Carbon monoxide total columns from SCIAMACHY 2.3 µm atmospheric reflectance measurements: towards a full-mission data product (2003–2012)

T. Borsdorff1, P. Tol1, J. E. Williams2, J. de Laat2, J. aan de Brugh1, P. Nédélec3, I. Aben1, and J. Landgraf1

1SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
2Royal Netherlands Meteorological Institute (KNMI), de Bilt, the Netherlands
3Laboratoire d’aérologie (LA), CNRS UMR-5560 et Observatoire Midi-Pyrénées, Université Paul-Sabatier, Toulouse, France

Correspondence to: T. Borsdorff (t.borsdorff@sron.nl)

Received: 10 August 2015 – Published in Atmos. Meas. Tech. Discuss.: 17 September 2015
Revised: 5 January 2016 – Accepted: 6 January 2016 – Published: 26 January 2016

Abstract. We present a full-mission data product of carbon monoxide (CO) vertical column densities using the 2310–2338 nm SCIAMACHY reflectance measurements over clear-sky land scenes for the period January 2003–April 2012. The retrieval employs the SICOR algorithm, which will be used for operational data processing of the Sentinel-5 Precursor mission. The retrieval approach infers simultaneously carbon monoxide, methane and water vapour column densities together with a Lambertian surface albedo from individual SCIAMACHY measurements employing a non-scattering radiative transfer model. To account for the radiometric instrument degradation including the formation of an ice-layer on the 2.3 µm detector array, we consider clear-sky measurements over the Sahara as a natural calibration target. For these specific measurements, we spectrally calibrate the SCIAMACHY measurements and determine a spectral radiometric offset and the width of the instrument spectral response function as a function of time for the entire operational phase of the mission. We show that the smoothing error of individual clear-sky CO retrievals is less than ±1 ppb and this error contribution does not need to be accounted for in the validation considering the much higher retrieval noise. The CO data product is validated against measurements of ground-based Fourier transform infrared spectrometers at 27 stations of the NDACC-IRWG and TCCON network and MOZAIC/IAGOS aircraft measurements at 26 airports worldwide. Overall, we find a good agreement with TCCON measurements with a mean bias $\bar{b} = -1.2$ ppb and a station-to-station bias with $\sigma = 7.2$ ppb. The negative sign of the bias means a low bias of SCIAMACHY CO with respect to TCCON. For the NDACC-IRWG network, we obtain a larger mean station bias of $\bar{b} = -9.2$ ppb with $\sigma = 8.1$ ppb and for the MOZAIC/IAGOS measurements we find $\bar{b} = -6.4$ ppb with $\sigma = 5.6$ ppb. The SCIAMACHY data set is subject to a small but significant bias trend of $1.47 \pm 0.25$ ppb yr$^{-1}$. After trend correction, the bias with respect to MOZAIC/IAGOS observation is 2.5 ppb, with respect to TCCON measurements it is $-4.6$ ppb and with respect to NDACC-IRWG measurements $-8.4$ ppb. Hence, a discrepancy of 3.8 ppb remains between the global biases with NDACC-IRWG and TCCON, which is confirmed by directly comparing NDACC-IRWG and TCCON measurements. Generally, the scatter of the individual SCIAMACHY CO retrievals is high and dominated by large measurement noise. Hence, for practical usage of the data set, averaging of individual retrievals is required. As an example, we show that monthly mean SCIAMACHY CO retrievals, averaged separately over Northern and Southern Africa, reflect the spatial and temporal variability of biomass burning events in agreement with the global chemical transport model TM5.
1 Introduction

Carbon monoxide (CO) is an important atmospheric trace gas for the understanding of tropospheric chemistry and air quality. Its main source is incomplete combustion of fossil fuel and biomass and the oxidation of atmospheric methane and other hydrocarbons. The reaction of CO with the hydroxyl radical (OH) represents its major atmospheric sink and thus CO regulates the self-cleaning capability of the atmosphere (Spivakovsky et al., 2000). Enhanced CO concentration can indicate anthropogenic air pollution (Logan et al., 1981) and as a precursor of ozone (O$_3$) formation it influences tropospheric air quality (Seiler and Fishman, 1981). Moreover, by constraining the depletion of methane (CH$_4$), CO affects indirectly global warming (Daniel and Solomon, 1998) and due to its moderately long lifetime of several weeks to several months (Holloway et al., 2000), it is a tracer for global transport and redistribution of pollutants in the atmosphere (e.g. Yurganov et al., 2004, 2005; Gloudemans et al., 2006).

Since 2000, the global concentration of CO has been measured by various satellite missions. For example, the MOPITT (Measurements of Pollution in the Troposphere) instrument uses the spectral measurements at 2.3 and 4.7 µm to retrieve CO (Deeter et al., 2003). AIRS (Atmospheric Infrared Sounder; McMillan et al., 2005), launched in 2002 onboard the Aqua satellite, TES (Tropospheric Emission Spectrometer; Rinsland et al., 2006) and IASI (Infrared Atmospheric Sounding Interferometer; Turquety et al., 2004), onboard of a series of three METOP (Meteorological Operational) satellites, employ spectral measurements at 4.7 µm to infer atmospheric CO abundances. The Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) was one of the first space-based instruments observing CO from the shortwave infrared (SWIR) range around 2.3 µm (Bovensmann et al., 1999) and it was fully operational from January 2003 until April 2012 when contact with its host ENVISAT (Environmental Satellite) was lost. In this period, an almost continuous long-term record of more than 9 years of SWIR measurements in the 2.3 µm spectral range from space was recorded. For cloud-free scenes, these spectra are sensitive to the total column density of CO with a good vertical sensitivity throughout the whole atmosphere (Buchwitz et al., 2004; Gloudemans et al., 2008).

In recent years, several algorithms have been developed to infer CO total columns from SCIAMACHY’s SWIR measurements – e.g. IMAP-DOAS (Iterative Maximum A Posteriori Differential Optical Absorption Spectroscopy; Frankenberger et al., 2005), WFM-DOAS (Weighting Function Modified Differential Optical Absorption Spectroscopy; Buchwitz et al., 2004), IMLM (Iterative Maximum Likelihood Method; Gloudemans et al., 2009), and the operational SCIAMACHY CO processor (Gimeno García et al., 2011). The global CO fields were used for a range of applications – e.g. the detection of biomass burning events (Buchwitz et al., 2004), to study the inter-annual variability of CO on the global scale (Gloudemans et al., 2009), to investigate pollution patterns of megacities (Buchwitz et al., 2007) and the long-range transport of CO in the Southern Hemisphere (Gloudemans et al., 2009), which indicates the broad scope of application for this data product. Furthermore, the SCIAMACHY CO measurements were compared with corresponding MOPITT CO retrievals (de Laat et al., 2010a) and additionally validated with CO observations of ground-based spectrometers (de Laat et al., 2010b) and MOZAIC/IAGOS aircraft measurements (de Laat et al., 2012). All these previous studies were dedicated to the early years of the mission before 2009. A possible reason for this is the extensive degradation of the instrument (Gloudemans et al., 2008), caused by the growing ice layer on the detector array and a considerable loss of detector pixels due to radiation damage in the later years of the mission. This reduces the radiometric quality of the SCIAMACHY spectra, which seriously complicates the processing of a SCIAMACHY CO product for the entire mission period.

