Quality assessment of ozone total column amounts as monitored by ground-based solar absorption spectrometry in the near infrared (> 3000 cm\(^{-1}\))

O. E. García\(^1\), M. Schneider\(^{1,2}\), F. Hase\(^2\), T. Blumenstock\(^2\), E. Sepúlveda\(^{1,3}\), and Y. González\(^1\)

\(^1\)Izaña Atmospheric Research Centre (IARC), Agencia Estatal de Meteorología (AEMET), Santa Cruz de Tenerife, Spain
\(^2\)Institute for Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
\(^3\)University of La Laguna, San Cristóbal de La Laguna, Spain

Received: 5 February 2014 – Accepted: 17 February 2014 – Published: 3 March 2014

Correspondence to: O. E. García (ogarcia@aemet.es)

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

This study examines the possibility of ground-based remote sensing ozone total column amounts (OTC) from spectral signatures at 3040 and 4030 cm\(^{-1}\). These spectral regions are routinely measured by the NDACC (Network for the Detection of Atmospheric Composition Change) ground-based FTIR (Fourier Transform InfraRed) experiments. In addition, they are potentially detectable by the TCCON (Total Carbon Column Observing Network) FTIR instruments. The ozone retrieval strategy presented here estimates the OTC from NDACC FTIR high resolution spectra with a theoretical precision of about 2\% and 5\% in the 3040 cm\(^{-1}\) and 4030 cm\(^{-1}\) regions, respectively. Empirically, these OTC products are validated by inter-comparison to FTIR OTC reference retrievals in the 1000 cm\(^{-1}\) spectral region (standard reference for NDACC ozone products), using a 8 year FTIR time series (2005–2012) taken at the subtropical ozone super-site of the Izaña Observatory (Tenerife, Spain). Associated with the weaker ozone signatures at the higher wavenumber regions, the 3040 cm\(^{-1}\) and 4030 cm\(^{-1}\) retrievals show lower vertical sensitivity than the 1000 cm\(^{-1}\) retrievals. Nevertheless, we observe that the rather consistent variations are detected: the variances of the 3040 cm\(^{-1}\) and the 4030 cm\(^{-1}\) retrievals agree within 90\% and 75\%, respectively, with the variance of the 1000 cm\(^{-1}\) standard retrieval. Furthermore, all three retrievals show very similar annual cycles. However, we observe a large systematic difference of about 7\% between the OTC obtained at 1000 cm\(^{-1}\) and 3040 cm\(^{-1}\), indicating a significant inconsistency between the spectroscopic ozone parameters (HITRAN 2012) of both regions. Between the 1000 cm\(^{-1}\) and the 4030 cm\(^{-1}\) retrieval the systematic difference is only 2–3\%. Finally, the long-term stability of the OTC retrievals has also been examined, observing that both near infrared retrievals can monitor the long-term OTC evolution in consistency to the 1000 cm\(^{-1}\) reference data.
1 Introduction

Atmospheric ozone concentrations (O$_3$) are monitored by ground- and space-based remote sensors, applying different measurement techniques. Very valuable is the ground-based FTIR (Fourier Transform InfraRed) experiment, since it can observe total column amounts and mixing ratio profiles often with high precision. Within the NDACC (Network for Detection of Atmospheric Composition Change, www.acd.ucar.edu/irwg) such FTIR experiments are operated at about 25 globally distributed sites. For NDACC FTIR ozone observations, the wide spectral region between 1000 and 1005 cm$^{-1}$ has been established as the reference spectral region. It theoretically offers the largest sensitivity and the smallest errors for retrieving atmospheric ozone (Barret et al., 2002; Lindenmaier et al., 2010). Furthermore, the high quality of the ozone products obtained in this region (total column amounts and vertical profiles) have been extensively documented by inter-comparing to other ozone measurement techniques (e.g., Dobson/Brewer/DOAS spectrometers and ozone sondes) (Barret et al., 2002; Schneider et al., 2008a, b; Vigouroux et al., 2008; Lindenmaier et al., 2010; García et al., 2012).

In this paper we examine the quality of ozone total column amount (OTC) time series retrieved in the near infrared spectral regions of 3040 cm$^{-1}$ and 4030 cm$^{-1}$. The former has demonstrated to provide the best quality of ground-based infrared ozone retrievals at > 1800 cm$^{-1}$ (e.g., Rinsland et al., 1996; Lindenmaier et al., 2010), and the latter corresponds to the highest infrared frequency with ozone signatures being still strong enough for ground-based retrievals (at higher infrared frequency ozone lines are too weak). This is of growing importance, since the number of FTIR instruments measuring in the near infrared region is steadily increasing. Several FTIR spectrometers within the TCCON (Total Carbon Column Observing Network, TCCON, www.tccon.caltech.edu, Toon et al., 2009) measure near infrared spectra above 3000 cm$^{-1}$, but not mid infrared spectra below 2000 cm$^{-1}$. A contribution of these new instruments to the global ozone dataset would be very desirable, but it is important to precisely examine the quality of
these data and to document its degree of consistency with the standard NDACC FTIR ozone data retrieved at 1000 cm$^{-1}$.

The quality of OTC data obtained from near infrared solar absorption spectra have already been empirically assessed in previous works (e.g., Rinsland et al., 1996; Lindenmaier et al., 2010; Virolainen et al., 2011). However, these studies have been mostly carried out in the context of campaigns of a few days, weeks, or months, thereby constituting no satisfactory long-term assessment. In this context, this study presents a theoretical and empirical long-term quality assessment of OTC obtained during eight years applying two near infrared spectral windows (3041.47–3045.66 and 4026.50–4029.14 cm$^{-1}$). For the empirical validation we use the middle infrared (1000–1005 cm$^{-1}$) retrievals as the reference from the NDACC FTIR solar spectra recorded at the Izaña Atmospheric Observatory (IZO) between 2005–2012. The FTIR Program at IZO is described in Sect. 2 together with the FTIR ozone retrieval strategy. Section 3 presents a theoretical quality assessment for the different ozone retrievals, while the inter-comparison of the analysed infrared spectral regions is shown in Sect. 4 (measurement-to-measurement, annual cycles, and long-term stability). Finally, the main results and conclusions are summarized in Sect. 5.

