



# Investigating the long-term evolution of subtropical ozone profiles applying ground-based FTIR spectrometry

O. E. García<sup>1</sup>, M. Schneider<sup>2,1</sup>, A. Redondas<sup>1</sup>, Y. González<sup>1</sup>, F. Hase<sup>2</sup>, T. Blumenstock<sup>2</sup>, and E. Sepúlveda<sup>3,1</sup>

<sup>1</sup>Izaña Atmospheric Research Centre (IARC), Agencia Estatal de Meteorología (AEMET), Santa Cruz de Tenerife, Spain

<sup>2</sup>Institute for Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

<sup>3</sup>University of La Laguna, La Laguna, Spain

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

Received: 9 February 2012 – Published in Atmos. Meas. Tech. Discuss.: 10 May 2012

Revised: 17 October 2012 – Accepted: 4 November 2012 – Published: 28 November 2012

**Abstract.** This study investigates the long-term evolution of subtropical ozone profile time series (1999–2010) obtained from ground-based FTIR (Fourier Transform InfraRed) spectrometry at the Izaña Observatory ozone super-site. Different ozone retrieval strategies are examined, analysing the influence of an additional temperature retrieval and different constraints. The theoretical assessment reveals that the FTIR system is able to resolve four independent ozone layers with a precision of better than 6% in the troposphere and of better than 3% in the lower, middle and upper stratosphere. This total error includes the smoothing error, which dominates the random error budget. Furthermore, we estimate that the measurement noise as well as uncertainties in the applied atmospheric temperature profiles and instrumental line shape are leading error sources. We show that a simultaneous temperature retrieval can significantly reduce the total random errors and that a regular determination of the instrumental line shape is important for producing a consistent long-term dataset. These theoretical precision estimates are empirically confirmed by daily intercomparisons with Electro Chemical Cell (ECC) sonde profiles. In order to empirically document the long-term stability of the FTIR ozone profile data we compare the linear trends and seasonal cycles as obtained from the FTIR and ECC time series. Concerning seasonality, in winter both techniques observe stratospheric ozone profiles that are typical middle latitude profiles (low tropopause, low ozone maximum concentrations) and in summer/autumn profiles that are typical tropical profiles (high tropopause, high maximum concentrations). The linear trends estimated from the FTIR and the ECC datasets agree within their error bars. For the FTIR time series, we observe a significant

negative trend in the upper troposphere/lower stratosphere of about  $-0.2\% \text{ yr}^{-1}$  and a significant positive trend in the middle and upper stratosphere of about  $+0.3\% \text{ yr}^{-1}$  and  $+0.4\% \text{ yr}^{-1}$ , respectively. Identifying such small trends is a difficult task for any measurement technique. In this context, super-sites applying different techniques are very important for the detection of reliable ozone trends.

## 1 Introduction

In the coming decades some kind of ozone recovery is expected, however, it is difficult to predict how, when and to what extent it will occur (Weatherhead and Andersen, 2006). Currently it is discussed how climate change will interact with ozone recovery. The multiple interactions between the components of the chemistry climate system complicate a clean attribution of changes in ozone to changes in ODSs (Ozone-Depleting Substances) and other factors such as the Brewer-Dobson circulation, anthropogenic emissions of greenhouse gases, stratospheric temperatures, etc. (WMO, 2011). For example, climate models predict an accelerated stratospheric circulation, leading to changes in the spatial distribution of stratospheric ozone and an increased stratosphere-to-troposphere ozone flux (Hegglin and Shepherd, 2009, and references therein). Furthermore, rising tropospheric temperatures can increase the water amounts that are injected into the stratosphere, thereby triggering ozone destruction (Anderson et al., 2012). Nonetheless, the effect of the decrease of anthropogenic halogen abundances in the upper stratosphere from the mid-1990s has to be considered

also (WMO, 2011). In order to verify or decline the different climate model simulations consistent long-term observations of the vertical distribution of ozone are required. Since the expected signals are rather small (e.g., expected trends from  $-3\%$  to  $+1\%$  per decade between 1960 and 2100, Hegglin and Shepherd, 2009; Li et al., 2009), only high precision observations are useful.

Within the NDACC (Network for the Detection of Atmospheric Composition Change, e.g., Kurylo and Zander, 2000) high resolution solar absorption infrared spectra have been measured by ground-based FTIR (Fourier Transform InfraRed) spectrometers for up to two decades at globally distributed sites. It has been shown that these measurements can provide very high quality ozone total column amounts (Schneider and Hase, 2008; Schneider et al., 2008a; Viatte et al., 2011) and profiles (Schneider et al., 2008b). Due to its long-term characteristic and its high precision, the FTIR data are very interesting for trend studies. Vigouroux et al. (2008) (updated in WMO, 2011) estimated ozone trends at several European NDACC FTIR sites. In this work, we examine in detail the FTIR error sources and discuss how they can affect the estimated ozone trends. We present three different FTIR ozone profile retrieval setups, including the setup applied by Vigouroux et al. (2008), and discuss their reliability for providing correct ozone trend estimates. This study is performed for the ozone super-site Izaña Observatory, where since 1999 the FTIR measurements have been performed coincidentally to several other high quality atmospheric ozone measurement techniques (e.g., Brewer spectrometer, Electro Chemical Cell, ECC, sondes, photometric in situ surface).