The Tropospheric Monitoring Instrument (TROPOMI) on board of the Sentinel 5 Precursor (SP-5) mission is expected to be launched in 2016. TROPOMI covers the same 2.3–2.4 µm spectral range as SCIAMACHY with the same spectral resolution but with an improved radiometric performance and a better spatial resolution of the TROPOMI instrument. For the S5-P mission, the highly efficient Shortwave Infrared Carbon Monoxide Retrieval algorithm (SICOR) (Vidot et al., 2012) was developed to meet the demanding requirements of operational data processing regarding calculation time. In this study, we apply the SICOR algorithm to the SWIR measurements of the SCIAMACHY instrument and infer a data set of CO vertical columns for the entire ENVISAT mission (2003–2012), limited to land and cloud-free scenes. This study represents the first application of the TROPOMI operational processor to real data.

Due to unexpected in-orbit problems of the SCIAMACHY measurements in the 2.3 µm spectral range (Gloudemans et al., 2005), recalibration of the radiometric measurements is needed. For this purpose, we use clear-sky measurements over the Sahara as a natural calibration target in combination with accurate a priori knowledge of the atmospheric methane abundances in this region. For the entire mission lifetime, we determine the temporal dependence of the spectral calibration, a spectral radiometric offset, and the width of the instrument spectral response function from these measurements. Furthermore, we use SCIAMACHY’s solar measurements to obtain a proper reflectance retrieval. Here, multiplicative radiometric errors common to both the radiance and irradiance measurement cancel out and are thus not relevant for the retrieval. Finally, the CO data set is validated with ground-based measurements of the TCCON (Total Carbon Column Observing Network) and NDACC-IRWG (Network for the Detection of Atmospheric Composition Change – Infrared Working Group) at 27 sites and aircraft measurements close to 26 airports of the MOZAIC (Measurement of Ozone
To obtain CO vertical column densities, we use SCIAMACHY SWIR vertical measurements in the spectral range 2310–2338.4 nm with a spectral resolution of 0.2 nm and a spectral sampling distance of 0.1 nm. The retrieval is based on the profile scaling approach, which was first applied by Gloudemans et al. (2008) to interpret SCIAMACHY data. The approach is discussed in detail by Borsdorff et al. (2014) and this section summarises its main characteristics.

Basically, the retrieval approach scales an \( n \)-dimensional reference profile \( \rho_{\text{ref}} \), which is the input to a radiative transfer model, to fit SCIAMACHY reflectance measurements. Hence, \( \rho_{\text{ref}} \) describes the vertical concentration of an atmospheric trace gas in arbitrary units. Subsequently, we estimate the retrieved CO vertical column density \( c \) by

\[
c = C^T \alpha \rho_{\text{ref}},
\]

with the profile scaling factor \( \alpha \). Here, the \( n \)-dimensional vector \( C = (f_1, \ldots, f_n) \) approximates the vertical integration, where \( f_k \) converts the \( k \)-th element of the state vector to the corresponding partial column amount of the trace gas. For the sake of simplicity, we refer to the retrieval of the total column density \( c \) when meaning this approach in the following.

For the inversion, a forward model \( F \) is needed, which describes the \( m \)-dimensional measurement \( y_{\text{meas}} \) within the spectral error \( e_y \), namely

\[
y_{\text{meas}} = F(x, b) + e_y.
\]

Here, state vector \( x \) contains all parameters to be retrieved including the column density of CO and other trace gases. The forward model vector \( b \) includes all parameters which are needed for the simulation but are assumed to be known a priori. For the measurement, we employ a non-scattering radiative transfer model (Vidot et al., 2012) which simulates atmospheric transmission including Lambertian reflection at the Earth surface. Figure 1 shows a typical transmission spectrum in the retrieval window for clear-sky conditions and the individual spectral contributions of the trace gases HDO, CO, H\(_2\)O, and CH\(_4\). The forward model employs the cross-section database by Gloudemans et al. (2009), which comprises CO and CH\(_4\) absorption cross-sections from the high-resolution transmission molecular absorption database (HI-TRAN) Rothman et al. (2005) and Predoi-Cross et al. (2006), respectively, and H\(_2\)O and HDO cross sections from Jenouvrier et al. (2007). The spectral fit window is extended significantly with respect to the window used by Gloudemans et al. (2008) to establish a stable retrieval for the entire mission period. This is particularly important for the later years of the mission with a significant loss of spectral pixels of the SCIAMACHY channel 8 detector due to radiation damage. The selected window includes strong absorption lines of CH\(_4\) between 2315 and 2320 nm and a nearly translucent range in the range 2310–2315 nm. Both spectral features are needed to mitigate the degradation of the instrument by fitting effective instrument parameters as described in the following section.

To invert Eq. (2), we employ a Gauss–Newton iteration scheme where the forward model \( F \) is linearised each iteration step around the solution \( x_0 \) of the previous iteration. Thus, we can rewrite Eq. (2) as

\[
y = Kx + e_y
\]
where \(\|\cdot\|_2\) represents the \(L_2\) norm and \(S_y \in \mathbb{R}^{m \times m}\) is the non-singular measurement error covariance matrix. Simultaneously with CO, we retrieve the vertical column densities of HDO, H\(_2\)O and CH\(_4\) from the SWIR measurements using per species the explained profile scaling approach. Additionally, we infer a wavelength-dependent albedo described by a quadratic polynomial with respect to wavelength. The solution of Eq. (4) can be expressed by the gain matrix \(G\):

\[
x_{\text{col}} = G y
\]

with

\[
G = (K^T S_y^{-1} K)^{-1} K^T S_y^{-1}.
\]

The retrieved vertical column density \(c_{\text{col}}\) is an effective column product due to the regularisation inherent to the profile scaling approach. The relation between the effective column and the true atmospheric abundance is described by the total column averaging kernel \(a_{\text{col}}\):

\[
c_{\text{col}} = a_{\text{col}} \rho_{\text{true}} + e_c,
\]

where \(e_c\) is the column retrieval error due to the measurement error \(e_y\) and \(\rho_{\text{true}}\) is the true trace gas profile. A numerically efficient algorithm to calculate \(a_{\text{col}}\) is presented in Borsdorff et al. (2014). The total column averaging kernel represents an altitude-weighted integration of the true profile taking into account the particular retrieval sensitivity. The differences between the true column, \(c_{\text{true}} = C^T \rho_{\text{true}}\), and the effective column, \(c_{\text{eff}} = a_{\text{col}} \rho_{\text{true}}\), cannot be inferred from the measurement and is also known as the null space or smoothing error of the retrieval (Borsdorff et al., 2014; Rodgers, 2000),

\[
e_{\text{null}} = (C^T - a_{\text{col}}) \rho_{\text{true}}.
\]

Finally, we characterise the noise on the retrieval product due to the measurement noise, described by the retrieval noise covariance

\[
S_t = GS_y G^T.
\]

In this manner, we have defined all diagnostic tools for our retrieval. A detailed overview of the profile-scaling approach is given in Borsdorff et al. (2014).