2 Ground-based FTIR ozone measurements

2.1 FTIR measurements at the Izaña Atmospheric Observatory

The Izaña Atmospheric Observatory (www.izana.org), run by the Spanish Meteorological Agency (AEMET), is a high mountain observatory at the Tenerife island (28.3° N, 16.5° W, and 2373 m.a.s.l.) and offers excellent conditions for atmospheric observations by remote sensing techniques (e.g., Sepúlveda et al., 2012; García et al., 2012).

Izaña’s FTIR activities started in 1999 with a Bruker IFS 120M spectrometer. In 2005 it was replaced by a Bruker IFS 120/5HR spectrometer. These activities have been contributing to the international networks NDACC and TCCON since 1999 and 2007,
respectively. For NDACC, the solar absorption spectra are measured in the middle infrared spectral region (740–4250 cm\(^{-1}\), corresponding to 13.5–2.4 \(\mu m\)), whereby two liquid nitrogen-cooled detectors are applied: a mercury cadmium telluride (MCT) for wavenumbers below 1850 cm\(^{-1}\) and an indium antimonide photodiode (InSb) for higher wavenumbers. The TCCON spectra are recorded in the near infrared spectral region (3500–9000 cm\(^{-1}\), corresponding to 2.9–1.1 \(\mu m\)) using a room-temperature indium gallium arsenide (InGaAs) detector.

In general, the NDACC spectra are highly-resolved with a spectral resolution of 0.005 cm\(^{-1}\), while the resolution of the TCCON spectra is typically limited to 0.02 cm\(^{-1}\). At IZO the average number of FTIR measurement days is about 100 per year (two or three days per week). For this study we only work with the IFS 120/5HR measurements from 2005 onward.

### 2.2 Ozone retrieval strategy

This study examines the OTC retrieved from NDACC FTIR spectra in three different spectral infrared regions: 1000.00–1005.00, 3041.47–3045.66, and 4026.50–4029.14 cm\(^{-1}\), measured with a spectral resolution of 0.005 cm\(^{-1}\) (in the following referred to as 1000 cm\(^{-1}\), 3040 cm\(^{-1}\) and 4030 cm\(^{-1}\), see Table 1). As aforementioned, the 1000 cm\(^{-1}\) region is the official NDACC ozone microwindow, thus, it will be our spectral region of reference. The fitted spectral microwindows containing the O\(_3\) absorption lines are shown in Fig. 1 (lower panel). The upper panel shows the O\(_3\) absorption signatures produced for typical measurement conditions at IZO.

For the different spectral regions we nearly use identical retrieval setups. We use the ground-based FTIR retrieval code PROFFIT (Hase et al., 2004), where the O\(_3\) isotopologues are retrieved on a logarithmic scale using an ad-hoc Tikhonov-Phillips slope constraint (TP1 constraint). In order to minimise interference errors due to water vapour (H\(_2\)O, main interference specie), we apply a two-step strategy. First, we perform a dedicated H\(_2\)O retrieval in order to get an optimal estimation of the atmospheric H\(_2\)O state (Schneider et al., 2006a). Second, we perform the ozone retrieval simultaneously.
with a H$_2$O scaling retrieval that uses the previously estimated H$_2$O state as a priori. This strategy reduces the interfering error due to H$_2$O (see Sect. 3.2) and makes the O$_3$ inversion process more stable. The rest of interfering species are simultaneously fitted with O$_3$ (Table 1).

As a priori profiles of O$_3$ as well as of the all interfering species, we take the climatological data from WACCM (Whole Atmosphere Community Climate Model-version 5, http://waccm.acd.ucar.edu) provided by NCAR (National Center for Atmospheric Research, J. Hannigan, private communication, 2009). The spectroscopic line parameters for O$_3$ are taken from the HITRAN 2012 database (Rothman et al., 2013) and for the rest of interfering species from the HITRAN 2008 database (Rothman et al., 2009), with 2009 update for H$_2$O (www.cfa.harvard.edu/hitran/).

As temperature and pressure profiles, we use the diurnal Vaisala R292 radiosondes (launched about 15 km southeast of the Izaña Observatory on the coastline) and extended them by the NCEP (National Centers for Environmental Prediction) 12:00 UT temperature and pressure profiles.

There are some retrieval settings that are specific for each region. For the 4030 cm$^{-1}$ spectral window methane (CH$_4$) is also an important absorber. In order to reduce its interference with the retrieved O$_3$ amounts we make a profile fit of CH$_4$, thereby an additional CH$_4$ microwindow has been added (see Table 1). For the 3040 cm$^{-1}$ region we enable our retrieval algorithm to disregard the residuals in the 3042.28–3042.48 and 3043.72–3044.04 cm$^{-1}$ ranges (dotted gray lines in Fig. 1). Thereby, we avoid that the relatively high and not well understood residuals observed for these spectral bins significantly affect our retrievals. For the 1000 cm$^{-1}$ region a simultaneous optimal estimation of temperature profile is performed, which assures very precise OTC and O$_3$ profiles (Schneider et al., 2008a; García et al., 2012). This temperature retrieval is not necessary for the O$_3$ retrievals in the 3040 and 4030 cm$^{-1}$ spectral regions, since there the temperature error is much smaller with regard to the total error (see Sect. 3.2). As a priori temperature profiles, we use the aforementioned temperature profiles.
In order to minimise errors due to uncertainties of the ILS (Instrumental Line Shape: the interferometer’s modulation efficiency amplitude and phase error), the ILS is monitored about every two months. These measurements consist in independent low pressure N$_2$O cell measurements and the ILS is retrieved by using the LINEFT code (version 14) as described in (Hase, 2012). Our retrieval works with the so-obtained actual ILS (Fig. 2).

3 Theoretical quality assessment

Trace gas profiles can be retrieved by observing the pressure broadening effect from highly-resolution FTIR solar absorption spectra. The atmospheric solution state $\hat{x}$ can be written down as a linear combination of the a priori state $x_a$ and the real state $x$, the estimated model parameters $\hat{b}$, and the measurement noise $\epsilon$:

$$\hat{x} = x_a + A(x - x_a) + GK_b(b - \hat{b}) + G\epsilon \quad (1)$$

where $G$ represents the gain matrix, $K_b$ a sensitivity matrix to model parameters, and $A$ the full averaging kernel matrix. Equation (1) will be the basis for the analytic error estimation of the retrieval products (for more details see Rodgers, 2000).

