Stutz, 2008). The HCHO
retrievals used in the present paper are described in detail in De Smedt et al. (2008). Therefore, we only give a very brief summary of the retrieval characteristics here.

Slant columns are fitted in the 328.5–346 nm wavelength range using the WINDOAS analysis software (Van Roozendael et al., 1999). This choice of the fitting interval minimizes uncertainties due to a polarization anomaly affecting the SCIAMACHY spectra around 350 nm, and to a major absorption band of the O₄ collision complex (centered near 360 nm). Furthermore, it decreases fitting residuals in tropical areas and reduces noise over the oceans (De Smedt et al., 2008). Vertical columns are obtained by dividing the slant columns by air mass factors (AMFs) (Palmer et al., 2001) calculated using scattering weights evaluated from radiative transfer calculations performed with the DISORT code (Mayer and Kylling, 2005). A correction for cloud effects is applied based on the independent pixel approximation (Martin et al., 2002), but aerosols are not explicitly considered in the current version of the dataset. The ground albedo is obtained from the climatology of Koelemeijer et al. (2003). Cloud information is provided by the FRESCO v5 algorithm (Koelemeijer et al., 2002; Wang et al., 2008). For the determination of the AMFs, HCHO vertical profiles are needed a priori (Palmer et al., 2001): they are provided by the IMAGESv2 global chemistry transport model (Stavrakou et al., 2009a) on a daily basis and interpolated for each satellite geolocation.

A random and a systematic error component of the slant columns is estimated (De Smedt et al., 2008). For a single pixel, the random error reaches 10^{16} mol/cm^2. However, when considering regionally and temporally averaged columns, this error is reduced by the square root of the number of observations included in the mean. The error on the air mass factors is mainly due to uncertainties on the cloud properties and on the HCHO vertical distribution. On average, the total uncertainty on the monthly HCHO vertical column ranges between 20 and 40% for regions with high HCHO signal to noise ratio (De Smedt et al., 2008). In this study, daily averaged HCHO columns have been calculated in a region of 500 km around Saint-Denis, using observations with a cloud fraction below 0.4. Only days for which at least 20 SCIAMACHY pixels
were available and for which the total error was smaller than $2.5 \times 10^{15} \text{ mol/cm}^2$ have been selected. In 2004, the mean total error on the daily averaged columns is 40%, divided into 15% of random errors and 37% of systematic errors. The mean number of satellite observations per day is 200. In 2007, total errors are larger. There are two reasons for this. First, the degradation of the satellite with years increases the fitting residuals. This impacts the random error. Second, the observed cloud fractions are larger, which increases the AMF uncertainties and reduces the number of pixels included in the daily means. The mean total error is 60% (30% random and 51% systematic) and the averaged number of observations is only 100.

5 Formaldehyde from the CTM IMAGESv2

The global IMAGESv2 chemistry-transport model (CTM) is an updated version of the IMAGES model (Müller and Brasseur, 1995; Müller and Stavrakou, 2005). It calculates the daily mean concentrations of 80 trace gases at a horizontal resolution of 4 × 5 degrees and 40 vertical levels extending from the surface up to the top of the troposphere. Monthly mean ECMWF wind fields are used to drive advection of chemical compounds, whereas boundary layer diffusion, deep convection and other cloud processes are parameterized using daily ECMWF fields. The NMVOC chemical mechanism of the model has been revised on the basis of box model calculations with the quasi-explicit Master Chemical Mechanism (MCM, Saunders et al., 2003) in order to provide a more reliable representation of the formaldehyde produced by pyrogenic and biogenic hydrocarbon emissions (Stavrakou et al., 2009a). The model time step is equal to one day. Diurnal cycle simulations with a time step of 20 min are performed to derive correction factors for the kinetic and photolysis rates (Stavrakou et al., 2009a).

Anthropogenic NMVOC emissions are obtained from the EDGAR v3.3 database for 1997 (Olivier et al., 2001), vegetation fire emissions from the GFEDv2 database (van der Werf et al., 2006), and isoprene emissions from the MEGAN-ECMWF inventory (Müller et al., 2008). A diurnal fire profile based on Giglio (2007) is applied to the di-
urnal cycle calculations (Stavrakou et al., 2009a). Further, the large-scale fire injection heights are obtained from Dentener et al. (2006).

Prescribed monthly mean climatological OH fields from Spivakovsky et al. (2000) are used to test the influence of OH on the calculated HCHO columns.

The HCHO columns have been calculated with and without accounting for the averaging kernels of the measured FTIR and MAX-DOAS columns.

The chemical simulations are conducted for 2004, 2005, and 2007 following a 4-month initialization.

6 Results

Figure 8 displays the time series of the daily mean total columns of HCHO from FTIR MAX-DOAS and SCIAMACHY observations in the period common to the FTIR and MAX-DOAS measurements campaigns, i.e., from August to November 2004. Analogously, Fig. 9 shows the time series of FTIR and SCIAMACHY daily mean total columns during the 2007 FTIR campaign. The error bars represent the total error on the daily mean column, as discussed in Sects. 2.3, 3.3.3, and 4, for FTIR, MAX-DOAS and SCIAMACHY, respectively. In the next section (6.1), we discuss the comparison between the FTIR and MAX-DOAS results. The comparisons between the FTIR and SCIAMACHY, and between the MAX-DOAS and SCIAMACHY results, are discussed in Sects. 6.2 and 6.3, respectively.

6.1 FTIR versus MAX-DOAS

It can be seen in Fig. 8 that the daily means of FTIR and MAX-DOAS total columns measured on coincident days are in very good agreement, considering their total error bars.

The statistical analysis of the differences between the FTIR and MAX-DOAS measurements has been performed as follows. We have first calculated the direct dif-
ferences (FTIR–DOAS) for the 24 days of coincident measurements. The mean and standard deviation of these differences should be compared to the systematic and random errors on the difference, respectively. As FTIR and MAX-DOAS use the same a priori profile, we can directly compare the retrieved profiles without adjusting them for different a priori, as described in Eq. (10) of Rodgers and Connor (2003).