The retrieval depends on a priori information, which is adopted from different sources. Surface pressure, temperature profiles and water vapour reference profiles (H\(_2\)O, HDO) are based on the ECMWF (The European Centre for Medium-Range Weather Forecasts) ERA-Interim (Re-Analysis Interim) data set, which is sampled every 6 h on 60 vertical layers and on a 0.75\(^\circ\) latitude by 0.75\(^\circ\) longitude grid (Dee et al., 2011). CO and CH\(_4\) reference profiles are taken from 3-dimensional global chemistry transport simulations of TM5 covering the entire mission period (Williams et al., 2013, 2014), which includes seasonal and interannual variability of the a priori information on the relative CO profile. Atmospheric trace gas profiles are provided every 3 h on 34 layers and on a 2\(^\circ\) latitude by 3\(^\circ\) longitude grid. For every SCIAMACHY measurement, the model data are spatially resampled to the satellite ground pixel and interpolated in time. Moreover, we account for the differences between the mean SCIAMACHY pixel elevation and the mean pixel elevation of the model data. First, we calculate the mean SCIAMACHY pixel height using the digital Shuttle Radar Topography Mission (SRTM) elevation map with a spatial resolution of 15 arcsec (Farr et al., 2007) and subsequently, all model profiles are interpolated to the mean altitude of a SCIAMACHY ground pixel.

One may question the relevance of the null-space error and the need for column averaging kernels for a proper validation of our data product. Generally, the correct use of Eq. (7) requires measurements of the CO vertical profile. However, this hampers any validation of the SCIAMACHY CO data product because measurements of CO profiles are hardly available for the mission period. On the other hand, a direct comparison of ground-based measurements of the total CO column with our data product suffers from the null-space error. Borsdorff et al. (2014) showed from simulations that in the presence of clouds the null-space error can easily exceed 30\% of the CO total column. The error is much smaller for clear sky conditions, depending on the reference profile used for scaling. To estimate the null-space error, we consider simulated retrievals for a set of solar zenith angles between 20 and 70\(^\circ\). Here, we used the US standard atmosphere (NOAA, 1976) for the profiles of dry air density, pressure, water and CO. The CH\(_4\) profile is adopted from the CAMELOT (Chemistry of the Atmosphere Mission Concepts and Sentinel Observations Techniques) European background model atmosphere (Levelt et al., 2009). The total column averaging kernels are shown in the left panel of Fig. 2. Subsequently, we investigate the null-space error due to the difference of 533
CO profiles measured by the HIAPER Pole-to-Pole Observations (HIPPO) of the Carbon Cycle and Greenhouse Gases Study (Wofsy, 2011; Wofsy et al., 2012) and two different choices for the reference profiles. First, we consider the CO US standard profile and, second, we make use of the collocated CO profiles from the TM5 chemical transport model, which is the baseline of our algorithm. After scaling to the same total column, the variation of the HIPPO profiles and the corresponding reference profiles are shown in the middle panel of Fig. 2. Finally, the right panel of the figure shows the corresponding distribution of the null-space error utilising the column averaging kernels of the same figure. For both cases, the null-space error is less than 1 ppb (< 1 % of a mean CO total column) and so far less than the SCIAMACHY measurement noise error that varies between 30 ppb and > 170 ppb for individual retrievals (see Figs. 6 and 7).

Because of its randomness the noise error can be reduced by averaging multiple SCIAMACHY CO retrievals, which is not necessarily the case for the null space. However, a null-space error of the order of < 1 % for clear-sky SCIAMACHY CO retrievals represents a minor contribution to the overall error and is ignored in the following. So a direct comparison between ground-based measurements and SCIAMACHY retrievals is possible. Following this approach, strict cloud filtering of SCIAMACHY data is required. For this purpose, we employ the SCIAMACHY polarisation device (PMD) Identification of Clouds and Ice (SPICI) algorithm (Krijger et al., 2005).

3 Instrument calibration

In this section, we consider SCIAMACHY nadir measurements for the full operational phase of the mission from January 2003 to April 2012, where we use Level 1b spectra disseminated by ESA. The measurements are corrected for memory-nonlinearity and dark current using the Netherlands SCIAMACHY Data Center (NADC) toolbox Version 1.2 (http://www.sron.nl/~richardh/). Data recorded during SCIAMACHY’s commissioning phase are not considered. The CO data processing relies on SCIAMACHY’s forward scans for a solar zenith angle smaller than 80° and with a ground pixel size of about 30 km × 120 km (along-track × across-track) for an integration time of 0.5 s. At higher northern and southern latitudes the integration time is increased to 1 s, which accordingly doubles the across-track pixel size.

The SCIAMACHY detector in the 2.3–2.4 µm range (channel 8) suffers from detector radiation damage and very noisy detector pixels. Buchwitz et al. (2007) and Gloudemans et al. (2008) showed the severe sensitivity of CO retrievals to the instrument malfunction. A stable retrieval performance for the entire mission lifetime requires careful and strict spectral filtering of bad detector pixels based on in-flight detector performance monitoring. For measurements after 11 January 2005, we utilise the NADC version 3.0 time-dependent pixel mask, and for earlier measurements we fixed the pixel mask to that of this reference date. The channel 8 SCIAMACHY measurement noise is dominated by detector dark noise, which is estimated from SCIAMACHY’s daily dark state measurements taken during the orbit eclipse.

Despite the strict filtering, the absolute radiometric calibration of SCIAMACHY Earthshine measurements is not accurate enough to retrieve CO. Figure 3 shows the time dependence of the mean solar signal. Besides the seasonal variation of the signal due to the change of the Earth–Sun distance throughout the year, the effect of the ice layer formation on the overall instrument transmission and the signal recovery due to the instrument heating during so-called decontami-
nation events (see Table 1) is clearly visible (Gloudemans et al., 2005). To mitigate effects on the CO data quality, our retrieval is based on the reflectance \( r_i \), which is the ratio of the Earth radiance measurement \( I_i \) divided by the solar measurement \( S_i \) by the same detector pixel \( i \). Here, we use SCIAMACHY’s daily Sun mean reference measurements, determined from the Sun measurements via the elevation scan mirror. Subsequently, the solar measurements are interpolated to the measurement time of the Earthshine observation to account for a rapidly changing instrument directly after a thermal decontamination event.

The approach implies that any common multiplicative radiometric error of the Earth and solar observations cancels out in the reflectance ratio. However, any additive error component, e.g. due to detector hysteresis, non-linear radiometric detector response, dark detector current and analogue offset, still affects the radiometric accuracy of the reflectance spectrum and has to be accounted for by the radiometric calibration procedure.

Moreover, the use of SCIAMACHY reflectance measurements is hampered by the different malfunctioning pixels of the SCIAMACHY channel 8 detector for the Earthshine and solar observation mode. Filtering on both types of detector performance results in insufficient spectral coverage. To overcome this problem, we identify outliers in the solar irradiance spectrum and replace them by interpolated values. We start with a solar reference measurement \( S(\tau_0, i) \) from 11 January 2003 (Fig. 4, upper panel), which is representative of a well-performing nearly ice-free detector (see Fig. 3). The nearly linear dependence of the solar signal on wavelength is due to the spectral variation of the detector pixel quantum efficiency. To detect spectral outliers, we determine the relative difference of the spectrum with respect to its running median spectrum, assuming an average over 1.4 nm. Differences between the original and spectrally smoothed solar spectrum of > 7% are classified as outliers and are replaced by the smoothed value of the running median spectrum. Smaller differences are attributed to the pixel-to-pixel gain variation of the detector and the measurement noise and these features are maintained in the spectrum. This approach is based on the assumption that spectral variations in the solar spectrum in the considered spectral range are smooth.