3.1 Averaging kernels and sensitivity

The full averaging kernel matrix ($A$) relates the real variability ($x - x_a$) to the measured variability of the considered atmospheric state ($\hat{x} - x_a$), such as $(\hat{x} - x_a) = A(x - x_a)$. This full matrix comprises sub-matrixes describing the smoothing of the target absorber profiles by the remote-sensing system (averaging kernel matrix of the target gas, avks), and the cross-dependence between the target absorbers and the interfering species.
Thus, this matrix can be written as:

$$
A = \begin{pmatrix}
A_{OO} & A_{O/I_1} & \cdots \\
A_{I_1/O} & A_{I_1/I_1} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix}
$$

(2)

where $A_{OO}$ is the averaging kernel matrix of $O_3$, $A_{O/I_1}$ describes the cross-dependence of the retrieved $O_3$ on the interfering specie $I_1$, like $H_2O$, etc.

The response of the $O_3$ retrievals on real atmospheric variability is significatively different for the three analyzed regions. This fact can be observed in the $A_{OO}$ and can be quantified by the trace of $A_{OO}$ (so-called the degrees of freedom for signal, DOFS). The DOFS is a measure for the number of independent $O_3$ partial columns that can be retrieved by the remote sensing system. Thus, we observe a decrease of the FTIR vertical resolution for the high wavenumber regions (see the plots of $A_{OO}$ in Fig. 3 and typical DOFS values as listed in Table 2), associated with the weaker $O_3$ signature in these regions. For the 1000 cm$^{-1}$ region four independent $O_3$ partial columns can be well detected: the troposphere, the tropopause region, the lower/middle stratosphere and the middle/upper stratosphere. For the 3040 cm$^{-1}$ region the number of independent layers is limited to two and the FTIR system only distinguishes the upper troposphere/lower stratosphere and the middle/upper stratosphere. The spectral $O_3$ signatures at 4030 cm$^{-1}$ are only sensitive to atmospheric $O_3$ changes in the stratosphere.

The fact that the mid and near infrared regions do not contain the same amount of information has to be considered when comparing the $O_3$ products obtained in the three spectral regions.

### 3.2 Error estimation

According to Eq. (1) the error in the retrieved profile can be calculated by summing up three error classes: smoothing and interference errors, errors due to uncertainties in the input parameters, and measurement noise. The covariance matrices of these errors are given by:
1. smoothing and interference errors: the covariance of the smoothing error can be calculated by $(A_{OO} - I)S_a(A_{OO} - I)^T$, where $A_{OO}$ is the $O_3$ averaging kernel, $I$ is the Identity matrix and $S_a$ the assumed a priori covariance of atmospheric $O_3$. The $S_a$ matrix used here is obtained from a Electro Chemical Cell (ECC) sonde climatology calculated from weekly measurements above IZO between 1999 and 2006 (Schneider et al., 2008b). For the interference error, the error covariance is $A_{OI}S_aI{A_{OI}}^T$ (Sussmann and Borsdorf, 2007), where $A_{OI}$ describes the cross-dependences of $O_3$ on the interfering species and $S_aI$ is the a priori covariance of the interfering species. In all spectral regions the most important interfering species is $H_2O$. As aforementioned, and in order to minimise its cross-dependence on the retrieved $O_3$, we use an $H_2O$ a priori specific for the considered spectra (the $H_2O$ a priori profile used has been obtained by a previous dedicated $H_2O$ fit on the same spectra). Therefore, as $S_aI$ we can use the $H_2O$ error covariance estimated for the dedicated $H_2O$ fit.

2. Errors due to uncertainties in the input parameters (instrumental characteristics, spectroscopy data, etc): $GK_bS_bG^TK_b^T$, with $S_b$ being the error covariance matrix of $b$. The assumed uncertainties in the input parameters are listed in Table 3. We assume that each error source has a statistical and systematical contribution: 80% and 20% respectively, except for spectroscopic parameters (line strength and pressure broadening coefficient), which are purely systematic (see also García et al., 2012).

3. Retrieval error due to measurement noise: $GS_eG^T$, whereby $S_e$ is the noise covariance matrix.

Table 4 shows our error estimations of OTC for the three analysed spectral regions (the error estimations for the $O_3$ profiles are included in Appendix A). The random errors are dominated by the ILS uncertainty (for the 1000 cm$^{-1}$ and 3040 cm$^{-1}$ retrievals) as well as by the solar lines and the measurement noise (for the 4030 cm$^{-1}$ retrieval). In this study, the measurement noise depends on the quality of the fitted spectra (Hase...
et al., 2004). Thereby, we observe high values of measurement noise error, where the fit residuals are slightly larger, especially in the 4030 cm\(^{-1}\) region (see Fig. 1). Also for this region the solar absorption lines are stronger than the O\(_3\) lines and responsible for a 1.1 \% uncertainty of OTC product. Likewise, only for this region, the H\(_2\)O interfering error is noticeable (about 0.1 \%), but not critical. Note that when H\(_2\)O is simultaneously fitted with O\(_3\), using an optimal H\(_2\)O estimation (O\(_3\) retrieval in one-step strategy), the H\(_2\)O interfering error is larger (about 0.3 \%, in square brackets in Table 4). This fact confirms our decision of using a two-step inversion strategy to estimate the H\(_2\)O profile in a dedicated H\(_2\)O profile fit prior to the O\(_3\) retrieval. Finally, while the contribution of the smoothing error to the total column random error is minor for the 1000 cm\(^{-1}\) region, it is dominating the error budget for the other regions due to the lower FTIR vertical sensitivity.

Considering all the uncertainty sources, the smoothing and H\(_2\)O interference error as well, the total random error (TE) is about 2 \% and 5 \% for the 3040 and 4030 cm\(^{-1}\) retrieval, respectively, while it is only 0.7 \% for the 1000 cm\(^{-1}\) region. Our theoretical quality assessment confirms that the 1000 cm\(^{-1}\) region is the optimal microwindow for retrieving high-quality OTC. It offers the largest sensitivity and the smallest errors whenever the temperature is simultaneously fitted with ozone. If not, the OTC error at 1000 cm\(^{-1}\) region is significantly larger (Schneider and Hase, 2008) and can be similar to the other regions (see errors in brackets in Table 4).