The error covariance matrix of the difference between the retrieved profiles from two different sounders is given by Eq. (13) of Rodgers and Connor (2003). Taking into account that, in our case, (1) the a priori profile $x_a$ and the covariance matrix $S_{\text{var}}$ are the same for FTIR and MAX-DOAS; and (2) the FTIR and MAX-DOAS retrievals use different vertical grids, we calculate the random error covariance matrix $S_{\delta_{\text{direct}}}^{\text{rand}}$ as follows:

$$S_{\delta_{\text{direct}}}^{\text{rand}} = W_{21} S_{1}^{\text{rand}} W_{21}^T + S_{2}^{\text{rand}} + (W_{21} A_1 W_{12} - A_2)^T S_{\text{var}} (W_{21} A_1 W_{12} - A_2),$$

with $A_1$ and $S_{1}^{\text{rand}}$ the averaging kernel matrix and the random error covariance matrix (noise and forward model parameters errors) for the MAX-DOAS retrievals, $A_2$ and $S_{2}^{\text{rand}}$ the averaging kernel matrix and the random error covariance matrix (noise and forward model parameters errors) for the FTIR retrievals, $S_{\text{var}}$ expressed in the FTIR grid, and $W_{21}$ and $W_{12}$ the grid transformation matrices defined in Calisesi et al. (2005).

As seen for the random error on the FTIR and MAX-DOAS daily mean total columns, $S_{1}^{\text{rand}}$ and $S_{2}^{\text{rand}}$ are the MAX-DOAS and FTIR random error covariance matrices divided respectively by the number of MAX-DOAS and FTIR measurements used in the daily mean comparisons.

The systematic error covariance matrix $S_{\delta_{\text{direct}}}^{\text{syst}}$ can be calculated as:

$$S_{\delta_{\text{direct}}}^{\text{syst}} = W_{21} S_{1}^{\text{syst}} W_{21}^T + S_{2}^{\text{syst}}.$$  

Or, as we are interested in the error on total column differences, we can obtain it directly
as:
\[
\Delta T_C^{\text{syst}} = \sqrt{(\Delta T_{\text{FTIR}}^{\text{syst}})^2 + (\Delta T_{\text{DOAS}}^{\text{syst}})^2}.
\] (9)

However, as discussed before, the FTIR measurements are sensitive to the whole partial column below about 20 km, while the MAX-DOAS measurements are sensitive only below 2 km. To take this into account, we follow Eq. (28) of Rodgers and Connor (2003) where we smooth the FTIR profile \(x_2\) with the MAX-DOAS averaging kernels \(A_1\), i.e., we degrade the profile with the higher vertical resolution (cf. DOFS FTIR=1.1) to the lower resolution (cf. DOFS MAX-DOAS=0.7):

\[
x_2^{\text{smooth}} = x_a + A_1 \ast (x_2 - x_a).
\] (10)

The random and systematic error covariance matrices on the differences, denoted as \(S_{\delta_{\text{smooth}}^{\text{rand}}}\) and \(S_{\delta_{\text{smooth}}^{\text{syst}}}\), respectively, are calculated according to Eq. (30) of Rodgers and Connor (2003) and Eq. (22) of Calisesi et al. (2005):

\[
S_{\delta_{\text{smooth}}^{\text{rand}}} = W_21 S_1^{\text{rand}} W_2^T
\]

\[
+ (W_21 A_1 W_12)^T S_2^{\text{rand}} (W_21 A_1 W_12)
\]

\[
+ (I_2 - A_2)^T (W_21 A_1 W_12)^T S_{\text{var}} (W_21 A_1 W_12) (I_2 - A_2)
\] (11)

\[
S_{\delta_{\text{smooth}}^{\text{syst}}} = W_21 S_1^{\text{syst}} W_2^T
\]

\[
+ (W_21 A_1 W_12)^T S_2^{\text{syst}} (W_21 A_1 W_12),
\] (12)

with \(I_2\) the identity matrix on the FTIR grid. We have expressed all matrices on the FTIR retrieval grid \((z_2)\), because it is the coarser grid (below 10 km) and therefore the interpolation errors are smaller.

Table 5 summarizes the comparisons between the FTIR and MAX-DOAS data sets in 2004. The table lists the weighted mean differences and standard deviations, as well as the mean values for the random and systematic errors on the differences.
Confirming the good agreement seen in Fig. 8, the mean difference FTIR–DOAS lies within the systematic error, and the standard deviation is close to the random error. We see from Table 5 the expected effect of the averaging kernels on the comparisons: the standard deviation of the difference is significantly decreased, as well as the random error on the difference. While the mean of the differences still lies within the systematic error, we see that the standard deviation is slightly larger than the random error, probably due to the errors associated with the grid conversions, as discussed in Calisesi et al. (2005).

6.2 FTIR versus SCIAMACHY

Figures 8 and 9 show the daily mean FTIR total columns compared to the daily SCIAMACHY columns in a region of 500 km around Saint-Denis, during the two FTIR campaigns in 2004 and 2007, respectively. The agreement between FTIR and SCIAMACHY total columns is very good: their values coincide within the given errors bars. The day-to-day variability is well reproduced by both instruments, as seen for example in Fig. 8 for the especially low total columns in mid-August 2004 or the very high values in mid-October 2004. We see in Fig. 9 that both instruments show the same seasonal behaviour with a minimum in August–September. We note that, due to larger cloud contamination, the SCIAMACHY errors are larger in 2007 than in 2004, as already mentioned in Sect. 4.

For the 21 and 28 coincidences in 2004 and 2007, respectively, the statistical analysis of the differences between FTIR and SCIAMACHY HCHO total columns has been performed as follows. Since the retrievals have used different a priori profiles, we adjust the two retrieved products to a comparison ensemble \( x_c \) by adding to each of them the term \((A_r-I)(x_{ar}-x_c)\) (Rodgers and Connor, 2003). We have chosen \( x_c \) to be the a priori FTIR profile \( x_{a,FTIR}=x_a \). The SCIAMACHY retrievals have been made using a zero linearization point (De Smedt et al., 2008, following the work of Eskes and Boersma, 2003), thus \( x_{a,SCIA}=0 \). So, the FTIR retrieval products to be used for comparisons remain unchanged while we have to add the term \((A_{SCIA}-I)(0-x_a)\) to the SCIAMACHY
products. Therefore, the SCIAMACHY total columns for comparison $c_{1}^{\text{compar}}$ become:

$$c_{1}^{\text{compar}} = c_1 + c_a - a_1^T x_a^{pc},$$  \hfill (13)

in which $c_1$ is the retrieved SCIAMACHY total column, $a_1$ the associated total column averaging kernel, $c_a$ the FTIR a priori total column and $x_a^{pc}$ the FTIR a priori profile of layers partial columns, expressed in mol/cm$^2$.