In the following, we assume that the degradation of the solar spectrum can be described by

\[
S_i(t) = \alpha(t) \cdot \beta_i(t) \cdot S_i(t_0) + \epsilon_i(t),
\]

where \( \alpha \) describes the relative degradation of the mean signal shown in Fig. 3, \( \beta \) represents the relative spectral degradation of detector pixel \( i \) and \( \epsilon_i \) summarises high-frequency error contributions including noise and outliers. The middle panel of Fig. 4 shows the ratio \( S_i(t)/(\alpha(t)S_i(t_0)) \) for three exemplary days in the year 2003. Applying a 1.0 nm running median suppresses high-frequency contributions and allows us to estimate the degradation function \( \beta \) from the data. Subsequently, this defines also the error contribution \( \epsilon_i(t) \) in Eq. (10), which can be used to detect spectral outliers in the solar measurement \( S_i(t) \). Any measurement with \( \epsilon_i(t) > 2\% \) is classified as an outlier and is replaced by the expected value \( \alpha(t)\beta_i(t)S_i(t_0) \). In summary, the approach allows us to replace corrupted signals by interpolated values assuming a spectrally smooth degradation. High-frequency pixel-to-pixel variation present in the reference spectrum \( S(t_0) \) are considered to be constant over the entire mission lifetime.

To account for an imperfect calibration resulting in an additive radiometric bias, we consider the Sahara region between 30° and 15° northern latitude and −15 and 30° longitude as a natural calibration target for the entire mission period. This region is chosen because of the high signal levels due to the high reflective desert surface, and because it is assured that the amount of CH₄ can be relatively well predicted using the TM5 model (Gloudemans et al., 2005). For measurements over this particular region, we modify our forward model by adding a polynomial expansion of an additive
radiometric bias,

\[ F(x, b, a) = \hat{F}(x, b) + \sum_{i=0}^{3} a_i \cdot p_i(\lambda). \]  

(11)

Here, \( \hat{F}(x, b) \) denotes the forward calculation in Eq. (2), and \( p_i \) are Chebyshev polynomials as function of wavelength \( \lambda \). The coefficients \( a_i \) can be determined as additional fit parameters of the retrieval algorithm because of the high radiometric signal over the desert region, where we fix the atmospheric methane abundance to the a priori model information. To fully exploit this approach, it was necessary to include the strong \( \text{CH}_4 \) absorption between 2315–2320 nm in our spectral fitting window (see Fig. 1).

Figure 5 shows the temporal evolution of coefficient \( a_0 \), which represents a spectrally constant additive bias of the measurement. Here \( a_0 \) increases with a growing ice layer while the overall optical throughput of the instrument declines (see Fig. 3). For a fully established ice layer, the offset is 20–30 %. We attribute this offset to photons scattered in the ice layer and then detected at a spectrally shifted position on the detector. In other words, the effective spectral instrument response function is altered by the ice layer (Gloudemans et al., 2005). To account for this significant bias in our overall retrieval, we smooth the data over a 40-day period and correct all SCIAMACHY measurements accordingly. To demonstrate the general applicability of our approach to global data sets, we applied the same procedure for corresponding cloud-free measurements over Australia with a lower surface albedo and with different solar geometries. We obtained very similar radiometric biases (see Fig. 5) which supports the overall validity of the approach. The difference between the coefficients derived over Sahara and Australia (about 3 % in spring and in agreement in autumn) is not fully understood yet and is the topic of further investigation.

Subsequently, we evaluate the spectral calibration and the SCIAMACHY instrument spectral response function. Based on gas-cell measurements during the on-ground calibration of the instrument, Schrijver (1999, 2000b, 2001b) suggested using a quadratic polynomial in pixel number for the wavelength calibration for the channel 8 detector,

\[ \lambda = a_0 + a_1 \cdot n + a_2 \cdot n^2. \]  

(12)
where wavelength $\lambda$ is given in nm and $n$ denotes the spectral pixel number. For the purpose of this study, we adopted coefficients $a_1 = 0.135254$ nm and $a_2 = -1.19719 \times 10^{-5}$ nm from the previous studies but re-evaluated coefficient $d_0$ using the Sahara calibration scenes giving $d_0 = 2259.24$ nm.

Finally, we utilise the instrument spectral response function $s$ as determined from pre-flight line source measurements (Schrijver, 2000a, 2001a),

$$s(n, n_0) = \frac{1}{N} \cdot \left( b_0 \cdot \frac{b_1^2}{b_1^2 + (n - n_0)^2} + (1 - b_0) \cdot \frac{b_1^2}{b_1^2 + (n - n_0)^4} \right).$$

with $b_0 = 0.7532$, $b_1 = 0.4313$. $N$ controls the overall normalisation of the response function and $n$ denotes the pixel number, where $n_0$ represents the centre pixel. Substitution of Eq. (12) in Eq. (13) allows us to adjust the full width half maximum (FWHM) of the response function for the Sahara calibration scenes. Here, the FWHM varies between 0.19 and 0.24 nm which correlates with the growth of the detector ice-layer (not shown). However, since the effect on the CO retrieval was minor, we fixed the FWHM to a representative value of 0.21 nm.

4 Validation

To validate our SCIAMACHY CO data product, we have to treat two main problems. First, the retrieved CO column suffers severely from measurement noise. The retrieval noise error for low radiance signal can exceed 100% of the retrieved column. Therefore, any validation can only be performed on quantities averaged in space and time. Second, a direct comparison with ground-based measurements is affected by representation errors. For example, a monthly mean CO concentration derived from ground-based measurements may differ from a corresponding monthly mean of SCIAMACHY measurements due to different temporal sampling. A strict temporal co-registration criterion for both ground-based and SCIAMACHY measurements may reduce the sampling effect but at the cost of fewer SCIAMACHY samplings, which in turn enhances the noise contribution. Both aspects have to be considered in an appropriate validation strategy of the SCIAMACHY CO data product.

4.1 Ground-based Fourier transform spectrometers

In this section, we validate the SCIAMACHY CO data product with the measurements of Fourier transform spectrometers used for observing CO column densities under clear-sky conditions allowing direct Sun measurements. Table 2 summarises the validation data set, which comprises measurements at various stations of the Infrared Working Group (IRWG) that is part of the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndsc.ncep.noaa.gov/) and of the TCCON (Wunch et al., 2010, 2011). The IRWG provides measurements in the mid-infrared with the aim of analysing the atmospheric composition of the troposphere and stratosphere. The NDACC-IRWG supplies CO total columns that we transformed to column mixing ratios by calculating air columns from the surface pressures at a station. At several sites, the data record covers the entire SCIAMACHY lifetime, and thus makes these data very suited for the validation of the SCIAMACHY data product. The TCCON network collects measurements in the same spectral range as recorded by SCIAMACHY from space. This results in a similar vertical sensitivity of both the SCIAMACHY and the TCCON product, which is in particular desirable for validation purposes (see e.g. Wunch et al., 2010, Fig. 3 and Borsdorff et al., 2014, Fig. 2). In 2004, TCCON started with the first instrument at Park Falls, WI, USA, and since then the network has grown gradually to 19 observation sites worldwide. Therefore, the TCCON data set is very well suited to validate SCIAMACHY measurements in the later years of the mission. This study is based on the TCCON GGG2014 data set (Deutscher et al., 2014; Wennberg et al., 2014c, e, a, b, d; Griffith et al., 2014a, b; Strong et al., 2014; Sussmann and Rettinger, 2014; Blumenstock et al., 2014; Kawakami et al., 2014; Sherlock et al., 2014; Warneke et al., 2014; Maziere et al., 2014; Kivi et al., 2014; Morino et al., 2014). Measurements at Ny-Ålesund, Bremen and Four Corners are taken from the GGG2012 data since those sites are not yet available in the 2014 data release.