Regarding systematic errors, the spectroscopy is the major contributor. It determines the total systematic error and is about 2–3 \% for the three spectral regions, considering 2 \% as uncertainty in the spectroscopy parameters.

4 Empirical validation

We empirically validate the near infrared retrievals taking the 1000 cm\(^{-1}\) retrievals as the reference from the NDACC FTIR solar spectra time series at IZO (2005–2012). FTIR OTC observations are compared when they are made within 1 h of each other.
The OTC comparison was addressed following two strategies: first, we directly compare the OTC retrieved in each spectral region (unsmoothed OTC), so the influence of the different sensitivities can be directly validated. Second, we convolve the vertically highly-resolved O$_3$ profile obtained from the 1000 cm$^{-1}$ retrievals ($\hat{x}_{1000}$), applying the averaging kernels of the vertically poorly-resolved profiles obtained from the near infrared retrievals ($\hat{x}'_{1000}$, smoothed OTC, Eq. 3). When comparing $\hat{x}'_{1000}$ with the near infrared retrievals the different sensitivities are accounted for.

$$\hat{x}'_{1000} = A(\hat{x}_{1000} - x_a) + x_a$$ (3)

4.1 Measurement-to-measurement comparison

Figure 4 summarizes the comparison between the OTC obtained in the three spectral regions. The straightforward comparison between 1000 cm$^{-1}$ and 3040 cm$^{-1}$ regions (Fig. 4a) shows a good agreement. More than 90 % of the OTC variance obtained for the 1000 cm$^{-1}$ and the 3040 cm$^{-1}$ regions agree (correlation coefficient, $R$, of 0.96, i.e., $R^2 = 0.92$). The agreement to the 4030 cm$^{-1}$ retrieval is slightly poorer ($R = 0.86$, i.e., $R^2 = 0.74$, meaning that about 75 % of the variances are in agreement, Fig. 4b).

The relative differences among regions are not uniformly distributed, but they depend on the season. For example, the relative differences between 1000 cm$^{-1}$ and 3040 cm$^{-1}$ regions show a marked annual cycle: maxima in spring/summer and minima in autumn/winter (see Fig. 5a). This seasonality is due to the different O$_3$ sensitivities of the two retrievals: while the 1000 cm$^{-1}$ retrieval is well able to capture tropospheric and stratospheric O$_3$ variations, the 3040 cm$^{-1}$ retrieval's tropospheric O$_3$ sensitivity is rather limited. Thus, the 1000 cm$^{-1}$ retrieval captures the tropospheric ozone seasonality (maxima in spring/summer and minima in autumn/winter), but the 3040 cm$^{-1}$ does not. For the 4030 cm$^{-1}$ region, we do not observe so clearly a seasonal cycle in the relative difference time series, since it is masked by the high variability of the relative differences, 3.1 % ($1\sigma$, $\sigma$ standard deviation with respect to the 1000 cm$^{-1}$ retrieval).
If we account for the different vertical resolutions and sensitivities, i.e. if we compare to the smoothed 1000 cm\(^{-1}\) data (\(\hat{x}_{1000}',\) Eq. 3), we can document that the retrievals in the 3040 cm\(^{-1}\) and the 1000 cm\(^{-1}\) regions reflect almost the same variation in OTC (Fig. 4c). Note, also, that the straightforward comparison improves (the slope is closer to one and the bias decreases). The scatter observed between the smoothed 1000 cm\(^{-1}\) and 3040 cm\(^{-1}\) retrievals is about 1\% (1\(\sigma\) of the relative differences) and both retrievals observe similar seasonality (Fig. 5b, the peak-to-peak amplitude of the relative differences is reduced from 3\% to 1\%). Part of this remaining scatter may be due to instrumental error sources that are common for both spectral regions (e.g., the ILS error may be correlated among regions), leading to the scatter found may slightly be lower than the real one. For the 4030 cm\(^{-1}\) region, the dispersion reaches 2\% (1\(\sigma\) with respect to the smoothed 1000 cm\(^{-1}\) retrieval). These scatter values agree well with our theoretical error estimation and with previous studies (e.g., Rinsland et al., 1996; Lindenmaier et al., 2010). For example, Lindenmaier et al. (2010) found the same range of uncertainty between OTC retrievals in 1000 and 3040 cm\(^{-1}\) regions (scatter between 1–2\%) using FTIR spectra measured at the Eureka arctic site (80.0\(^{\circ}\)N, 86.4\(^{\circ}\)W).

However, we observe significant systematic differences, especially between the 1000 cm\(^{-1}\) and the 3040 cm\(^{-1}\) retrievals. The 3040 cm\(^{-1}\) region systematically gives OTC 7\% lower than the 1000 cm\(^{-1}\) region. These systematic differences might indicate discrepancies in the applied near and middle infrared spectroscopic parameters. In fact, when using the spectroscopic line parameters given by the HITRAN 2004 database (Rothman et al., 2005) the systematic difference between the OTC obtained at 3040 cm\(^{-1}\) with respect to the OTC reference data is reduced to about +2\%, which agrees with our theoretical error estimation (Table 4). For the 4030 cm\(^{-1}\) region the systematic underestimation of the 1000 cm\(^{-1}\) OTC is within the expected uncertainty (about 2–3\%) and no significant differences are observed between the two HITRAN spectroscopic databases (2004 and 2012).

Since 2007 the 4030 cm\(^{-1}\) region is also measured at IZO in the framework of TCCON. A detailed comparison between the 1000 cm\(^{-1}\) (NDACC spectra) and the
4030 cm\(^{-1}\) retrievals (TCCON spectra) is included in Appendix B. The TCCON’s objective is the monitoring of tropospheric greenhouse gases at a very high precision. Despite the fact that TCCON is not meant to measure stratospheric trace gases, we found a reasonable agreement between the OTC obtained from TCCON spectra in the 4030 cm\(^{-1}\) region and from NDACC spectra in the 1000 cm\(^{-1}\) region \((R = 0.69, i.e., R^2 = 0.48, \text{meaning that about 50\% of the variances agree})\). The scatter is about 4\% (1\(\sigma\) of the relative differences between TCCON and NDACC OTC retrievals). To obtain the TCCON OTC retrievals the ozone retrieval strategy was slightly modified, since the resolution of the TCCON spectra (0.02 cm\(^{-1}\)) is too low to perform an accurate O\(_3\) profile retrieval. Instead, we scale a O\(_3\) profile from the WACCM climatology.