The Izaña Observatory and its Ozone Program is described in Sect. 2. In Sect. 3, we present the three different FTIR retrieval setups, perform detailed theoretical error estimations and discuss the error sources that can affect trend estimations. In Sect. 4, we briefly discuss the quality of the ECC sonde data and in Sect. 5 we show a day-to-day comparison between the three different FTIR datasets and the ECC dataset. In Sect. 6, we present the ozone seasonality and the trends obtained at different altitudes from the different FTIR datasets and discuss their consistency to the values obtained for the ECC dataset. Finally, the main results are summarised in Sect. 7.

## 2 The Izaña Observatory and its ozone program

The Izaña Observatory ( $28.3^\circ$  N,  $16.5^\circ$  W) belongs to the Spanish Meteorological Agency (AEMET). It is a subtropical high mountain observatory, located at 2.37 km altitude on Tenerife Island, and typically above a temperature inversion layer which acts as a natural barrier for local pollution. Hence, it offers excellent conditions for the remote-sensing of the upper atmosphere.

Since many years the Izaña Observatory has been a WMO/GAW (World Meteorological Organisation/Global

Atmospheric Watch) station and an NDACC site, monitoring a large variety of atmospheric constituents, among them, ozone total column amounts and ozone profiles. Different ozone measuring techniques are applied: Brewer spectrometer, FTIR spectrometer, Differential Optical Absorption Spectroscopy (DOAS), in situ ultraviolet photometric analysers, and ECC sondes. In this study, we focus on ozone profiles and in the following two subsections we briefly describe the techniques that can measure ozone concentrations at different altitudes above the Izaña Observatory.

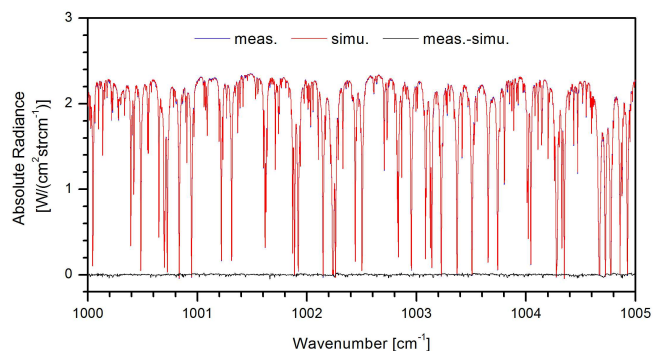
### 2.1 FTIR program

Ground-based FTIR systems measure solar absorption spectra applying a high resolution Fourier Transform spectrometer. The FTIR activities at Izaña started in 1999 when, in the framework of a collaboration between AEMET and KIT (Karlsruhe Institute of Technology, Germany), a Bruker IFS 120M was installed at the Observatory. In January 2005 KIT scientists substituted this spectrometer by a Bruker IFS 120/5HR, which is one of the best performing FTIR spectrometers commercially available. During March and April 2005 both instruments measured side-by-side, which allows for documenting the consistency of both FTIRs.

Since 1999, the Izaña experiment provides data to the NDACC network. Currently there are about 25 ground-based FTIR NDACC experiments. For NDACC, the solar absorption spectra are measured in the mid-infrared spectral region ( $740$  and  $4250\text{ cm}^{-1}$ , corresponding to  $13.5$  and  $2.4\text{ }\mu\text{m}$ ), which is covered by six individual measurements applying different filters in order to achieve an optimal signal-to-noise ratio. In this spectral region, the FTIR spectra are recorded using a potassium bromide (KBr) beamsplitter, whereby two liquid nitrogen-cooled detectors are applied: a mercury cadmium telluride (MCT) for wavenumbers below  $1850\text{ cm}^{-1}$  and an indium antimonide photodiode (InSb) for higher wavenumbers. For operational ozone measurements the FTIR spectra covering the  $1000\text{ cm}^{-1}$  region were measured with an aperture of  $1.5\text{ mm}$ , which corresponds to a field-of-view of only  $0.2^\circ$ . Therefore, the FTIR instrument only analyses sunlight coming from the centre of the solar disc (diameter of  $0.5^\circ$ ). In order to increase the signal-to-noise ratio several scans, with a high spectral resolution of  $0.005\text{ cm}^{-1}$  (maximum Optical Path Difference,  $\text{OPD}_{\text{max}}$  of  $180\text{ cm}$ ), are co-added (8 for the ozone measurements). Thereby, the measurement of one spectrum takes about 10 min. At Izaña the FTIR spectra typically are measured on two or three days per week.

### 2.2 ECC and in situ surface program

The Ozone Sonde Program on Tenerife started in November 1992 and since March 2001 these ECC sonde activities form part of the NDACC. The sondes (type: Scientific Pump 6A) are launched weekly very close to the Izaña



**Fig. 1.** Spectral microwindow applied for ozone retrievals.

Observatory: from Santa Cruz de Tenerife (35 km northeast of the Observatory) and since October 2006 from Güímar (15 km east of the Observatory). Smit et al. (2007) demonstrate that the expected uncertainty of these ECC sonde profiles is  $\pm 5$ –10 %.

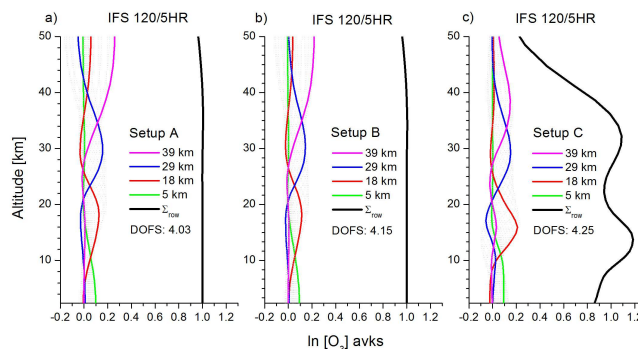
Since 1987, and in the framework of the GAW Program, in situ surface ozone has been monitored. During this period, different ultraviolet absorption instruments have been applied. Since 1999, two TEI analysers (Thermo Electron corporation environmental Instruments) are in operation and continuously record one-minute average ozone values. Zero-checks with an activated-carbon absorber are performed on a daily basis to detect instrumental offset drifts. This Surface Ozone Program has been audited by the World Calibration Center for Surface Ozone, Carbon Monoxide and Methane each two or four years since 1996. The expected uncertainty is to be  $\pm 1$  ppb (Zellweger et al., 2009). At the Izaña high altitude Observatory, the surface measurements are well representative of the free troposphere and, thus, are well suited for validating the tropospheric ozone concentrations observed by the FTIR system (see also Sepúlveda et al., 2012).