To achieve the best quality of the SCIAMACHY data, we apply an a posteriori quality filter based on the following criteria:

1. The spectral fit residual $\chi^2$ must be $< 10$.
2. The mean signal-to-noise ratio of the measurements in the fit window must be $> 10$.
3. The noise $\epsilon$ of the retrieved CO, CH$_4$ and H$_2$O column must be below an upper threshold, namely $\epsilon_{CO} < 1 \times 10^{19}$, $\epsilon_{CH_4} < 6 \times 10^{18}$, $\epsilon_{H_2O} < 2 \times 10^{22}$ molec cm$^{-2}$.
4. Only SCIAMACHY measurements are used, which are classified as cloud free by the SPICI algorithm.

Moreover, we selected SCIAMACHY measurements over land, which fall within a radius of 850 km around a TCCON or NDACC-IRWG station site. To derive one representative monthly value for both data sets, we interpolate the FTIR measurements to the point of the SCIAMACHY measurements at time $t$. For this purpose, we consider the ratio of the FTIR columns divided by the co-aligned TM5 columns at two adjacent FTIR samples, $\delta(t_1)$ and $\delta(t_2)$ at time $t_1$ and $t_2$. The temporal interpolated FTIR column $c^{FTIR}(t)$ is then given by

$$c^{FTIR}(t) = \delta(t) \cdot c^{TM5}(t),$$

(14)
Figure 6. The 30-day medians of CO wet air column averaged mixing ratios measured by SCIAMACHY (black) and at various NDACC-IRWG stations (pink). To derive mixing ratios from the CO columns supplied by NDACC-IRWG we calculated total air columns from the surface pressure per station. Open circles denote interpolated values for periods where no NDACC-IRWG measurements are available.
Figure 7. As Fig. 6, but for TCCON measurements indicated in blue. The GGG2014 release of TCCON was used but the stations Ny-Ålesund, Bremen and Four Corners are still based on GGG2012.
where $c^{TM5}(t)$ is the corresponding TM5 CO column and $\delta(t)$ is the linear function through the adjacent points $\delta(t_1)$ and $\delta(t_2)$. Beforehand, we applied an additive bias correction to the TM5 model values such that the overall mean of the FTIR and TM5 values are the same. This simple interpolation scheme makes use of the precise FTIR measurement where the relative temporal trend in CO due to meteorology and photo-chemistry is adopted from the TM5 model. Subsequently, we correct $c^{FTIR}(t)$ for differences between the surface elevation at the station site and the mean altitude of the individual satellite ground pixels using also TM5 CO profiles.

In this manner, we obtain two co-aligned data sets, which are subsequently used to derive monthly median CO column concentrations. The scattering of the individual SCIAMACHY retrievals, which underlies each monthly median, is described by the half difference of the 15.9th and the 84.1th percentile $e_S$ to be an analogue for the standard deviation of a normal distribution. For the same SCIAMACHY retrievals, we also calculate the mean retrieval noise $e_N$. To characterise the retrieval performance per station, we determine the bias $b$ as the mean difference between the monthly median CO concentrations of the ground-based and SCIAMACHY retrievals. A negative sign of a bias means that the SCIAMACHY retrievals are biased low in comparison with the FTIR measurements. Moreover, we use the standard deviations $\sigma$ of these difference and the standard error of the mean $s_e$ to characterise the accuracy of $b$. Finally to characterise the overall performance, we determine the global mean bias $\bar{b}$ as the mean of the individual station biases weighted by their standard error $s_e$ and the corresponding mean standard deviation $\bar{\sigma}$ and the mean standard error $\bar{s_e}$.

For all NDACC and TCCON stations in Table 2, Figs. 6 and 7 show time series of CO monthly median columns, and Fig. 8 summarises the validation diagnostics. Overall, the larger scatter of the individual SCIAMACHY CO columns is mainly caused by the large measurement noise indicated by the similar values of $e_S$ and $e_N$. For some stations, $e_N$ even exceeds a typical mean CO concentration, indicating the need to average data for validation purposes. For the sites Eureka, Ny-Ålesund, Sodankylä, Thule, Kiruna, Harestua, Mauna Loa, Reunion, Tsukuba, Saga and Lauder the noise in
Table 2. Ground-based FTIR stations used for validation. The latitude and longitude values are given in degrees, the surface elevation in km, and the overlap in time of the stations (NDACC-IRWG/TCCON) with SCIAMACHY in years.

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Lat</th>
<th>Long</th>
<th>Alt</th>
<th>NDACC, TCCON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Eureka</td>
<td>80.05</td>
<td>-86.42</td>
<td>0.61</td>
<td>2006–2011, 2010–2011</td>
</tr>
<tr>
<td>2</td>
<td>Ny-Ålesund</td>
<td>78.92</td>
<td>11.92</td>
<td>0.02</td>
<td>2003–2011, 2005–2011</td>
</tr>
<tr>
<td>3</td>
<td>Thule</td>
<td>76.52</td>
<td>-68.77</td>
<td>0.22</td>
<td>2003–2011</td>
</tr>
<tr>
<td>4</td>
<td>Kiruna</td>
<td>67.84</td>
<td>20.40</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Sodankylä</td>
<td>67.37</td>
<td>26.63</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Harestua</td>
<td>60.20</td>
<td>10.80</td>
<td>0.60</td>
<td>2003–2012, 2005–2012</td>
</tr>
<tr>
<td>7</td>
<td>Bialystok</td>
<td>53.23</td>
<td>23.03</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Bremen</td>
<td>53.10</td>
<td>8.85</td>
<td>0.03</td>
<td>2003–2012, 2005–2012</td>
</tr>
<tr>
<td>9</td>
<td>Orleans</td>
<td>47.97</td>
<td>2.11</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Garmisch</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td>2004–2010, 2007–2012</td>
</tr>
<tr>
<td>11</td>
<td>Zugspitze</td>
<td>47.42</td>
<td>10.98</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Jungfraujoch</td>
<td>47.42</td>
<td>10.98</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Park Falls</td>
<td>47.42</td>
<td>10.98</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Toronto</td>
<td>47.42</td>
<td>10.98</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Four Corners</td>
<td>36.80</td>
<td>-90.27</td>
<td>0.44</td>
<td>2004–2012, 2005–2012</td>
</tr>
<tr>
<td>16</td>
<td>Lauder</td>
<td>36.60</td>
<td>-97.49</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>JPL</td>
<td>36.05</td>
<td>140.12</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Saga</td>
<td>33.24</td>
<td>130.29</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Garmisch</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Zugspitze</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Jungfraujoch</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Park Falls</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Toronto</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Reunion</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Wollongong</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Lauder</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Arrival heights</td>
<td>-77.82</td>
<td>166.65</td>
<td>0.20</td>
<td>2003–2012</td>
</tr>
</tbody>
</table>

The data is so large that monthly median values are still dominated by measurement noise. For the remaining stations, the scatter of the monthly median is reasonable, and for stations with a mean instrumental noise error $\bar{\epsilon}_N < 60$ ppb, the seasonal CO cycle becomes clearly visible in the SCIAMACHY time series. The high noise variability can be explained by a corresponding change of the mean signal strength because of varying surface albedo and solar zenith angle, both governing the amount of solar light reflected at the Earth surface.