### 4.2 Annual cycle

The OTC annual cycle at subtropical latitudes is mainly controlled by the joint effect of the annual shift of the tropopause’s altitude and the annual cycle of the O\(_3\) photochemical production, as a result of tropical insolation. These phenomena produce a marked OTC annual cycle at subtropical latitudes: peak values in spring and minimum in autumn-winter, as observed in Fig. 6. This figure displays the annual cycle of the OTC multi-year variability (multi-year monthly mean minus multi-year annual mean) calculated for the 2005–2012 period.

The agreement between the OTC annual cycles from the different infrared spectral regions is rather satisfactory and all regions show coherent results. The peak-to-peak amplitude of the OTC annual cycle is similarly captured by the three retrievals (Fig. 6a and b) and the OTC variabilities are perfectly correlated (Fig. 6c and d), with correlation coefficients higher than 95\%. The largest discrepancies occur between 1000 cm\(^{-1}\) and 4030 cm\(^{-1}\) data: the peak-to-peak amplitude calculated from 4030 cm\(^{-1}\) region is about 10 DU lower than the one retrieved by the unsmoothed 1000 cm\(^{-1}\) region, 38 DU. A large part of this difference is due to the poor sensitivity of the 4030 cm\(^{-1}\) retrieval. The difference in the peak-to-peak amplitude is reduced
to less than 1 DU when the smoothed 1000 cm\(^{-1}\) O\(_3\) profiles are considered. For the 3040 cm\(^{-1}\) region, the peak-to-peak amplitude is about 32 and 37 DU for unsmoothed and smoothed 1000 cm\(^{-1}\) O\(_3\) profiles, respectively. The largest differences are observed during summer and the late autumn-winter, due to the missing tropospheric sensitivity of the 3040 cm\(^{-1}\) retrieval.

4.3 Long-term stability

In this section the long-term stability of the near infrared retrievals is checked. For this purpose, we examine possible drifts and discontinuities/change-points in the 3040 cm\(^{-1}\) and the 4030 cm\(^{-1}\) retrievals. We defined a drift as the linear trend of the de-seasonalised monthly mean differences with respect to the 1000 cm\(^{-1}\) OTC data (reference data). The change-points (changes in the median of the difference time series) are analysed by using a robust rank order change-point test (Lanzante, 1996). The Lanzante’s procedure is an iterative method that applies a (single) change-point test, based on summing the ranks of the data from the beginning to each point in the series, and followed by an adjustment step (the median computed for the segments enclosed by the change-points identified is used to adjust the series). In the subsequent iteration the change-point test is applied to the adjusted series and the iterative process finishes when the significance of each new change-point is less than an a priori specified level. Since this test uses non-parametric, resistant and robust principles, it is likely to be highly resilient in the presence of outliers and gaps in the time series (Lanzante, 1996).

A systematic change-point was detected in December 2009 in the monthly time series of the relative differences among regions at 99\% confidence level (both for unsmoothed and smoothed 1000 cm\(^{-1}\) OTC time series, Fig. 7a and b). This discontinuity coincides with the discontinuity detected in the de-seasonalised monthly DOFS time series (especially for the 1000 cm\(^{-1}\) retrievals, Fig. 7c and d) and in the signal-to-noise ratio of the measured spectra time series (figure not shown). This is likely due to the
increase of the noise level since the end of 2009, when we made modifications on the FTIR instrument (failure of the interferometer’s scanner motor and its subsequent replacement). Nevertheless, we observe that there are no significant drifts in the relative differences time series (Table 5) before and after this systematic change point at 95% of confidence level. This fact is also confirmed in the DOFS time series (data not shown). Therefore, we can assume, first, that the different FTIR’s sensitivities do not seem to affect the multi-annual OTC evolution, and secondly, that the near infrared retrievals are consistent over time. Furthermore, these findings show that the analysis of the DOFS (or the instrumental signal-to-noise ratio) time series by homogeneity tests (e.g., Lanzante’s method) offers additional tools for identifying inconsistencies and documenting the stability of long-term series.

5 Summary and conclusions

This paper documents, theoretically and empirically, the quality of the ozone (O₃) total column amounts (OTC) obtained by solar absorption spectrometry in the near infrared spectral regions: 3041.47–3045.66 cm⁻¹ (so-called 3040 cm⁻¹) and 4026.50–4029.14 cm⁻¹ (so-called 4030 cm⁻¹), measured with a spectral resolution of 0.005 cm⁻¹. These retrievals become increasingly important, since the number of high quality ground-based FTIR systems operating in these spectral regions has been continuously increasing during recent years.

The theoretical error assessment showed that uncertainties in the ILS (Instrumental Line Shape) are important error sources. Please recall that our error estimation considers an uncertainty of the modulation efficiency of only 1%. Such low uncertainty can only be achieved if the ILS is carefully and continuously documented. Furthermore, the measurement noise and the limitation of the remote sensing technique in resolving fine vertical O₃ structures (smoothing error) are important error factors for the near infrared retrievals. We estimate an overall precision of about 2% and 5% for the OTC obtained by the 3040 cm⁻¹ and 4030 cm⁻¹ retrievals, respectively. Note that the FTIR’s precision
for the OTC obtained in the 1000 cm\(^{-1}\) region, used as reference in this study, is less than 1% whenever a simultaneous temperature fitting is done. In addition, we estimate that the H\(_2\)O interference error is not critical and is reduced when a dedicated H\(_2\)O fit is performed previous to the ozone retrieval (two-step strategy). This finding demonstrates that the presented near infrared OTC retrievals may work similarly well at more humid FTIR sites.