The direct difference between both instruments is then:

$$\delta_{\text{direct}} = c_2 - c_1 - c_a + a_1^T x_a^{pc},$$  \hfill (14)

$c_2$ being the FTIR retrieved total column.

The variance of this difference is given by Eq. (24) of Rodgers and Connor (2003). If we decompose it into its random and systematic components, we have:

$$\sigma_{\text{direct}}^{\text{rand}} = \sigma_{1}^{\text{rand}} + \sigma_{2}^{\text{rand}}$$

$$+ (a_1 - a_2)^T S_{\text{var}}^{pc} (a_1 - a_2)$$

$$\sigma_{\text{direct}}^{\text{syst}} = \sigma_{1}^{\text{syst}} + \sigma_{2}^{\text{syst}},$$  \hfill (15)

(16)

where $a_2$ is the FTIR total column averaging kernel and $S_{\text{var}}^{pc}$ is expressed in the FTIR grid and in partial column units. $a_1$ can easily be interpolated in the FTIR grid, since it is independent of the layer width in case of SCIAMACHY.

As in the previous section, in order to take into account the difference in vertical sensitivity of both instruments, we calculate the “smoothed FTIR total columns” following Eq. (10):

$$c_2^{\text{smooth}} = c_a + a_1^T (x_2^{pc} - x_a^{pc}),$$  \hfill (17)

with $x_2^{pc}$ the FTIR retrieved profile expressed in partial column units.

The difference between the smoothed FTIR total column and the SCIAMACHY total columns is then, combining Eqs. (13) and (17):

$$\delta_{\text{smooth}} = a_1^T x_2^{pc} - c_1.$$  \hfill (18)
The variance of this difference is given by Eq. (27) of Rodgers and Connor (2003). If we separate the random and systematic component, we obtain:

\[
(\sigma_{\text{smooth}}^{\text{rand}})^2 = (\sigma_{1}^{\text{rand}})^2 + a_1^T S_{2}^{\text{rand,pc}} a_1 + a_1^T (I_2 - A_2^{\text{pc}}) S_{\text{var}}^{\text{pc}} (I_2 - A_2^{\text{pc}})^T a_1
\]

\[\text{(19)}\]

\[
(\sigma_{\text{smooth}}^{\text{syst}})^2 = (\sigma_{1}^{\text{syst}})^2 + a_1^T S_{2}^{\text{syst,pc}} a_1
\]

\[\text{(20)}\]

where \(A_2^{\text{pc}}\) and \(S_{2}^{\text{pc}}\) are the FTIR averaging kernel and error covariance matrices, expressed in partial column units.

The comparison between FTIR and SCIAMACHY columns is summarized in Table 6, for the 2004 and 2007 campaigns. The agreement is very good in 2004: the means of the differences are within the systematic error for both direct and FTIR smoothed comparisons, and the standard deviations correspond nicely to the random error budget. In 2007, the means are still within the systematic errors but the observed standard deviations (about 54%) are larger than the random errors (only about 35%). One explanation can be given for the latter discrepancy. The SCIAMACHY total columns being spatial averages, a geolocation error exists, which is not taken into account. This could have a larger effect in 2007, when the number of pixels included in the daily means is reduced due to enhanced cloud fractions.

We notice also from Table 6, that the smoothing of the FTIR profiles with the SCIAMACHY total column averaging kernels does not significantly improve the comparisons. This is probably due to the fact that both instruments have similar total column averaging kernels, as illustrated in Fig. 10.

### 6.3 MAX-DOAS versus SCIAMACHY

Although the total column averaging kernels, given in Fig. 10, show that MAX-DOAS and SCIAMACHY instruments have a maximum sensitivity to formaldehyde in two different parts of the atmosphere (0–2 km for MAX-DOAS and above 2 km for SCIAMACHY), the comparison of both data sets is found to be relevant because (1) between
0 and 2 km of altitude both total column averaging kernels are at least in the 0.3–0.8 range (see Fig. 10), and (2) about 40% of the formaldehyde total column is present in this altitude range.

The time series of MAX-DOAS and SCIAMACHY total columns from August 2004 to August 2005 are depicted in Fig. 11. It can be seen that the MAX-DOAS time series of total columns shows a smaller day-to-day variability than the SCIAMACHY one. This could be explained by the fact that MAX-DOAS measurements underestimate the variability of formaldehyde above 2 km, while SCIAMACHY tends to overestimate it above 3 km (total column averaging kernels larger than 1). We can note that the FTIR total column averaging kernel is larger than 1 above 5 km (Fig. 10), which means that the FTIR variability could also be overestimated. Above this altitude only about 24% of the formaldehyde total column is present (and only 10% above 10 km), thus the effect on the total column variability should not be too large. However, the effect of underestimation of the variability for MAX-DOAS and overestimation for FTIR is taken into account in their respective error bars via the smoothing error. One should keep in mind that this smoothing error has been calculated with our best estimate of the formaldehyde natural variability (PEM-Tropics-B measurements, see Sect. 2.3) which could be an underestimation of the variability at Réunion Island: as discussed in the following Sect. 7, Réunion Island is affected by long-range transport of HCHO precursors from Madagascar.

Using the same formalism as in Sect. 6.2, we give in Table 6 the statistical analysis and error budget for the MAX-DOAS and SCIAMACHY comparisons. The agreement is reasonably good with weighted mean and standard deviation of the differences between “MAX-DOAS smoothed” and SCIAMACHY total columns of −16.9±31.0%. The standard deviation corresponds nicely to the random error on the difference (32.5%).