Overall, Fig. 8 shows a good agreement between SCIAMACHY and TCCON ground-based measurements with a global bias of $-1.2 \pm 7.2$ ppb. For some stations, we observe higher bias, e.g. at Reunion $b = 39$ ppb. These biases come along with large standard error due to a small number of measurements indicating a large uncertainty of $b$ (see Fig. 7). For the NDACC-IRWG sites, we find a negative global bias $b = -9.2 \pm 8.1$ ppb of the SCIAMACHY CO retrieval with respect to the NDACC-IRWG observations. Here, biases for mountain stations like Zugspitze, Jungfraujoch and Izana differ significantly from those at other sites. For the mountain sites, our correction for altitude differences between validation site and the SCIAMACHY ground pixel exceeds 50% of the CO column and therefore our validation is dominated by uncertainties of the TM5 model. The different global biases for TCCON and NDACC-IRWG measurements can be partly explained by the different temporal sampling of the validation sets combined with a small bias trend of $\bar{\epsilon} = 1.47 \pm 0.25$ ppb yr$^{-1}$ in the SCIAMACHY CO columns, which we consider to be significant, keeping in mind both noise and a priori errors in the vertical shape of the reference profile scaled by the inversion (i.e. the null-space error) and the length of the time series. Figure 9 resolves this bias trend for seven NDACC-IRWG and TCCON stations, which cover the full SCIAMACHY mission period combined with low retrieval noise. The average bias trend $\bar{\epsilon}$ is calculated by an average of the individual bias trends weighted by their uncertainty, where we excluded measurements at Toronto because of a discontinuity of the NDACC-IRWG time series (see Fig. 6). This issue is already under investigation and does appear to be instrumental. When correcting the SCIAMACHY data for this bias trend, the bias with NDACC-IRWG becomes $-8.4$ ppb and with TCCON $-4.6$ ppb. Hence, a difference of 3.8 ppb remains between the TCCON and NDACC-IRWG validation. We consider this difference to be significant due to the small mean standard error $\bar{\epsilon}_N$, and we conclude that it is most probably caused by
Furthermore, the distance covered by the MOZAIC/IAGOS ever, biases due to highly polluted airports are still possible. When calculating the mean over longer time periods, how- ascent flight paths will average out for CO total columns (2014) showed that the errors caused by aircraft descent and"nother gases by several long-distance passenger airliners were per- formed during ascent and descent phases (in total more than 40,000 flights). Aircraft descents and ascents are not strictly vertical profiles: the CO concentration at the surface is rep- resentative for the airports but the top of the height cruise can be far away from the take-off location. The represent- ation error of the derived CO column for individual flight profile paths can reach up to 100 %, related to real spatio- temporal variability in CO total columns. However, Nédélec et al. (2003) indicated that the aircraft CO profile measure- ments reach a precision of about ±5 % and de Laat et al. (2014) showed that the errors caused by aircraft descent and ascent flight paths will average out for CO total columns when calculating the mean over longer time periods. How- ever, biases due to highly polluted airports are still possible. Furthermore, the distance covered by the MOZAIC/IAGOS profiles used in this study is about 200–400 km and therefore is within the collocation area around the airports considered for the comparison with SCIAMACHY.

Table 3 summarises the validation data set, which comprises CO profile measurement at 26 airports worldwide. At many airports the data set covers the early years of the SCIA- MACHY mission and therefore forms a complement to the TCCON data set used in the previous section. More information about the MOZAIC/IAGOS programme and its data products is provided by Marenco et al. (1998) and Nédélec et al. (2015) and can be found at http://www.iagos.org/.

For the comparison with the SCIAMACHY CO retrieval, we only select MOZAIC/IAGOS profiles that reach at least 300 hPa and have measurements in every 100 hPa altitude bin. Above the maximum flight altitude, the profiles are extended using the Monitoring Atmospheric Composition and Climate (MACC) reanalysis data at 12:00 Coordinated Universal Time (UTC). MACC is pre-operational Copernicus Atmosphere Service, which provides data records of CO and other atmospheric trace gases (ozone, nitrogen oxides) as well as aerosols and covers the 10 years from 2003 to 2012 (Inness et al., 2013, 2015). The derived CO profiles are vertically integrated to obtain an estimate of the CO total columns. For comparison, individual SCIAMACHY retrievals are quality filtered a posteriori as described in Sect. 4.1. Because the MOZAIC/IAGOS data set is tempor- ally more sparse than the ground-based FTIR data set in Sect. 4.1, we apply a slightly different collocation approach, proposed by de Laat et al. (2012). Here, SCIA- MACHY CO columns are spatially averaged within a 8° × 8° (±4°) area surrounding an airport location. Temporal av- erages are calculated around each MOZAIC/IAGOS sam- ple, where the time window of averaging is chosen such that the retrieval noise of the average is equal to or smaller than 10^{17} molecules cm^{-2} (3.7 ppb). This yields a nonuni- form sampling in time with samples of comparable retrieval noise.

We found that the MOZAIC/IAGOS measurements are on average biased low compared to the TM5 model used as pri- ori for the SCIAMACHY CO retrieval (15 % ± 25 %). This bias is a general known chemistry-transport model issue. The difference between the MOZAIC/IAGOS measure- ments and the ECMWF MACC model used for extending the MOZAIC/IAGOS profiles is low (1 % ± 31 %). This is not surprising given that MACC ingests both MOPITT and IASI CO measurements in its data assimilation scheme.

Figure A1 shows time series of collocated SCIAMACHY and MOZAIC/IAGOS CO total columns for 26 airport loca- tions with more than 13 collocations (similar to Figs. 6 and 7). Part of the data scatter is related to the spatio- temporal variability in CO. This affects both data sets differ- ently, where SCIAMACHY samples represent an average for a larger area surrounding the airport, and MOZAIC/IAGOS columns are derived from slant profiles measured during de- scent and ascent of the aircraft over horizontal differences of 200–400 km. The good agreement of both data sets in their seasonal cycle is noticeable for Windhoek airport. Here, CO is subject to a strong seasonal cycle due to biomass burn-
Table 3. As Table 2, but for MOZAIC/IAGOS airports sites.