### 3 Ground-based FTIR ozone data

#### 3.1 Ozone retrieval strategy

The high resolution spectra allow an observation of the pressure broadening effect and, thus, the retrieval of trace gas profiles. The inversion problems faced in atmospheric remote sensing are in general ill-determined and the solution has to be properly constrained. An extensive treatment of this topic is given in the textbook by C. D. Rodgers (Rodgers, 2000).

We apply the retrieval code PROFFIT (Hase et al., 2004) for retrieving the FTIR ozone profiles and investigate three different retrieval setups (in the following referred as retrieval setups A, B and C, see Table 1). All setups apply the spectral ozone microwindow suggested by Barret et al. (2002) (see Fig. 1) and the different ozone isotopologues ( $^{666}\text{O}_3$ ,  $^{686}\text{O}_3$  and  $^{668}\text{O}_3$ ) are retrieved on a logarithmic scale. Likewise, all setups use the same a priori profile taken from



**Fig. 2.** Columns of the FTIR averaging kernels, avks, for retrieved ozone values using the setup A (a), setup B (b), and C (c), expressed as  $\ln[\text{O}_3]$ , for the Izaña IFS 120/5HR.  $\Sigma_{\text{row}}$  (black line) is the total sensitivity of FTIR system and DOFS are the degrees of freedom for signal.

an ECC sonde climatology calculated from measurements between 1996 and 2006 and approximated to HALOE climatology above 30 km (Schneider et al., 2008b).

Setup A can be considered as the “NDACC” approach except for the logarithmic instead of the linear scale retrieval of ozone. This strategy has already been used for the ozone trend estimations at Izaña and Kiruna as presented in Vigouroux et al. (2008) (updated in WMO, 2011). Setup B is further refined by an additional temperature retrieval, for which we simultaneously fit four  $\text{CO}_2$  microwindows between 962 and 970  $\text{cm}^{-1}$ . Schneider and Hase (2008) demonstrated that such temperature retrieval is very important when aiming for high quality ozone data. For the setups A and B the inversion problem is solved using an ad-hoc Tikhonov-Phillips slope constraint (TP1 constraint). This constrains the vertical profile slope and the absolute value for the uppermost atmospheric model altitude. The strength of the constraint is determined by starting with a weak constraint and then increasing it until we observed a significant increase in the residual of the spectral fit (L-curve criterion). This is different to setup C, for which an Optimal Estimation (OE) constraint instead of the ad-hoc TP1 constraint is used. In this case, the a priori covariance matrix,  $\mathbf{S}_a$ , is taken from an ECC sonde climatology (Schneider et al., 2008b). In addition, setup C includes an inter-species constraint between the different ozone isotopologues ( $^{666}\text{O}_3$ ,  $^{686}\text{O}_3$  and  $^{668}\text{O}_3$ , Schneider et al., 2006). As a priori for the ozone isotopologue ratios, we assume a heavy ozone enrichment of 10 % throughout the atmosphere (Johnson et al., 2000; Mauersberger et al., 2001). Setup C can be considered as the setup with the most realistic constraint, since the actual covariances of ozone and of the different ozone isotopologues are taken into account. Applying a realistic constraint facilitates a correct interpretation of the day-to-day variability in the measured spectra. Assuming that atmospheric trends will be largest for altitudes with largest atmospheric day-to-day

**Table 1.** Description of FTIR ozone retrieval setups. Note that the level of refinement increases from the setup A to the setup C. <sup>1</sup>Four temperature microwindows (MW) with isolated CO<sub>2</sub> signatures [cm<sup>-1</sup>]: 962.80–963.80, 964.25–965.25, 967.20–968.20, 968.60–969.60.

| Retrieval Setup        | A: “NDACC”   | B + Temperature  | C + (TP → OE)  |
|------------------------|--|--|--|
| MW [cm <sup>-1</sup> ] | 1000–1005  | 1000–1005 + Temp. MW <sup>1</sup>  | 1000–1005 + Temp. MW <sup>1</sup>  |
| Interfering species    | <sup>666</sup> O <sub>3</sub> , <sup>686</sup> O <sub>3</sub> , <sup>668</sup> O <sub>3</sub><br>H <sub>2</sub> O, CO <sub>2</sub> , C <sub>2</sub> H <sub>4</sub> | <sup>666</sup> O <sub>3</sub> , <sup>686</sup> O <sub>3</sub> , <sup>668</sup> O <sub>3</sub><br>H <sub>2</sub> O, CO <sub>2</sub> , C <sub>2</sub> H <sub>4</sub> | <sup>666</sup> O <sub>3</sub> , <sup>686</sup> O <sub>3</sub> , <sup>668</sup> O <sub>3</sub><br>H <sub>2</sub> O, CO <sub>2</sub> , C <sub>2</sub> H <sub>4</sub> |
| Logarithmic scale      | Yes  | Yes  | Yes  |
| Temp. retrieval        | No   | Yes  | Yes  |
| Int.-Spec. Const.      | No   | No   | Yes  |
| Inversion method       | Tikhonov-Phillips (TP) with slope constraint (TP1)   | Tikhonov-Phillips (TP) with slope constraint (TP1)   | Optimal Estimation (OE) with S <sub>a</sub> from Tenerife’s ECC sondes   |

variability, this constraint might also be advantageous for trend estimations.