6.4 FTIR and MAX-DOAS versus IMAGESv2

Figures 12 and 13 to the formaldehyde columns simulated by the IMAGESv2 model and interpolated at the measurement site. The performed model simulations are sum-
marized in Table 7. The sensitivity run S2 is motivated by the important role played by the hydroxyl radical (OH). The oxidation of methane by OH is by far the largest source of formaldehyde in the troposphere, in particular over remote areas. The simulation S3 includes the oxidation of methane by chlorine atoms in the marine boundary layer (MBL), which could account for about 5% of the global tropospheric sink of methane according to an estimation based on isotopic measurements (Allan et al., 2007). The chlorine concentrations in the MBL are parameterized as in Allan et al. (2007) as a function of latitude and season, with an annual average of $18 \times 10^3$ mol cm$^{-3}$ and a large seasonal cycle. The chemistry of chlorine has been considered previously as a possible explanation for the underestimation of modeled formaldehyde and peroxy radical concentrations against ship measurements in the Indian ocean (Burkert et al., 2003; Wagner et al., 2002).

Model results obtained in simulations S1 and S2 are shown in Figs. 12 and 13. In both cases, the averaging kernels of the instruments were applied to the modeled profiles in the determination of the total column. For illustration purposes, vertical columns calculated without taking the averaging kernels into account are also shown. The results of simulation S3 are omitted since they are found to differ only very marginally from simulation S1. More precisely, although the formaldehyde production due to the presence of chlorine in the MBL increases the formaldehyde concentration by 5–15% at surface level, its impact on the total column is only 1–2% (taking the averaging kernels into account).

Table 8 summarizes the results of the comparisons for the two campaigns of FTIR measurements in 2004 and 2007 (52 and 84 coincidences, respectively), and for the MAX-DOAS campaign measurements (70 coincidences).

7 Discussion

Réunion Island, being a small island located in the Indian Ocean, can be seen as a remote marine site where the methane oxidation is the dominant formaldehyde
source. However, it is situated only 700 km from the East coast of Madagascar and about 2000 km from southeastern Africa, regions with large biogenic NMVOC emissions as well as extensive vegetation fires during the May–November period. Although formaldehyde itself has a too short lifetime (a few hours) to be directly transported over such distances, pyrogenic and biogenic NMVOC precursors (e.g. ethane, methanol, acetic acid, etc.) and oxidation products (e.g. certain organic nitrates, hydroperoxides, etc.) can be transported to Réunion Island and lead to enhanced formaldehyde formation.

As seen in Figs. 9 and 11, ground-based and SCIAMACHY measurements show a minimum of formaldehyde in local winter, which is also found in the IMAGESv2 model (Figs. 12 and 13). This minimum is primarily due to the lower radiation and humidity levels prevailing during this period, which lead to lower OH concentrations and therefore to lower methane oxidation rates. The model predicts a maximum of formaldehyde in January–March (Fig. 13), also found in the MAX-DOAS and SCIAMACHY measurements (Fig. 11). Unfortunately, only few ground-based measurements are available in this period due to the prevalence of cloudy conditions.

However, discrepancies are found regarding the amplitude of the seasonal variations: for example, the FTIR columns increase by about 50–60% between July–August and October, whereas the modeled values (taking FTIR averaging kernels into account) increase by about 40% in the same time period (Fig. 12). The MAX-DOAS measurements show a smaller seasonal amplitude (about 30%, Fig. 13), primarily because of the rather low degrees of freedom for signal (DOFS=0.7) and the fact that MAX-DOAS is sensitive only below 2.5 km. This is confirmed in Fig. 13 where we see that the amplitude of the modeled seasonal cycle is strongly reduced by the application of the MAX-DOAS averaging kernels.

Both the FTIR and SCIAMACHY total columns show a high day-to-day variability (Figs. 8 and 9), which is not captured by the model. The largest disagreements between FTIR and the model appear in October 2004 and in June 2007. The day-to-day variability of the MAX-DOAS columns is also higher than in the model, e.g. in June
2005, when values above $4 \times 10^{15}$ mol/cm$^2$ are found (Fig. 13) while the modeled values are at their minimum (around $3 \times 10^{15}$ mol/cm$^2$).

The large temporal variability of FTIR formaldehyde columns might be partly elucidated by their good correlation (0.7) with CO total columns measured with the same instrument (Senten et al., 2008) during the period from August to November, when the fire season is at its maximum in southeastern Africa and Madagascar (see Fig. 14). This result suggests that long-range transport of HCHO precursors from these regions can explain the large temporal variability of observed HCHO columns. The general underestimation of the modeled columns against FTIR measurements (by ca. 25% on average, see Table 8) would therefore reflect either an underestimation of NMVOC emissions in Southeast Africa and Madagascar in the model, or an underestimation of the role of long-range transport from these source regions to the measurement site. This interpretation is consistent with the comparatively lower bias found between the model and the MAX-DOAS measurements, since the DOAS technique is mostly sensitive to the lowermost atmospheric layers (below 2.5 km), where long-range transport is ineffective (see below).