The theoretical quality assessment has been completed by a detailed empirical inter-comparison study. For this purpose, we use the coincident FTIR measurements in the 1000 cm\(^{-1}\) region as a reference taken at the Izaña Atmospheric Observatory. During the 8 year period between 2005–2012, we observe a good consistency between OTC variations (day-to-day and annual variability) obtained by the 3040 cm\(^{-1}\) and 1000 cm\(^{-1}\) retrievals. The scatter for the relative difference between the two regions is about 1%, when the different sensitives are accounted for, and agree well with our theoretical estimation. However, significant systematic differences (about 7%) were found. It is likely that this discrepancy is caused by inconsistencies between the spectroscopic O\(_3\) parameters at 1000 cm\(^{-1}\) and 3040 cm\(^{-1}\) (HITRAN 2012). For the 4030 cm\(^{-1}\) region the systematic differences are only 2–3%, however, there the scatter is larger (about 2%). This poorer agreement is associated with the weaker O\(_3\) signatures in this region (the O\(_3\) signatures of the 4030 cm\(^{-1}\) region are one order of magnitude smaller than the O\(_3\) signatures of the 1000 cm\(^{-1}\) region).

An important but very difficult scientific task is documenting the long-term evolution of the OTC. The inter-annual OTC changes are rather small and any instrumental drifts/changes might alter the OTC estimated trends. In fact, when analysing the long-term stability of the different OTC retrievals, we find a discontinuity due to modifications on our FTIR instrument. Nevertheless, both near infrared retrievals have shown to be consistent to the 1000 cm\(^{-1}\) retrieval over time, thereby they may be used for correctly estimating the inter-annual OTC changes.

The final conclusion of our study is that solar absorption spectra of the 3040 cm\(^{-1}\) region – if measured at a high spectral resolution (about 0.005 cm\(^{-1}\)) – are well suited...
for monitoring OTC. NDACC FTIR sites routinely measure this spectral region, but currently only a subset of TCCON FTIR operates an InSb detector to cover near infrared spectra > 3000 cm\(^{-1}\). In the light of the above results, adding such detector might be desirable and useful for TCCON sites. Hence, both NDACC and TCCON ground-based FTIR experiments might contribute to global ozone databases.

**Appendix A**

**Error estimation for ozone profiles**

The theoretical quality assessment for the ozone profiles retrieved in the spectral regions of 1000 cm\(^{-1}\), 3040 cm\(^{-1}\), and 4030 cm\(^{-1}\) is shown in Fig. A1, considering the assumed uncertainty sources listed in Table 3 and typical measurement conditions of Izaña spectrometer 120/5HR (25 October 2010). The error profiles are depicted as the root-square of the diagonal elements of the error covariance matrix for the different error sources considered. The smoothing error profile is also shown for each spectral region.

**Appendix B**

**Long-term ozone retrievals in the 4030 cm\(^{-1}\) region from TCCON spectra (spectral resolution of 0.02 cm\(^{-1}\))**

This section presents the same ozone inter-comparison study in the 4030 cm\(^{-1}\) region as shown in the paper, but for measured spectra that are a bit less well resolved (0.02 instead of 0.005 cm\(^{-1}\)). This resolution is typical for TCCON spectra. For these measurement conditions, the theoretical error estimation reveals that the OTC from TCCON spectra in the this spectral region can be estimated with an overall precision of about 7% (Total Error, TE), where about 4% comes from the smoothing error and about 6%...
from all input parameters and measurement noise (assumed uncertainty sources listed in Table 3).

Figure B1 shows the comparison of OTC retrievals and ozone variability annual cycle obtained in the 4030 cm$^{-1}$ region from TCCON spectra and the 1000 cm$^{-1}$ retrievals from NDACC spectra. The scatter between the two data sets is in good agreement with the theoretical error as estimated for the TCCON OTC data. Note that TCCON solar absorption spectra are only measured at Izaña Atmospheric Observatory since 2007.

Appendix C

Multi-annual evolution

The multi-annual evolution of the relative differences and of the ozone total column amounts has been estimated by using a bootstrap re-sampling method (Gardiner et al., 2008; Kohlhepp et al., 2011), which fits the following function to the corresponding time series:

$$F(t) = f_0 + f_{\text{trend}} t + \sum_{i=1}^{p} [a_i \cos(\omega_i t) + b_i \sin(\omega_i t)]$$

(C1)

where $t$ is measured in days, $f_0$ is a baseline constant, and $f_{\text{trend}}$ the linear trend in change per year. The annual cycle is modeled in terms of a Fourier series, where $a_i$ and $b_i$ are the parameters of the Fourier series to be determined and $\omega_i = 2\pi i / T$ with $T = 365.25$ days. We consider frequencies up to $3 \text{yr}^{-1}$ ($p = 3$), since the third order Fourier series provides the best overall results. The significance of linear trends is estimated by assuming that the residuals are Gaussian and uniform over the whole analysed time period (Gardiner et al., 2008).

Acknowledgements. Since 1999 the Izaña FTIR activities have been supported by different funding agencies: European Commission, European Space Agency, Deutsche Forschungsgemeinschaft, Deutsches Zentrum für Luft- und Raumfahrt, and the Ministerios de Ciencia.
e Innovación and Educación from Spain. Furthermore, the research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 284421 (NORS project) and from the Ministerio de Economía and Competitividad from Spain for the project CGL2012-37505 (NOVIA project). M. Schneider and Y. González are supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007–2013)/ERC Grant agreement no. 256961 and E. Sepúlveda enjoys a pre-doctoral fellowship from the Ministerio de Educación from Spain. Finally, we are grateful to the Goddard Space Flight Center for providing the temperature and pressure profiles of the National Centers for Environmental Prediction.

References


Kohlhepp, R., Barthlott, S., Blumenstock, T., Hase, F., Kaiser, I., Raffalski, U., and Ruhnke, R.: Trends of HCl, ClONO$_2$, and HF column abundances from ground-based FTIR measure-
ments in Kiruna (Sweden) in comparison with KASIMA model calculations, Atmos. Chem. Phys., 11, 4669–4677, doi:10.5194/acp-11-4669-2011, 2011. 2088
Lanzante, J. R.: Resistant, robust and non-parametric techniques or the anayliss of climate data: theory and examples, including applications to historical radiosonde station data, Int. J. Climatol., 16, 1197–1226, 1996. 2084
Table 1. Spectral regions and interfering species considered in the FTIR ozone retrievals. MW1 stands for additional microwindows containing well-isolated CO$_2$ lines allowing for a temperature fit: 962.80–963.80, 964.25–965.25, 967.20–968.20, 968.20–969.60 cm$^{-1}$; MW2 stands for an additional microwindow with CH$_4$ lines: 4210.40–4211.70 cm$^{-1}$.