For all setups H<sub>2</sub>O, CO<sub>2</sub> and C<sub>2</sub>H<sub>4</sub> are considered as main interfering absorbers. Spectroscopic line parameters of these interfering absorbers as well as of ozone are taken from the HITRAN 2008 database (Rothman et al., 2009), except for H<sub>2</sub>O, where the parameters from the HITRAN 2009 update are applied. The a priori mean profiles of the interfering absorbers are taken from the WACCM climatology (Whole Atmosphere Community Climate Model, <http://waccm.acd.ucar.edu>) provided by James W. Hanningan (National Center for Atmospheric Research).

It is important to mention that we do not vary any of the a priori trace gas profiles that depend on the season. Thereby, all variability observed in our retrieved ozone profiles comes from the measurements. As a priori for the temperature retrievals, we use the diurnal radiosondes (Vaisala RS92) up to 30 km and extended it by the NCEP (National Centers for Environmental Prediction) 12:00 UT daily temperature profiles. The radiosondes are launched twice per day (23:15 UT and 11:15 UT), just about 15 km southeast of the Izaña Observatory on the coastline.

For all retrievals, we apply the ILS (Instrumental Line Shape: the interferometer’s modulation efficiency and phase error) as derived regularly from low pressure gas cell measurements by means of the code LINEFIT (Hase et al., 1999).

### 3.2 Vertical resolution of FTIR ozone profiles

The vertical structures that are detectable by a ground-based FTIR system are given by the averaging kernel matrix (avks,  $\hat{A}$ ). The columns of this matrix describe how an atmospheric perturbation is smoothed out by the remote-sensing system. As example, Fig. 2 depicts the avks columns for a typical measurement of the IFS 120/5HR and for setups A, B, and C. We can observe that the maxima of these response functions generally peak at the altitude of the perturbations: the green line describes the response for an 1.0 disturbance at 5 km and it peaks close to 5 km, the red line represents the response on a disturbance at 18 km and it peaks close to 18 km, etc. The different widths of the avks and the different

**Table 2.** Mean ( $M$ ) and standard deviation ( $\sigma$ ) of the DOFS time series of the retrieved ozone obtained from the Izaña IFS 120/5HR for all setups. These values are shown for each layer (2.37–13 km, 12–23 km, 22–29 km, 28–42 km) and for the total column (2.37–120 km).

| Layer<br>[km] | Retrieval Setup  |                  |                  |
|---------------|------------------|------------------|------------------|
|               | A<br>$M, \sigma$ | B<br>$M, \sigma$ | C<br>$M, \sigma$ |
| 2.37–13       | 1.01, 0.06       | 1.09, 0.05       | 1.31, 0.08       |
| 12–23         | 1.10, 0.06       | 1.23, 0.08       | 1.49, 0.11       |
| 22–29         | 1.01, 0.06       | 1.06, 0.05       | 1.02, 0.05       |
| 28–42         | 1.25, 0.06       | 1.28, 0.06       | 1.18, 0.06       |
| 2.37–120      | 3.84, 0.23       | 4.10, 0.14       | 4.20, 0.17       |

sensitivities (sum along the row of the avks,  $\Sigma_{\text{row}}$ ) are mainly due to the differences in the applied a priori constraint. For all setups, and for altitudes below 40 km, the sensitivity is better than 80 %. Beyond 40 km the sensitivity significantly decreases for setup C, whereas for the setups A and B it remains very close to 1.0 (which is a typical behaviour for TP1 constraints).

Following Rodgers (2000), the number of independent layers of a retrieved trace gas profile can be quantified by the trace of the averaging kernel matrix, the so-called “number of degrees of freedom for signal” (DOFS). For ozone, a mean total DOFS of around four is observed for all setups (see Table 2), meaning that the FTIR system is able to resolve four independent atmospheric layers: the troposphere (2.37–13 km), the tropopause region (12–23 km), the lower/middle stratosphere around the ozone maximum (22–29 km) and the middle/upper stratosphere (28–42 km). We have highlighted the ozone kernels at the altitudes of 5, 18, 29 and 39 km in representation of these layers (see Fig. 2). Note that for all layers the achieved DOFS are typically larger than one, indicating that the FTIR system is well sensitive for these layers.

The influence of the retrieval settings is also visible in the DOFS: for setup C the DOFS below 29 km is higher if compared to the setups A and B. Up to 29 km the setup C profiles have a better vertical resolution than the setup A



**Fig. 3.** Time series of total DOFS of retrieved ozone values using the setup C for the whole FTIR time series between 1999 and 2010 (number of data,  $N$ , is 1887).

and B profiles. This is a consequence of the rather loose constraint applied for setup C at these altitudes (the constraint of setup C is based on real ozone data, which show high variabilities from 15 to 25 km due to a vertically shifting tropopause). In addition, including a simultaneous temperature fit improves the retrieval quality (i.e., reduces the residues of the fitted spectra), thus, allowing for a more detailed interpretation of the spectra (see also the increased DOFS when including the temperature fit).