To assess the role of long-range transport on the observations at Réunion Island, we used the Lagrangian particle dispersion model FLEXPART v.6.2 (Stohl et al., 2005). The model was driven by 3-hourly $1^\circ \times 1^\circ$ windfields from ECMWF. For this study, for each day during our measurement campaign, the model released 1 million particles between 0–10 km above Saint-Denis at 05:00 UT. Backward trajectories were calculated using a CO tracer. The calculated backward trajectory output (i.e., the response function) is related to the particles residence time in the output grid cells. In our case the response function (s m$^{-3}$ kg$^{-1}$) at a given location equals the residence time divided by the local total air mass concentration. When multiplying by a 3-D field of emission mass fluxes (kg m$^{-3}$ s$^{-1}$), one can derive the strength of the contribution of a particular location (source strength) to the observations at Réunion Island at a given time. We use CO as a qualitative proxy for the emissions of HCHO and its precursors. The CO emission fluxes were obtained from the GFEDv2 fire emission database. Apart from
its limited temporal resolution (8 days) and intrinsic uncertainties, an additional major uncertainty is the emission injection height. At Réunion Island, the dominant transport pathway at low altitudes (below 4 km during most of the fire season) originates from the South-East (Taupin et al., 1999). At these low altitudes, African-Madagascar emissions reach Réunion Island only via a roundabout way. Given the short lifetime of HCHO and of its pyrogenic precursors that could be emitted by fires (a few hours to a few days) it is unlikely that these pathways contribute significantly to the observed HCHO concentrations. Above 4 km, the air masses generally take a direct eastward pathway (Taupin et al., 1999) and any African-Madagascar emissions that reach this altitude band are transported within a short timeframe towards Réunion Island. To assess the emission injection heights and the timescales at which the transport events take place, we have calculated the CO tracer volume mixing ratio detected at Réunion Island for different backward trajectory runtimes. It was found that the best correlation between the observed HCHO and simulated CO data was obtained when restricting the runtime to \( \approx 1 \) day and by allowing emission into the 4–6 km altitude layer. The IMAGEsv2 model uses the same emission height distribution as given by Dentener et al. (2006) and the EDGAR database (Olivier et al., 2001), setting the top emission height for large scale wild-fires from Madagascar at 1 km. However, according to Freitas et al. (2007), injection heights above 4 km are certainly attainable, even for African Savanna fires. An alternative explanation is the potential underestimation of vertical (convective or other) transport within the FLEXPART model itself (to simulate moist convection, FLEXPART uses the convective parameterization scheme by Emanuel and Živković-Rothman, 1999). Such an underestimation of vertical transport processes can only be offset by allowing an overestimation of high altitude emissions in the model run.

Due to the various uncertainties (limited temporal resolution of GFEDv2 database, CO used as a proxy, transport in FLEXPART), we give here qualitative examples only, of how emissions followed by transport to Réunion correlate with the high peak values of FTIR HCHO total columns. Within the 1 day backward trajectory, the FLEXPART simulations indicate that emissions/transport originate from Madagascar only. Longer
runtimes yielded lower correlations, indicating that sources further inland do not significantly contribute to the observations at Réunion Island.

During the 2004 campaign, we can observe from Fig. 12 that for example the 13 and 14 August are days with high HCHO values compared to the previous and following days (8th and 16th). The results of the FLEXPART simulations (1 day backward trajectories) for those four days are shown in Fig. 15. FLEXPART calculates indeed a significant contribution of Madagascar emissions to the observed signal at Réunion Island on the 13th and 14th, while no transport is seen on the 8th and 16th. Using 3 days backward trajectories, we observe (not shown) that the transport patterns of pyrogenic CO from Africa are not associated with any enhancement in HCHO, presumably because of the short lifetimes of its precursors. Similarly, we observe high HCHO events on 14 and 29 September as well as 12 October, that are well reproduced by FLEXPART (not shown). On the contrary, the 21–22 October peaks in 2004 are missed by FLEXPART, and are not seen in the FTIR CO measurements (Fig. 14). Note that MAX-DOAS and SCIAMACHY also show higher values of HCHO on the 21–22 October (see Fig. 8). It can be speculated that these enhancements appear below 2.5 km, and are due to local sources, which would explain that the model and the FTIR CO measurements miss them.

Similarly, the HCHO peaks in August–November 2007 can be explained using FLEXPART simulations. We show as an example in Fig. 15 the well isolated peak observed on the 6 August (Fig. 12). The peak values of 18 and 31 August, and 24 September 2007 (Fig. 12) are also well reproduced by FLEXPART (not shown). On the contrary, the peak value of 11 October 2007 is not explained by FLEXPART, but an enhancement of CO is also seen on this day by FTIR measurements (Fig. 14).

However, we note that there is hardly any correlation between the observed FTIR’s CO and HCHO in May–July 2007 (see Fig. 14), suggesting that the long-range transport of biomass burning plumes could not be the only explanation to the large temporal variability of FTIR columns. We give in Fig. 16 the biomass fire emissions of CO observed at Madagascar during 2004 and 2007 as given by the GFEDv2 database. While
the fire intensity peaks in September–October, there are indeed ample fires throughout the whole FTIR measurement campaign. As illustrated in Fig. 15 with an example of FTIR data in June, FLEXPART explains the observed day-to-day variability: a low HCHO value on the 3rd, high values on the 5th and 11th, and again smaller values on the 12th. An explanation for the absence of correlation with FTIR CO, is that CO’s lifetime is much larger than that of HCHO, and therefore will be primarily impacted by the global fluctuation and accumulation of biomass burning emissions, which are more important in August–November period.

In conclusion, it appears plausible that both the large temporal variability of observed HCHO columns and the general underestimation of HCHO columns by the IMAGESv2 model in comparison with FTIR data (Fig. 12) are primarily caused by an underestimated contribution of Madagascar emissions to the signal measured at Réunion Island. Uncertainties in the determination of OH concentrations in the model might also possibly contribute to the underestimation, as suggested by the results of simulation S2 given in Table 8: the biases between ground-based and modelled total columns are reduced to 8% using prescribed OH fields. The role of chlorine-initiated methane oxidation appears to be very small in this context (simulation S3). Pyrogenic and biogenic NMVOC emissions might be largely underestimated in the model in South-East Africa, as indicated by an inverse modeling study based on SCIAMACHY data (Stavrakou et al., 2009b). The fire injection heights and the convective updraft fluxes used in the model are also very uncertain.

8 Conclusions

Ground-based FTIR and MAX-DOAS daily mean total columns of formaldehyde have been retrieved during campaign measurements, in 2004 and 2007, and in 2004–2005, respectively, at the NDACC site of Réunion Island (21°S, 55°E). The two datasets are compared in the common measurement period August to October 2004. FTIR and MAX-DOAS total columns have also been compared to correlative SCIAMACHY
data. The comparisons include an error analysis, and account for differences in vertical sensitivities by including the respective averaging kernels. A good agreement is found (no significant bias, standard deviations within the random error budget) between the three instruments except for the FTIR and SCIAMACHY comparisons in 2007 where the standard deviation is larger (54%). The three instruments show the same seasonal cycle with a minimum in local winter, due to the lower OH concentrations during this period. However, the MAX-DOAS measurements show a lower seasonal amplitude due to their lack of sensitivity in the free troposphere.