<table>
<thead>
<tr>
<th>MW designation</th>
<th>Spectral region [cm$^{-1}$]</th>
<th>Interfering species</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000 cm$^{-1}$</td>
<td>1000.00–1005.00 + MW1</td>
<td>$^{666}$O$_3$, $^{668}$O$_3$, $^{668}$O$_3$, $^{676}$O$_3$, $^{667}$O$<em>3$, H$</em>{16}$O, CO$_2$, C$_2$H$_4$</td>
</tr>
<tr>
<td>3040 cm$^{-1}$</td>
<td>3041.47–3045.66</td>
<td>$^{666}$O$<em>3$, H$</em>{16}$O, H$<em>{17}$O, H$</em>{18}$O, HDO, CH$_4$, HCl, CO$_2$, solar</td>
</tr>
<tr>
<td>4030 cm$^{-1}$</td>
<td>4026.50–4027.88 + 4028.30–4029.14 + MW2</td>
<td>$^{666}$O$_3$, H$<em>2$O, H$</em>{18}$O, HDO, CO$_2$, CH$_4$, solar</td>
</tr>
</tbody>
</table>
Table 2. Mean ($M$) and standard deviation ($\sigma$) of the DOFS time series of the retrieved OTC of Izaña IFS 120/5HR (2005–2012) for each spectral region.

<table>
<thead>
<tr>
<th>Microwindow</th>
<th>DOFS: $M$, $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000 cm$^{-1}$</td>
<td>4.19, 0.21</td>
</tr>
<tr>
<td>3040 cm$^{-1}$</td>
<td>2.33, 0.15</td>
</tr>
<tr>
<td>4030 cm$^{-1}$</td>
<td>1.05, 0.05</td>
</tr>
</tbody>
</table>
Table 3. Assumed experimental and temperature uncertainties.

<table>
<thead>
<tr>
<th>Error source</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Offset</td>
<td>0.1 %</td>
</tr>
<tr>
<td>Modulation Efficiency</td>
<td>1 %</td>
</tr>
<tr>
<td>Phase Error</td>
<td>0.01 rad</td>
</tr>
<tr>
<td>Line Of Sight (LOS)</td>
<td>0.1°</td>
</tr>
<tr>
<td>Solar lines (intensity and scale)</td>
<td>1 %, $10^{-6}$</td>
</tr>
<tr>
<td>Temperature profile</td>
<td>2 K below 50 km</td>
</tr>
<tr>
<td></td>
<td>5 K above 50 km</td>
</tr>
<tr>
<td>Spectroscopic parameters</td>
<td>2 %</td>
</tr>
</tbody>
</table>
Table 4. Estimated random and systematic errors [%] of OTC for typical measurement conditions of the Izaña spectrometer 120/5HR (25 October 2010) for the three analyzed spectral regions. ILS: joint error due to the modulation efficiency amplitude and phase error uncertainties; TE: Total Error due to input parameters, measurement noise, smoothing error and H$_2$O interfering error. For the latter the error when H$_2$O is simultaneously fitted with O$_3$ (O$_3$ retrieval in one-step strategy) is shown in square brackets. Note that the OTC for the 1000 cm$^{-1}$ region is retrieved considering a simultaneous optimal estimation of the temperature profile (the error when the temperature profile is not fitted with O$_3$ is given in brackets).

<table>
<thead>
<tr>
<th>Error source</th>
<th>1000 cm$^{-1}$</th>
<th>3040 cm$^{-1}$</th>
<th>4030 cm$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random</td>
<td>Systematic</td>
<td>Random</td>
</tr>
<tr>
<td>Baseline Offset</td>
<td>0.3 (0.1)</td>
<td>&lt; 0.1 (&lt; 0.1)</td>
<td>0.4</td>
</tr>
<tr>
<td>ILS</td>
<td>0.5 (0.1)</td>
<td>0.1 (&lt; 0.1)</td>
<td>1.0</td>
</tr>
<tr>
<td>Line Of Sight (LOS)</td>
<td>&lt; 0.1 (&lt; 0.1)</td>
<td>&lt; 0.1 (&lt; 0.1)</td>
<td>0.1</td>
</tr>
<tr>
<td>Solar lines</td>
<td>&lt; 0.1 (&lt; 0.1)</td>
<td>&lt; 0.1 (&lt; 0.1)</td>
<td>0.2</td>
</tr>
<tr>
<td>(intensity and scale)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>0.4 (2.3)</td>
<td>&lt; 0.1 (0.6)</td>
<td>0.3</td>
</tr>
<tr>
<td>Spectroscopy</td>
<td>–</td>
<td>2.0 (2.0)</td>
<td>–</td>
</tr>
<tr>
<td>Measurement Noise</td>
<td>0.2 (&lt;0.1)</td>
<td>–</td>
<td>0.4</td>
</tr>
<tr>
<td>Smoothing Error (SE)</td>
<td>0.1 (0.1)</td>
<td>–</td>
<td>1.7</td>
</tr>
<tr>
<td>H$_2$O Interfering Error</td>
<td>&lt;0.1 [&lt;0.1]</td>
<td>–</td>
<td>&lt;0.1 [&lt;0.1]</td>
</tr>
<tr>
<td>Total Error (TE)</td>
<td>0.7 (2.3)</td>
<td>2.0 (2.1)</td>
<td>2.1</td>
</tr>
</tbody>
</table>
Table 5. Linear trends [%yr⁻¹] of the de-seasonalised monthly relative differences with respect to the 1000 cm⁻¹ OTC data for the IFS 120/5HR time series (2005–2012): 1000 cm⁻¹ unsmoothed (third column), and the 1000 cm⁻¹ smoothed by the avks from 3040 cm⁻¹ and 4030 cm⁻¹ regions (fourth column). The significance interval of the linear trends is estimated by assuming that the residuals are Gaussian and considering 2σ standard deviations (i.e., 95% confidence level).