When interpreting the FTIR ozone profile time series it is important to consider that trends in the sensitivity of the remote-sensing system (avks or DOFS) can influence the ozone trend: decreasing DOFS might underestimate gradually increasing differences between the a priori ozone amounts used as constraint and the real ozone amounts, i.e., the ozone trends will be underestimated. Furthermore, there might be a bias between the climatologic ozone data (the a priori) and the FTIR ozone data due to systematic error sources like spectroscopic line parameters. In this case, the magnitude of the bias will decrease with decreasing DOFS, thereby leading to a trend even though real atmospheric ozone remains stable. The variability and the drifts in the DOFS can be observed in Fig. 3, which shows the time series of the total DOFS for setup C. The apparent drift between 1999 and 2004 means that trends estimated for this time period have to be treated with care. The DOFS time series also shows some kind of annual cycle, which is strongly anti-correlated with the annual cycle of the ozone slant column amounts (for low slant columns – less saturated ozone lines – the DOFS is higher than for high slant columns – partially saturated ozone lines). Likewise, in 2005 the change from the IFS 120M to the IFS 120/5HR instrument can be observed in the DOFS time series. As expected, we observe that the DOFS obtained from IFS 120M spectra are smaller than the DOFS obtained from the IFS 120/5HR spectra: the total DOFS values differ by 7%. This difference is smaller in the troposphere (6–7%) than in the stratosphere (up to 10%).

**Table 3.** Assumed experimental and temperature uncertainties.

| Error Source                      | Uncertainty                        |
|-----------------------------------|------------------------------------|
| Modulation Efficiency             | 1 %                                |
| Phase Error                       | 0.01 rad                           |
| Baseline Offset                   | 0.1 %                              |
| Temperature profile               | 2 K below 50 km<br>5 K above 50 km |
| Line Of Sight (LOS)               | 0.1°                               |
| Solar lines (intensity and scale) | 1 %, $10^{-6}$                     |
| Spectroscopic parameters          | 1 %                                |

### 3.3 Error estimation

The theoretical error estimation is analytically performed by the retrieval code PROFFIT (Hase et al., 2004). It is based on the formalism suggested by Rodgers (2000). We consider three error types: (a) errors due to uncertainties in the input parameters (instrumental characteristics, spectroscopy data, etc.), (b) the smoothing error, and (c) errors due to measurement noise.

As uncertainties in the input parameters we assume the values as listed in Table 3. The uncertainties are split into statistical and systematic contributions, 80 % and 20 %, respectively, except for spectroscopic parameters (line strength and pressure broadening coefficient), for which the entire uncertainty in the input parameter is systematic. The assumptions of Table 3 are reasonable for the IFS 120/5HR, but are very likely too optimistic for the IFS 120M (see, for instance, the discussion in Sect. 3.4). Hence, Table 3 only describes the errors for the IFS 120/5HR. For the IFS 120M the errors due to LOS (Line Of Sight) and ILS uncertainties are by a factor of 2–3 larger.

The propagation of uncertainty sources for a typical measurement of the IFS 120/5HR, and using the different retrieval setups, is displayed in Fig. 4. The error profiles are shown as the root-square of the diagonal elements of the error covariance matrix for the different error sources considered (see Table 3). The error covariance matrices are calculated following the matrix multiplication formalism as suggested by Rodgers (2000). The smoothing error (SE), associated with the smoothing of the real vertical distribution of ozone by the FTIR measurement process, is the leading error for all setups. It is  $(\hat{\mathbf{A}} - \mathbf{I})\mathbf{S}_a(\hat{\mathbf{A}} - \mathbf{I})^T$ , whereby  $\mathbf{I}$  is a unity matrix,  $\hat{\mathbf{A}}$  is the averaging kernel, and  $\mathbf{S}_a$  the assumed a priori covariance of atmospheric ozone. We use the same  $\mathbf{S}_a$  matrix for the three setups, which is obtained from an ECC sonde climatology (Schneider et al., 2008b). Note that the inverse of this matrix ( $\mathbf{S}_a^{-1}$ ) has been used for the optimal estimation constraint of setup C. The SE reaches about 40 % in the tropopause region, where the ozone concentrations are very variable and the profile might be highly-structured. The FTIR system is not able to resolve such fine vertical structures. Excluding the smoothing error, below 20 km the

random errors are dominated by the measurement noise, the temperature and the ILS uncertainties. Above 20 km and for setup A the error due to temperature uncertainties notably increases, reaching about 5 % at 40 km. In contrast to the setups retrieving the temperature (setups B and C), where the respective errors are lower than 2.5 %.