The FTIR HCHO total columns show a high day-to-day variability at Réunion Island, which is correlated with FTIR CO measurements during the intense fire season in Madagascar in August-November. Simulations performed using the Lagrangian particle dispersion model FLEXPART, allowing emissions into the 4–6 km layer and using 1 day backward trajectory runtimes, explain qualitatively this day-to-day variability, even in the May–June period where bad correlation is seen between FTIR HCHO and CO. These results indicate an underestimation of the fire emission injection heights as given by EDGAR (Olivier et al., 2001) and/or the underestimation of vertical transport within the FLEXPART model. Quantitative studies would require less uncertainties in the FLEXPART simulations: an increased temporal resolution of the GFEDv2 database (currently 8 days), vertical and horizontal transport, fire injection heights, ...

The FTIR and MAX-DOAS data have also been compared to modeled HCHO columns from the global CTM IMAGESv2. We found that the standard model underestimates HCHO total columns at Réunion Island (biases of 25–31 % and 15% are obtained with FTIR and MAX-DOAS comparisons, respectively). This could be due to underestimate of OH concentrations in the model as suggested by a simulation performed using prescribed OH fields. More probably, the model underestimates the contribution of large-scale transport of HCHO precursors from Madagascar to Réunion Island. The pyrogenic and biogenic emissions at Madagascar could be underestimated. The fire injection heights and the convective updraft fluxes used in the model might be also too low in the model.
Acknowledgements. We acknowledge for financial support the Belgian Federal Science Policy (AGACC; SSD program, SD/AT/01B), the PRODEX office (SECPEA), the Communauté Française de Belgique (Actions de Recherche Concertées), the EU (SCOUT-O3), and ESA (TEMIS and PROMOTE). We thank A. Perrin, D. Jacquemart, F. Kwabia Tchana and N. La- come for providing their new HCHO spectroscopic linelist before its publication. We are grateful to all people that have participated to the ground-based measurements. We acknowledge J.-L. Baray, R. Delmas, J. Leveau, and G. Clain, from Université de La Réunion, for their collaboration. We thank also G. Dufour for discussion on ACE HCHO measurements above La Réunion.

References

Calisesi Y., Soebijanta V. T., and van Oss R.: Regridding of remote soundings: For-
FTIR and MAX-DOAS HCHO observations at Réunion Island

C. Vigouroux et al.

Abstract

Introduction

Conclusions

References

Tables

Figures


Steck, T., Glatthor, N., von Clarmann, T., Fischer, H., Flaud, J. M., Funke, B., Grabowski,


Taupin, F. G., Bessafi, M., Baldy, S., and Bremaud, P. J.: Tropospheric ozone above the southwestern Indian Ocean is strongly linked to dynamical conditions prevailing in the tropics, J. Geophys. Res., 104, 8057–8066, 1999. 15921


Vandaele, A.-C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M.-F., Jenouvrier, A., and Coquart, B.: Measurements of the NO₂ absorption cross-section from 42 000 cm⁻¹ to 10 000 cm⁻¹ (238–1000 nm) at 220 K and 294 K, J. Quant. Spectrosc. Ra.,
Table 1. Microwindows (in cm\(^{-1}\)) used for the independent retrievals of H\(_2\)O, HDO, CH\(_4\), N\(_2\)O, and HCHO. The retrieved profiles of the first four compounds are used as a priori profiles in the retrievals of HCHO.

<table>
<thead>
<tr>
<th>Target gas</th>
<th>Microwindows (cm(^{-1}))</th>
<th>Interfering species</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(_2)O</td>
<td>2925.10–2925.30</td>
<td>CH(_4)</td>
</tr>
<tr>
<td></td>
<td>2941.60–2941.90</td>
<td>CH(_4), O(_3), solar CO</td>
</tr>
<tr>
<td>HDO</td>
<td>2660.00–2661.20</td>
<td>CH(_4), CO(_2)</td>
</tr>
<tr>
<td>CH(_4)</td>
<td>2613.70–2615.40</td>
<td>HDO, CO(_2)</td>
</tr>
<tr>
<td></td>
<td>2650.60–2651.30</td>
<td>HDO, CO(_2)</td>
</tr>
<tr>
<td></td>
<td>2835.50–2835.80</td>
<td>HDO</td>
</tr>
<tr>
<td></td>
<td>2903.60–2904.03</td>
<td>HDO, H(_2)O</td>
</tr>
<tr>
<td></td>
<td>2921.00–2921.60</td>
<td>HDO, H(_2)O</td>
</tr>
<tr>
<td>N(_2)O</td>
<td>2806.20–2806.48</td>
<td></td>
</tr>
<tr>
<td>HCHO</td>
<td>2763.425–2763.600</td>
<td>CH(_4), HDO, N(_2)O, O(_3), CO(_2), solar CO</td>
</tr>
<tr>
<td></td>
<td>2765.725–2765.975</td>
<td>CH(_4), HDO, N(_2)O, O(_3), CO(_2), solar CO</td>
</tr>
<tr>
<td></td>
<td>2778.200–2778.590</td>
<td>CH(_4), HDO, N(_2)O, O(_3), CO(_2), solar CO</td>
</tr>
<tr>
<td></td>
<td>2780.800–2781.150</td>
<td>CH(_4), HDO, N(_2)O, O(_3), CO(_2), solar CO</td>
</tr>
<tr>
<td></td>
<td>2810.000–2810.350</td>
<td>CH(_4), HDO, N(_2)O, O(_3), solar CO</td>
</tr>
<tr>
<td></td>
<td>2855.650–2856.400</td>
<td>CH(_4), HDO, H(_2)O, O(_3), CO(_2), solar CO</td>
</tr>
</tbody>
</table>
Table 2. Mean of the FTIR error budget (in %) on formadehyde individual total columns, and on the daily mean total columns used for the comparisons with MAX-DOAS, SCIAMACHY and IMAGESv2.