<table>
<thead>
<tr>
<th>Microwindow</th>
<th>Period</th>
<th>1000 cm⁻¹ Unsmoothed</th>
<th>1000 cm⁻¹ Smoothed</th>
</tr>
</thead>
<tbody>
<tr>
<td>3040 cm⁻¹</td>
<td>2005–2009</td>
<td>-0.05 ± 0.11</td>
<td>-0.03 ± 0.09</td>
</tr>
<tr>
<td></td>
<td>2010–2012</td>
<td>+0.01 ± 0.22</td>
<td>+0.10 ± 0.17</td>
</tr>
<tr>
<td>4030 cm⁻¹</td>
<td>2005–2009</td>
<td>-0.09 ± 0.23</td>
<td>+0.02 ± 0.15</td>
</tr>
<tr>
<td></td>
<td>2010–2012</td>
<td>+0.11 ± 0.55</td>
<td>+0.19 ± 0.29</td>
</tr>
</tbody>
</table>
Fig. 1. Example of simulated transmission spectra of most abundant ozone isotopologue, $^{666}\text{O}_3$ (upper panel, green solid line), and the measured transmission spectra (blue solid line), simulated spectra (red dashed line) and residuals (black solid line) for the 1000, 3040, and 4030 regions and for typical measurement conditions of the Izaña spectrometer 120/5HR (25 October 2010) (bottom panel). The dotted gray lines enclose the spectral regions with less weight in the ozone retrieval. Note that the $^{666}\text{O}_3$ transmission spectra has been multiplied by a factor of 2 and 10 for 3040 and 4030 cm$^{-1}$ regions, respectively.
Fig. 2. Times series of the modulation efficiency amplitude [%] and phase error [rad] at different optical path differences (OPD) for the MCT detector of the Izaña spectrometer 120/5HR. Individual data points indicate individual low pressure N$_2$O cell measurements. The lines are the smoothed efficiency/phase error curves used during the FTIR retrievals. Black at 38 cm, red at 85 cm, green at 133 cm and blue at 180 cm.
Fig. 3. FTIR $O_3$ averaging kernel matrix, avks, for the three analyzed spectral regions, expressed as $\ln[O_3]$, for typical measurement conditions of Izaña spectrometer 120/5HR (25 October 2010). The dotted lines are the diagonals ($x = y$). The panel on the right hand side shows the total sensitivity of FTIR system ($\Sigma_{row}$) and the DOFS (degrees of freedom for signal).
Fig. 4. Scatter plots of OTC [DU] retrieved in the 3040 cm\(^{-1}\) region (a, c, \(N = 2723\)) and in the 4030 cm\(^{-1}\) region (b, d, \(N = 2300\)) vs. 1000 cm\(^{-1}\) region unsmoothed and smoothed by the averaging kernels (avks) from 3040 cm\(^{-1}\) and 4030 cm\(^{-1}\) \(O_3\) retrievals. The black solid lines are the linear regression line of the least square fits, whose parameters are shown in the legend (\(S\) and \(B\) are the slope and the bias of the regression fit respectively, and \(R\) the correlation coefficient). The dotted lines are the diagonals \((x = y)\).
Fig. 5. Consistency of OTC time series at Izaña between 2005 and 2012 using different spectral regions: (a) Relative difference [%] between the 3040 cm\(^{-1}\) and 1000 cm\(^{-1}\) OTC. (b) Same as (a), but after smoothing the 1000 cm\(^{-1}\) OTC data with the 3040 cm\(^{-1}\) averaging kernels (see Eq. 3). (c) Time series of the 1000 cm\(^{-1}\) OTC. (d, e) Same as (a, b), respectively, but for the OTC retrievals in the 4030 cm\(^{-1}\) region. The mean, the standard error of the mean (SEM) and the standard deviation (1\(\sigma\), in brackets) of the relative differences are shown in the legend for each spectral region. The green solid line is the multi-annual evolution fitted according to Eq. (C1) in Appendix C.
Fig. 6. Annual cycle of the OTC variability [DU] for the three spectral regions: (a) 1000 cm$^{-1}$ unsmoothed, 1000 cm$^{-1}$ smoothed by the avks from 3040 cm$^{-1}$ and 3040 cm$^{-1}$, (b) same as (a) but for the 4030 cm$^{-1}$ region. The error bars indicate ±1SEM. Scatter plots of the OTC variability [DU]: (c) 1000 cm$^{-1}$ (unsmoothed and smoothed) vs. 3040 cm$^{-1}$, and (d) same as (c) but for the 4030 cm$^{-1}$ region (SF = Scaling Factor). The dotted lines are the diagonals ($x = y$).
Fig. 7. Time series of the de-seasonalised monthly relative differences with respect to the 1000 cm\(^{-1}\) OTC (a, 3040 cm\(^{-1}\) and b, 4030 cm\(^{-1}\)). The solid and dashed green lines represents the medians of the relative differences time series for the unsmoothed and smoothed 1000 cm\(^{-1}\) OTC, respectively. These medians are calculated for the periods before and after the systematic change point detected in December 2009 (solid orange line). (c) and (d) same as (a) and (b), but for the de-seasonalised monthly DOFS time series. The solid green lines represents the medians of the DOFS time series for the 1000 cm\(^{-1}\) OTC and the dashed ones are for 3040 and 4030 cm\(^{-1}\) regions.
Fig. A1. Random and systematic error profiles [%] for typical measurement conditions of Izaña spectrometer 120/5HR (25 October 2010) for the spectral regions: 1000 cm\(^{-1}\) (a), 3040 cm\(^{-1}\) (b), and 4030 cm\(^{-1}\) (c). ILS means the joint error due to the modulation efficiency amplitude and phase error uncertainties and TPE (Total Parameter Error, black line) is the quadratic sum of all errors except for smoothing error.
Fig. B1. Same as Figs. 4–6, but for OTC obtained from TCCON spectra in the 4030 cm⁻¹ region (InGaAs detector, spectral resolution of 0.02 cm⁻¹, N = 2949).