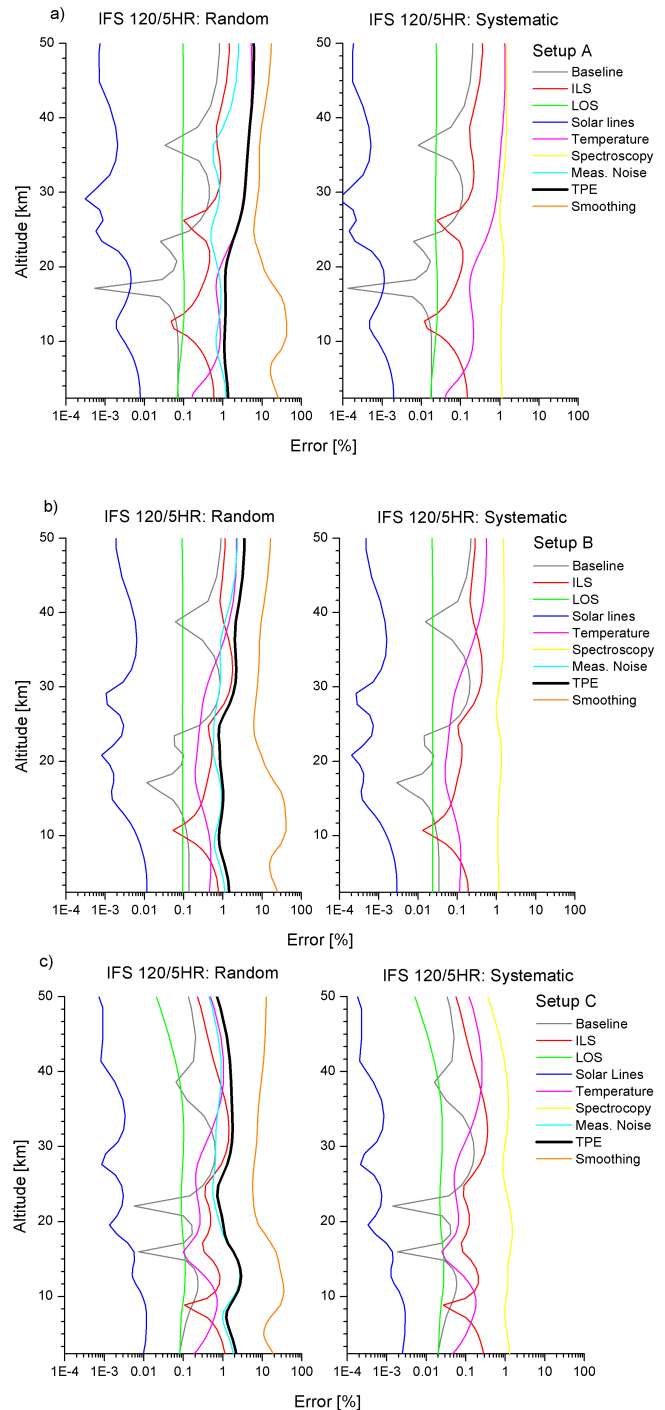
As summary, Table 4 shows the random error budget for the different partial column amounts and for the total column amount. It lists the total random errors (TRE) estimated as the root-square-sum of all parameter errors (TPE, input parameters and measurement noise) and the smoothing error (SE):  $\sigma_{\text{TRE}} = \sqrt{\sigma_{\text{TPE}}^2 + \sigma_{\text{SE}}^2}$ . The significant contributors to the TPE (ILS, temperature and measurement noise) are also shown. Note that the inclusion of a simultaneous temperature fit (setup B and C) reduces significantly the random error associated with the temperature for all layers. For setup A, and especially for the higher layers (22–29 km and 28–42 km), the temperature uncertainty accounts for most of error on the ozone partial columns. This fact illustrates that a simultaneous temperature retrieval is important when aiming on high quality middle stratospheric ozone data (Schneider and Hase, 2008). Applying the retrieval setup C, the FTIR technique provides ozone partial columns with an overall precision of better than 3 % for the tropopause and middle/upper stratosphere and of better than 6 % for the troposphere.

Regarding setup B and C, the spectroscopic parameters are responsible for most of the systematic errors, whereas for setup A the temperature uncertainty becomes an important systematic error source, especially above the troposphere.

High quality measurements are very important for trend studies, since they minimise possible artificial trends caused by drifts in the error sources. For example, a simultaneous temperature fit minimises the artificial trend that might be caused by a drift in the temperature uncertainty (e.g., it might be  $-1^\circ\text{C}$  in 2000 and gradually improve to  $\sim 0^\circ\text{C}$  in 2010). Another example is the ILS uncertainty (see Fig. 4). If a possible drift in the ILS is not adequately considered in the retrieval an artificial trend will be the consequence (in the following subsection, we demonstrate that it is very important to regularly monitor the ILS by laboratory cell measurements). Furthermore, a realistic constraint assures a correct interpretation of the variability as seen in the measured spectra, thereby leading to ozone data of a best possible quality (compare error budgets of setup B and C).

### 3.4 Long-term consistency of the ILS

Figure 4 shows that the ILS uncertainties are an important error source (in particular, middle and upper stratosphere). When aiming on a consistent long-term quality of ozone profiles, a continuous and precise documentation of the ILS is mandatory. Therefore, at Izaña we make regular low pressure  $\text{N}_2\text{O}$  cell measurements. These measurements allow for retrieving the actual ILS by means of the LINEFIT code

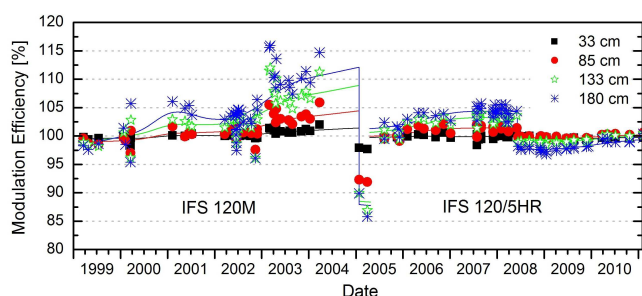


**Fig. 4.** VMR random and systematic errors relative to actual ozone VMR profiles [%] for the Izaña IFS 120/5HR for the setups A (a), B (b) and C (c). ILS means the joint error due to the modulation efficiency and phase error uncertainties and TPE (Total Parameter Error, black line) is the sum of all random errors except for smoothing error.