<table>
<thead>
<tr>
<th>Errors</th>
<th>Individual [%]</th>
<th>Daily mean [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothing</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Measurement noise</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Temperature</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Interfering species</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Instrumental Line Shape</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total random error</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>Intensity</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pressure broadening</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total systematic error</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Total error</td>
<td>20</td>
<td>18</td>
</tr>
</tbody>
</table>
Table 3. Forward model parameters b, their uncertainty, and the corresponding forward model parameters errors. 2.5 K is the annual variability of the NCEP temperature at 1000 hPa. 0.03 is the annual variability of the surface albedo at 20.9° S, 55.5° E (Koelemeijer et al., 2003). For ozone, a variability of 10% has been arbitrarily chosen. Regarding the aerosols, the variability in the AOD (Aerosol Optical Depth) estimated from the aerosol profiles shown in Fig. 5 is 0.1. The total forward model parameters error is calculated by adding in quadrature the different error sources.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Uncertainty on b</th>
<th>Error on HCHO total column (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>±2.5 K</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Albedo</td>
<td>±0.03</td>
<td>1</td>
</tr>
<tr>
<td>Aerosols</td>
<td>±0.1</td>
<td>12</td>
</tr>
<tr>
<td>Ozone</td>
<td>±10%</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>12.1</td>
</tr>
</tbody>
</table>

For ozone, a variability of 10% has been arbitrarily chosen.
Table 4. Error budget (in %) for the HCHO total columns retrieved from MAX-DOAS observations.

<table>
<thead>
<tr>
<th>Errors</th>
<th>Individual [%]</th>
<th>Daily mean [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothing error</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Retrieval noise</td>
<td>9</td>
<td>1.5</td>
</tr>
<tr>
<td>Forward model parameter error</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Total random error</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Systematic error</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Total error</td>
<td>27</td>
<td>22</td>
</tr>
</tbody>
</table>
Table 5. Weighted mean differences, and standard deviations, (1) between the FTIR and MAX-DOAS daily mean total columns, and (2) between the FTIR “smoothed total columns” (calculated from the FTIR profiles which have been smoothed with the MAX-DOAS averaging kernels) and the MAX-DOAS daily mean total columns. The mean of the random and systematic errors on the differences are also given. All values are in %.

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Mean ± Std [%]</th>
<th>Systematic error [%]</th>
<th>Random error [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTIR – DOAS</td>
<td>−0.3±18.3</td>
<td>13.1</td>
<td>26.4</td>
</tr>
<tr>
<td>FTIR\textsubscript{smooth} – DOAS</td>
<td>−8.4±7.8</td>
<td>10.0</td>
<td>4.7</td>
</tr>
</tbody>
</table>
Table 6. Weighted mean differences, and standard deviations, (1) between the ground-based (FTIR or MAX-DOAS) and SCIAMACHY daily mean total columns, and (2) between the ground-based “smoothed total columns” (calculated from the FTIR (or MAX-DOAS) profiles which have been smoothed with the SCIAMACHY total column averaging kernels) and the SCIAMACHY daily mean total columns. The mean of the random and systematic errors on the differences are also given. All values are in %.

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Mean ± Std [%]</th>
<th>Systematic error [%]</th>
<th>Random error [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTIR – SCIA 2004</td>
<td>−4.6±20.3</td>
<td>39.9</td>
<td>22.1</td>
</tr>
<tr>
<td>FTIR_smooth – SCIA</td>
<td>−3.6±20.6</td>
<td>39.8</td>
<td>21.7</td>
</tr>
<tr>
<td>FTIR – SCIA 2007</td>
<td>+17.0±54.3</td>
<td>56.8</td>
<td>37.0</td>
</tr>
<tr>
<td>FTIR_smooth – SCIA</td>
<td>+18.1±54.4</td>
<td>56.5</td>
<td>35.4</td>
</tr>
<tr>
<td>DOAS – SCIA</td>
<td>−11.2±30.5</td>
<td>42.3</td>
<td>30.6</td>
</tr>
<tr>
<td>DOAS_smooth – SCIA</td>
<td>−16.9±31.0</td>
<td>42.3</td>
<td>32.5</td>
</tr>
</tbody>
</table>
Table 7. Performed simulations.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Standard</td>
</tr>
<tr>
<td>S2</td>
<td>Use OH fields from Spivakovsky et al. (2000)</td>
</tr>
<tr>
<td>S3</td>
<td>Use Cl fields in MBL from Allan et al. (2007)</td>
</tr>
</tbody>
</table>
Table 8. Weighted mean differences, and standard deviations (in %), between the ground-based and the modeled HCHO total columns. In all cases, the averaging kernels of the instruments were applied in the calculation of the simulated columns.

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Mean ± Std [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTIR – IMAGES S1 2004</td>
<td>30.5±22.1</td>
</tr>
<tr>
<td>FTIR – IMAGES S2 2004</td>
<td>8.5±19.6</td>
</tr>
<tr>
<td>FTIR – IMAGES S1 2007</td>
<td>25.0±26.9</td>
</tr>
<tr>
<td>FTIR – IMAGES S2 2007</td>
<td>9.4±25.5</td>
</tr>
<tr>
<td>DOAS – IMAGES S1</td>
<td>14.5±12.7</td>
</tr>
<tr>
<td>DOAS – IMAGES S2</td>
<td>8.0±12.9</td>
</tr>
</tbody>
</table>
Fig. 1. Formaldehyde a priori profile used in the FTIR and MAX-DOAS retrievals, in vmr (left) and concentration (right) units.
Fig. 2. FTIR volume mixing ratio averaging kernels (ppv/ppv) for the altitudes listed in the legend.
Fig. 3. Extinction profile given by the aerosol model of Shettle (1989) included in UVspec/DISORT and corresponding to oceanic conditions with a visibility of 100 km (black dashed line), and example of a “perturbed” extinction profile (red solid line) corresponding to values of parameters A, B, C equal to 0.2 km, 0.4 km$^{-1}$, and 0.35 km$^{-1}$.
Fig. 4. Example of aerosol extinction profile selection (25/11/2004 morning) based on O₄ Langley-plot. The selected extinction profile and corresponding Langley-plot appear on the lefthand side while an example of a not selected extinction profile and corresponding Langley-plot is presented on the righthand side. The RMS of the linear fit is $166 \times 10^{40}$ mol²/cm⁵ for the selected aerosol extinction profile (left) and $813 \times 10^{40}$ mol²/cm⁵ for the other one (right).
Fig. 5. Different aerosol extinction profiles selected for the clear-sky days during the August 2004–July 2005 period at Réunion Island. Parameters A, B, and C are in km, km$^{-1}$, and km$^{-1}$, respectively.
Fig. 6. Examples of retrieval fit results obtained on 25 November 2004 morning for some MAX-DOAS scans between 58° and 68° SZA.