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Lat</th>
<th>Long</th>
<th>Alt</th>
<th>MOZAIC/IAGOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>London</td>
<td>50.39</td>
<td>0.85</td>
<td>0.03</td>
<td>2006–2009</td>
</tr>
<tr>
<td>2</td>
<td>Frankfurt</td>
<td>50.32</td>
<td>7.71</td>
<td>0.11</td>
<td>2003–2011</td>
</tr>
<tr>
<td>3</td>
<td>Munich</td>
<td>48.85</td>
<td>12.70</td>
<td>0.45</td>
<td>2003–2011</td>
</tr>
<tr>
<td>4</td>
<td>Vienna</td>
<td>48.51</td>
<td>17.36</td>
<td>0.18</td>
<td>2003–2007</td>
</tr>
<tr>
<td>5</td>
<td>Portland</td>
<td>46.54</td>
<td>−122.18</td>
<td>0.02</td>
<td>2003–2009</td>
</tr>
<tr>
<td>6</td>
<td>Montreal</td>
<td>46.06</td>
<td>−72.65</td>
<td>0.04</td>
<td>2003–2006</td>
</tr>
<tr>
<td>7</td>
<td>Toronto</td>
<td>44.46</td>
<td>−78.59</td>
<td>0.17</td>
<td>2003–2011</td>
</tr>
<tr>
<td>8</td>
<td>Boston</td>
<td>43.01</td>
<td>−69.95</td>
<td>0.01</td>
<td>2003–2011</td>
</tr>
<tr>
<td>9</td>
<td>Chicago</td>
<td>42.79</td>
<td>−86.67</td>
<td>0.20</td>
<td>2003–2010</td>
</tr>
<tr>
<td>10</td>
<td>New York</td>
<td>41.23</td>
<td>−72.75</td>
<td>0.00</td>
<td>2003–2006</td>
</tr>
<tr>
<td>11</td>
<td>Beijing</td>
<td>40.38</td>
<td>115.22</td>
<td>0.04</td>
<td>2003–2005</td>
</tr>
<tr>
<td>12</td>
<td>Philadelphia</td>
<td>40.23</td>
<td>−74.23</td>
<td>0.01</td>
<td>2003–2006</td>
</tr>
<tr>
<td>13</td>
<td>Washington</td>
<td>39.94</td>
<td>−76.43</td>
<td>0.10</td>
<td>2004–2012</td>
</tr>
<tr>
<td>14</td>
<td>Tehran</td>
<td>35.98</td>
<td>50.24</td>
<td>1.01</td>
<td>2003–2005</td>
</tr>
<tr>
<td>15</td>
<td>Los Angeles</td>
<td>34.48</td>
<td>−117.47</td>
<td>0.04</td>
<td>2003–2005</td>
</tr>
<tr>
<td>16</td>
<td>Atlanta</td>
<td>34.43</td>
<td>−83.56</td>
<td>0.31</td>
<td>2003–2011</td>
</tr>
<tr>
<td>17</td>
<td>Dallas</td>
<td>33.83</td>
<td>−96.04</td>
<td>0.19</td>
<td>2003–2010</td>
</tr>
<tr>
<td>18</td>
<td>Tel Aviv</td>
<td>32.77</td>
<td>33.81</td>
<td>0.04</td>
<td>2003–2011</td>
</tr>
<tr>
<td>19</td>
<td>Cairo</td>
<td>30.72</td>
<td>31.15</td>
<td>0.12</td>
<td>2003–2011</td>
</tr>
<tr>
<td>20</td>
<td>Houston</td>
<td>30.63</td>
<td>−94.26</td>
<td>0.03</td>
<td>2003–2004</td>
</tr>
<tr>
<td>21</td>
<td>New Delhi</td>
<td>28.94</td>
<td>76.14</td>
<td>0.24</td>
<td>2003–2006</td>
</tr>
<tr>
<td>22</td>
<td>Kuwait City</td>
<td>28.74</td>
<td>48.44</td>
<td>0.06</td>
<td>2003–2004</td>
</tr>
<tr>
<td>23</td>
<td>Abu Dhabi</td>
<td>25.90</td>
<td>54.32</td>
<td>0.03</td>
<td>2003–2005</td>
</tr>
<tr>
<td>24</td>
<td>Hyderabad</td>
<td>17.54</td>
<td>77.40</td>
<td>0.63</td>
<td>2005–2011</td>
</tr>
<tr>
<td>25</td>
<td>Caracas</td>
<td>11.26</td>
<td>−66.37</td>
<td>0.07</td>
<td>2003–2009</td>
</tr>
<tr>
<td>26</td>
<td>Windhoek</td>
<td>−21.43</td>
<td>17.34</td>
<td>1.72</td>
<td>2005–2012</td>
</tr>
</tbody>
</table>

ing and the high surface albedo permits SCIAMACHY CO retrieval with a low instrument noise error. For airports like Beijing and Tehran, we notice high value outliers, where the MOZAIC/IAGOS columns are much larger than those measured by SCIAMACHY. This bias can be attributed to representation errors comparing localised pollution with spatial averages of SCIAMACHY CO observations. Overall, our results are in agreement with the findings of de Laat et al. (2012) analysing the SCIAMACHY CO measurements before 2009. Figure 10 summarises the comparison between SCIAMACHY and MOZAIC/IAGOS. We find a global bias of $-6.4 \pm 5.6$ ppb. The difference between SCIAMACHY and MOZAIC/IAGOS shows a small but significant positive bias trend of $1.2 \pm 0.7$ ppb yr$^{-1}$, which is in agreement with Sect. 4.1. When correcting the SCIAMACHY data for this bias trend, the global bias reduces to 2.5 ppb which is in the range of the MOZAIC/IAGOS CO column uncertainty as reported by Nédélec et al. (2003).

5 Potential data application

One fundamental limitation of the SCIAMACHY CO data product is its large noise contribution. For most applications, individual CO columns must be averaged to reduce the retrieval noise to an acceptable level. The degree of averaging depends on the signal-to-noise ratio of the corresponding SCIAMACHY observations and therefore on the brightness of the observed scenes. For example, for regions in Africa and Australia with high surface albedo, the retrieval noise is much lower than over dark scenes at high northern latitudes with low solar zenith angle. To mitigate this effect, one can consider data which are averaged both spatially and temporally. Averaging over the full-mission period (January 2003–April 2012), we obtain the CO global distribution shown in Fig. 11. This illustrates that a high spatial resolution can be achieved with the SCIAMACHY CO retrievals sacrificing temporal resolution. One of the most striking features of Fig. 11 is the enhanced CO column concentrations over central Africa due to biomass burning. To illustrate the seasonal variation of CO in this region, Fig. 12 shows the 30-day median of the SCIAMACHY CO concentration for northern hemispheric Africa (averaged between 0 and $10^\circ$ latitude) and southern hemispheric Africa (averaged between 0 and $-35^\circ$ latitude). The figure also includes corresponding averages of TM5 model simulations which use the Global Fire Emissions Database (GFED) version 3 for the biomass burning input. Overall, the SCIAMACHY and TM5 fields agree well. The seasonal variation is present in both data sets including the phase shift between the northern and southern hemispheric CO concentration. In this case, the seasonal
Our recommendation for the use of the data set is to average measurements but its limitation is its high retrieval noise. Hence, lies in the availability of more than 9 years continuous measurements and the collocated and 30-day averaged SCIAMACHY CO retrievals shown in Figs. 6, 7, and 10. For sites with a low CO retrieval noise error and a sufficient temporal coverage, we found a strong correlation showing that the seasonal variation is in agreement – e.g. for MOZAIC/IAGOS (0.7 for Windhoek and 0.8 for Los Angeles), for NDACC/IRWG (0.6 for Wollongong and 0.7 for Kitt Peak) and for TCCON (0.7 for Darwin and 0.7 for Wollongong).

This example illustrates nicely the limitations but also the strength of the presented SCIAMACHY CO data product. Clearly, the strength of the SCIAMACHY CO data product lies in the availability of more than 9 years continuous measurements but its limitation is its high retrieval noise. Hence, our recommendation for the use of the data set is to average individual CO retrievals reducing the retrieval noise error to an acceptable level. Dependent on the considered application, this averaging can be performed spatially as well as temporally. For example, 30-day means averaged spatially over 850 km are already sufficient for most sites worldwide. This is shown by the NDACC/TCCON validation in Sect. 4, where the SCIAMACHY CO retrieval showed a surprisingly homogeneous performance over the full mission time range. The SCIAMACHY CO data set must be seen as complementary to other measurements – e.g. of MOPITT (Deeter et al., 2003) – that provide a finer spatial and temporal resolution. Together with the future TROPOMI instrument, these missions will provide a unique long-term CO data set with global coverage from 2003 onward. In this context, a satellite inter-comparison of the CO retrieval from SCIAMACHY and MOPITT provides the perspective of an interesting and important follow-up study.