**Table 4.** Estimated random errors relative to actual ozone partial columns and total column [%] for the Izaña IFS 120/5HR for all setups and for the different layers.

| Layer<br>[km] | Retrieval Setup                        |                                      |                                      |
|---------------|--|--------------------------------------|--------------------------------------|
|               | A                                      | B                                    | C                                    |
|               | TPE (ILS, Temp, Noi), SE, <b>TRE</b>   | TPE (ILS, Temp, Noi), SE, <b>TRE</b> | TPE (ILS, Temp, Noi), SE, <b>TRE</b> |
| 2.37–13       | 1.0 (0.4, 0.6, 0.8), 7.1, <b>7.2</b>   | 0.9 (0.4, 0.5, 0.6), 6.8, <b>6.9</b> | 1.0 (0.3, 0.5, 0.8), 5.5, <b>5.6</b> |
| 12–23         | 1.4 (0.4, 1.3, 0.5), 2.4, <b>2.8</b>   | 0.7 (0.5, 0.3, 0.4), 2.3, <b>2.4</b> | 0.7 (0.5, 0.3, 0.4), 1.8, <b>1.9</b> |
| 22–29         | 2.7 (< 0.1, 2.7, 0.4), 3.2, <b>4.2</b> | 0.9 (0.6, 0.3, 0.5), 3.2, <b>3.3</b> | 0.9 (0.5, 0.3, 0.5), 2.8, <b>2.9</b> |
| 28–42         | 4.3 (0.7, 4.2, 0.5), 3.4, <b>5.5</b>   | 2.0 (1.6, 0.7, 0.7), 2.9, <b>3.5</b> | 1.6 (1.3, 0.6, 0.6), 2.4, <b>2.9</b> |
| 2.37–120      | 2.3 (0.1, 2.3, 0.1), 0.2, <b>2.3</b>   | 0.7 (0.4, 0.4, 0.2), 0.1, <b>0.7</b> | 0.6 (0.4, 0.3, 0.2), 0.2, <b>0.6</b> |

TPE [%]: Total Parameter Error due to input parameters and measurement noise; ILS [%]: Error due to instrumental line shape; Temp [%]: Error due to temperature; Noi [%]: Error due to measurement noise; SE [%]: Smoothing Error; TRE [%], in bold: Total Random Error.

**Fig. 5.** Times series of the modulation efficiency [%] at different optical path differences (OPD) for the Izaña spectrometers. Individual data points indicate individual cell measurements. The lines are the smoothed efficiency curves used during the FTIR retrievals. Black at 38 cm, red at 85 cm, green at 133 cm and blue at 180 cm.

(Hase et al., 1999). LINEFIT estimates the ILS at 20 equidistant optical path intercepts. At Izaña we have performed such cell measurements about every two months. Figure 5 shows the 1999–2010 time series of the modulation efficiency obtained for the two spectrometers. The jumps in the time series (beginning of 2005 for IFS 120M and June 2008 for the IFS 120/5HR) are due to realignments of the instruments.

As expected, we observe that the ILS of IFS 120M is significantly less stable than the ILS of IFS 120/5HR. While for the IFS 120/5HR the modulation efficiency can be determined with a precision of about 1 %, for the IFS 120M we only achieve a precision of 3 %. Concerning the phase error, we make a similar observation: with our regular cell measurements we can achieve a precision of 0.005 rad for the IFS 120/5HR, but only of 0.025 rad for the IFS 120M. It is important to note we apply the ILS as obtained from the cell measurement for our ozone retrieval. Otherwise, the resulting error in the ozone profiles would be much larger than estimated in Fig. 4 and Table 4.

### 3.5 Comparison between IFS 120M and IFS 120/5HR

During March and April 2005 both instruments (IFS 120M and IFS 120/5HR) measured side-by-side. The comparison between these coincidental measurements is displayed in Fig. 6, where the ozone partial columns as obtained from retrieval setup C are shown.

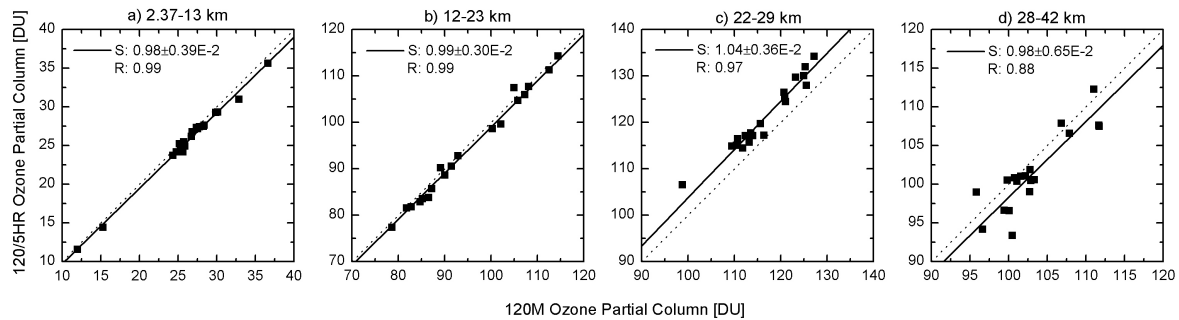
The agreement is excellent, except for the highest layers, where the FTIR data are more sensitive to instrumental uncertainties (ILS and measurement noise, see Fig. 4). For example, we observe a mean ratio between the IFS 120/5HR and 120M data of  $0.98 \pm 0.39 \times 10^{-2}$  ( $\pm 1$  standard error of the mean value) with a correlation of 99 % in 2.37–13 km layer and  $0.98 \pm 0.65 \times 10^{-2}$  with a correlation of 88 % in 28–42 km layer. These values are for setup C results. When comparing the IFS 120M and IFS 120/5HR results obtained by retrieval setups A and B, we find more scatter and lower correlation coefficients at the highest altitudes. This is expected from the error estimation (more uncertainty in setup A and B than in setup C retrieval data, see Table 4). Note also that higher errors are expected for the IFS 120M ozone retrievals due to its higher ILS uncertainties (see Fig. 5) and its higher measurement noise and, consequently, lower sensitivity. The lower sensitivity is clearly observed by comparing the DOFS time series of retrieved ozone from the two spectrometers (see Fig. 3).