- 25/11/2004 AM – 67.9° SZA
- 25/11/2004 AM – 64.6° SZA
- 25/11/2004 AM – 61.4° SZA
- 25/11/2004 AM – 58.1° SZA

The information content is characterized through the calculation of the average correlation coefficient (ACC) as a function of the number of elevation angles (2, 8, 15, and 30°) available for each measurement scan. As seen in Sect. 2.2.3, the averaging kernels (rows of the matrix) are rotated in the vertical direction. A sensitivity analysis indicates that a DOFS of up to 10 could be achieved by reducing the noise on HCHO measurements.

As already discussed in Sect. 2.3, the total retrieval error consists of the sum of the smoothing error, the forward model parameters error, and the measurement noise error. They are typically small and all less than 2% of the HCHO DSCDs (0.45 to 1.6 × 10^16 molecules/cm^2) for this campaign down to 1.6 × 10^15 molecules/cm^2 for our most recent instrument operational performance. The error budget is presented in Table 1 and in Sect. 3.3.3.

We see that, according to the MAX-DOAS geometry, most of the information on the vertical distribution of HCHO is contained in the measurements located below 2.5 km.
Fig. 7. Averaging kernels (ppv/ppv) obtained for MAX-DOAS observations of HCHO at Réunion Island on 26 August 2004 AM using measurements at all elevation angles at a SZA of about 64°.
Fig. 8. Daily means of formaldehyde total columns above Saint-Denis from FTIR, MAX-DOAS and SCIAMACHY measurements during the 2004 campaign.
Fig. 9. Formaldehyde daily mean FTIR total columns and daily averaged SCIAMACHY columns in a region of 500 km around Saint-Denis during the 2007 campaign.
Fig. 10. FTIR, MAX-DOAS and SCIAMACHY total column averaging kernels.

Table 6. Weigthed mean differences, and standard deviations, (1) between the ground-based (FTIR or MAX-DOAS) and SCIAMACHY daily mean total columns, and (2) between the ground-based “smoothed total columns” (calculated from the FTIR (or MAX-DOAS) profiles which have been smoothed with the SCIAMACHY total column averaging kernels) and the SCIAMACHY daily mean total columns. The mean of the random and systematic errors on the differences are also given. All values are in %.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-4.6 ± 20.3</td>
<td>-3.6 ± 20.6</td>
<td>+17.0 ± 54.3</td>
<td>+18.1 ± 54.4</td>
</tr>
<tr>
<td>Std</td>
<td>39.9</td>
<td>39.8</td>
<td>56.8</td>
<td>56.5</td>
</tr>
<tr>
<td>Random</td>
<td>22.1</td>
<td>21.7</td>
<td>37.0</td>
<td>35.4</td>
</tr>
<tr>
<td>System</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 11: Daily means of formaldehyde total columns above Saint-Denis from MAX-DOAS and SCIAMACHY measurements from August 2004 to August 2005.
Table 6. Weighted mean differences, and standard deviations, (1) between the ground-based (FTIR or MAX-DOAS) and SCIAMACHY daily mean total columns, and (2) between the ground-based “smoothed total columns” (calculated from the FTIR (or MAX-DOAS) profiles which have been smoothed with the SCIAMACHY total column averaging kernels) and the SCIAMACHY daily mean total columns. The mean of the random and systematic errors on the differences are also given. All values are in %.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± Std</td>
<td>Systematic Error</td>
<td>Random Error</td>
<td>Systematic Error</td>
</tr>
<tr>
<td></td>
<td>-4.6 ± 20.3</td>
<td>39.9</td>
<td>22.1</td>
<td>39.8</td>
</tr>
<tr>
<td></td>
<td>-3.6 ± 20.6</td>
<td>39.8</td>
<td>21.7</td>
<td>39.7</td>
</tr>
<tr>
<td></td>
<td>+17.0 ± 54.3</td>
<td>56.8</td>
<td>37.0</td>
<td>56.5</td>
</tr>
<tr>
<td></td>
<td>+18.1 ± 54.4</td>
<td>56.5</td>
<td>35.4</td>
<td>56.4</td>
</tr>
<tr>
<td>DOAS – SCIA</td>
<td>-11.2 ± 30.5</td>
<td>42.3</td>
<td>30.6</td>
<td>42.3</td>
</tr>
<tr>
<td>DOAS smooth – SCIA</td>
<td>-16.9 ± 31.0</td>
<td>42.3</td>
<td>32.5</td>
<td>42.4</td>
</tr>
</tbody>
</table>

Fig. 11. Daily means of formaldehyde total columns above Saint-Denis from MAX-DOAS and SCIAMACHY measurements from August 2004 to August 2005.
Fig. 12. FTIR and IMAGESv2 formaldehyde total columns during the two FTIR campaigns in 2004 (left) and 2007 (right). The arrows correspond to dates mentioned in the discussion (Sect. 7): the 14 August, 14 and 29 September, 12 and 21 October 2004 (left plot); and the 3, 5, 11 and 12 June, 6, 18, and 31 August, 24 September, and 11 October 2007 (right plot).
Fig. 13. MAX-DOAS and IMAGESv2 formaldehyde total columns.
Fig. 14. Daily means of FTIR HCHO (in blue) and CO (in red) (Senten et al., 2008) total columns at Réunion Island in 2004 (left) and 2007 (right). The CO total columns have been multiplied by a factor $2 \times 10^{-3}$. Only coincident days are shown.
**Fig. 15.** CO tracer source strength (in ppb) as obtained from FLEXPART/GFEDv2, after a 1 day backward run. The total emission mass flux has been evenly distributed between 0 and 6 km. See text for details.
Fig. 15. CO tracer source strength (in ppb) as obtained from FLEXPART/GFEDv2, after a 1 day backward run. The total emission mass flux has been evenly distributed between 0 and 6 km. See text for details.

Fig. 16. Average CO emission source strength from biomass-burning in Madagascar (in g m\(^{-2}\) month\(^{-1}\)), as derived from the GFEDv2 database.