6 Summary and conclusions

We presented a full-mission data set of SCIAMACHY CO vertical column densities for cloud-free scenes over land. The retrieval employs the operational SICOR algorithm of the Sentinel-5 Precursor mission and is based on a profile scaling approach using SCIAMACHY 2.3 μm reflectance measurements. For the first time, a stable CO retrieval approach is presented for the entire mission period (January 2003–April 2012), which has to deal with the severe instrument degradation over the nearly 10-year mission period. While previous studies focused on the early years of the SCIAMACHY mission period, we were able to mitigate effects of a changing instrument performance in space on the CO column product. For this purpose, we optimised the retrieval window to account for the serious loss of useful detector pixels caused by radiation damage. Furthermore, we estimated effective instrument parameters, which describe the temporal degradation of SCIAMACHY, using the Sahara region as a natural calibration target. These parameters describe the spectral calibration, a spectral radiometric offset, and the width of the instrument spectral response function. The CO total column amount is inferred simultaneously with methane and water vapour abundances and a Lambertian surface albedo from individual SCIAMACHY measurements assuming a non-scattering model atmosphere.

To obtain atmospheric CO abundances, the retrieval scales a CO reference profile, which represents a specific regularisation of the inversion. Consequently when interpreting the retrieved CO column as an estimate of the true column abundance, the data product suffers from a null-space error which describes the error in the inferred trace gas column due to the assumed profile to be scaled. Using 533 HIPPO CO profile measurements, we showed that for clear-sky conditions the null-space error is typically < ±1 ppb. This represents a minor error source and thus is not further considered in...
Figure 11. CO wet air column averaged mixing ratios over land and clear-sky scenes. The values are averaged from January 2003 to April 2012 on a cylindrical equal area projection with 180 grid points in latitude and 360 in longitude.

Figure 12. Time series of SCIAMACHY CO wet air column averaged mixing ratios (blue) over northern and southern hemispheric Africa compared with the calculation of the TM5 model (pink); 30-day medians are shown.

the validation of our data product. To ensure clear-sky conditions, SCIAMACHY observations are filtered strictly employing the onboard polarisation measurement device of the same instrument (SPICI algorithm).

The full-mission data set is validated with ground-based FTIR measurements at 27 stations of the NDACC-IRWG and TCCON network and MOZAIC/IAGOS airborne measurements at 26 airports worldwide. Here, measurements of the NDACC-IRWG network cover the entire mission period. TCCON measurements can only be used to validate the CO product in the later phase of the mission, whereas IAGOS/MOZAIC measurements are mainly available for the early years of the mission. For the validation, it is important to realise the main and principal limitation of the SCIAMACHY CO product, which is its high retrieval noise of individual CO columns. It varies between 30 ppb over high albedo scenes and more than 170 ppb over dark ground scenes with low signal-to-noise measurements. Con-
sequently, averaging of individual data points is essential for practical data usage. Hence, we base our validation on monthly median column abundances for the comparison with the FTIR measurement and instrument error weighted means for the comparison with MOZAIC/IAGOS airborne observations. Overall, we found a good agreement with TCCON measurements with only a global mean bias of $\bar{b} = -1.2$ ppb with a station-to-station bias variation of $\sigma = 7.2$ ppb. The negative sign of the bias means that SCIAMACHY CO is biased low in comparison with TCCON. For the NDACC-IRWG network, we obtained a significant mean station bias $\bar{b} = -9.2$ ppb with $\sigma = 8.1$ ppb. Moreover, for the IAGOS/MOZAIC measurements, we find a mean station bias of $\bar{b} = -6.4$ ppb with $\sigma = 5.6$ ppb. We detected a small but significant bias trend of about $1.47 \pm 0.25$ ppb yr$^{-1}$ in the SCIAMACHY data. Correcting this bias trend, the bias with the IAGOS/MOZAIC measurements becomes 2.5 ppb, which is within the uncertainty of the IAGOS/MOZAIC measurements. The bias between SCIAMACHY and NDACC-IRWG measurements becomes $-8.4$ ppb and with the TCCON measurements $-4.6$ ppb. A discrepancy of 3.8 ppb remains between the global biases with NDACC-IRWG and TCCON, which is confirmed by directly comparing NDACC-IRWG and TCCON measurements. There are some possible reasons why the NDACC-IRWG and TCCON retrievals differ in that magnitude. NDACC-IRWG retrievals are done from the 5 µm and TCCON from the same 2.3 µm spectral regions as SCIAMACHY using different retrieval approaches. A disagreement of the line parameters of this region can easily lead to differences and is under investigation. Further, the retrieval of the two networks are based on different isotopic lines. NDACC-IRWG is using two $^{13}$CO and one $^{12}$CO line while TCCON retrievals are solely based on $^{12}$CO lines. Furthermore, TCCON retrievals are calibrated by scaling the retrieved CO columns to the ones obtained from simultaneous in situ measurements (aircraft sampling or AirCore measurements) which is not done with the NDACC-IRWG data. Both ground-based FTIR data sets are very valuable for satellite validation, although for the validation of future satellite missions, like the Sentinel 5 Precursor (S5-P) mission to be launched in 2016, it is desirable to improve the comparability of NDACC-IRWG and TCCON measurements.

Finally, to demonstrate potential data use, we showed the seasonal cycle of biomass burning events in central Africa. Averaging the entire mission data set, the biomass burning area can be detected with good spatial resolution. On the other hand, the monthly median SCIAMACHY CO fields averaged over the northern and southern parts of central Africa, reflect the spatial and temporal variability of biomass burning events in this region, in good agreement with the global chemical transport model TM5.

This study represents the first application of the retrieval algorithm SICOR, which was developed for the operational data processing of the S5-P mission on real measurements of the shortwave infrared spectral range. Using the same retrieval approach for both satellite instruments will make the CO data sets of both missions more compatible, which is highly desirable from the perspective of long-term atmospheric monitoring. In a follow-up study, we will focus on extending the presented CO data set to SCIAMACHY ocean measurements.
Appendix A

Figure A1. As Fig. 6, but for MOZAIC/IAGOS measurements indicated in yellow.
Acknowledgements. SCIAMACHY is a joint project of the German Space Agency DLR and the Dutch Space Agency NSO with contribution of the Belgian Space Agency. The work performed is (partly) financed by NSO and through the SCIAMACHY Quality Working Group (SQWG) by ESA. We acknowledge the European Commission for the support to the MOZAIC project (1994–2003) and the preparatory phase of IAGOS (2005–2012), the partner institutions of the IAGOS Research Infrastructure (FZJ, DLR, MPI, KIT in Germany, CNRS, CNES, Métope-France in France and the University of Manchester in UK), ETHER (CNES-CNRS/INSU) for hosting the database, the participating airlines (Lufthansa, Air France, Austrian, China Airlines, Iberia, Cathay Pacific) for the transport free of charge of the MOZAIC/IAGOS instrumentation. MACC data were obtained from http://www.copernicus-atmosphere.eu. We acknowledge the NDACC-IRWG and TCCON ground-based FTIR networks for providing data.

Edited by: H. Worden

References


Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., Dowell, P., Teske, K., Martin, C., and Martin, J.: TCCON Data from Lamont, Oklahoma, USA, Release GGG2014R0, TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center,