We decided not to correct the ozone partial column time series from IFS 120M (1999–2004) by the bias as derived during the two months side-by-side intercomparison period. We think that this could introduce artificial trends, since a two months inter-comparison period in 2005 cannot be perfectly representative for the whole IFS 120M time series (1999–2005). Instead, we document the long-term consistency of the FTIR data by an intercomparison to the independent ECC sonde time series (see Sect. 5).

**Table 5.** Statistics of coincident measurements from the IFS 120/5HR and IFS 120M ozone partial columns for each setup and layer ( $N = 19$ ).

| Layer<br>[km] | Retrieval Setup   |   |   |
|---------------|---|---|---|
|               | A<br>$M \pm \text{SEM}, \sigma, S, R$                   | B<br>$M \pm \text{SEM}, \sigma, S, R$                   | C<br>$M \pm \text{SEM}, \sigma, S, R$                   |
| 2.37–13       | $-0.3 \pm 0.4, 1.8, 1.00 \pm 0.42 \times 10^{-2}, 0.99$ | $+2.8 \pm 0.4, 2.0, 0.97 \pm 0.41 \times 10^{-2}, 0.99$ | $-2.6 \pm 0.4, 1.9, 0.98 \pm 0.39 \times 10^{-2}, 0.99$ |
| 12–23         | $+1.0 \pm 0.4, 1.6, 0.99 \pm 0.36 \times 10^{-2}, 0.99$ | $+1.4 \pm 0.3, 1.4, 0.99 \pm 0.32 \times 10^{-2}, 0.99$ | $+1.1 \pm 0.3, 1.3, 0.99 \pm 0.30 \times 10^{-2}, 0.99$ |
| 22–29         | $-6.2 \pm 0.4, 1.9, 1.07 \pm 0.48 \times 10^{-2}, 0.96$ | $-4.0 \pm 0.4, 1.7, 1.04 \pm 0.41 \times 10^{-2}, 0.97$ | $-3.7 \pm 0.3, 1.5, 1.04 \pm 0.36 \times 10^{-2}, 0.97$ |
| 28–42         | $-3.5 \pm 0.8, 3.4, 1.04 \pm 0.83 \times 10^{-2}, 0.78$ | $+0.4 \pm 0.7, 2.9, 1.00 \pm 0.63 \times 10^{-2}, 0.84$ | $+1.7 \pm 0.6, 2.4, 0.98 \pm 0.65 \times 10^{-2}, 0.88$ |

$M \pm \text{SEM}$  [%]: mean and standard error of the mean of relative differences (120M-120/5HR)/120/5HR;  $\sigma$  [%]: standard deviation of relative differences;  $S$ : slope of the linear regression through origin;  $R$ : correlation coefficient of the linear fit.



**Fig. 6.** Comparison between the IFS 120/5HR and IFS 120M ozone partial columns [DU] for the 120M-120/5HR coincidences during March and April 2005 ( $N = 19$ ) and for setup C: (a) 2.37–13 km, (b) 12–23 km and (c) 22–29 km and (d) 28–42 km. The black solid lines are the linear regression lines through origin, whose parameters are shown in the legend ( $S$  is the slope of regression fit and  $R$  the correlation coefficient). The dotted lines are the diagonals.

#### 4 Consistency of ECC sonde time series

Before using the ECC data as reference for empirically assessing the quality of the FTIR ozone data and their representativeness for annual cycles and trends, it is very important to check the consistency of the ECC ozone time series. In this section, we use the coincident measurements of ozone total column from Brewer spectrometers and of surface ozone from in situ analysers for empirically documenting the quality and the long-term stability of the ECC sonde dataset.

##### 4.1 ECC sonde vs. Brewer

The consistency and quality of the ECC sonde ozone profile ( $x_{\text{ECC.org}}$ ) time series can be estimated by comparing it to independent and coincident very high quality measurements of ozone total amounts. At the Izaña Observatory such high quality measurements are performed by the FTIR system and by Brewer spectrometers. The Brewers have been operative since May 1991, and like the ECC sonde and FTIR Programs, they have been part of NDACC since March 2001. Furthermore, since November 2003 they represent the Regional Brewer Calibration Center for Europe (<http://www.rbcc-e.org/>) of WMO/GAW, which guarantees the high quality of their ozone total column measurements (better than 1 %, Redondas and Cede, 2006).

The ECC sondes normally burst between 30 and 34 km. In order to homogenise the study, we only consider the ECC data measured up to 29 km. Thus, the ozone total columns from the ECC sondes ( $\text{TC}_{\text{ECC.org}}$ ) have been calculated by integrating the ozone profiles up to 29 km and adding a typical residual ozone column up to the top of atmosphere. This residual has been estimated from the mean difference between the daily Brewer total ozone columns ( $\text{TC}_{\text{Brewer}}$ ) and ECC ozone partial amounts up to 29 km ( $\text{PC}_{\text{ECC.org}}$ ) for all 515 Brewer/ECC coincidences from 1999 to 2010 (for more details see Schneider et al., 2008b). We found a mean and a standard deviation ( $1\sigma$ ) for the ozone residual of 81.5 and 10.6 DU, respectively. Note that in our study, we apply Brewer measurements in order to remain independent from the FTIR data.

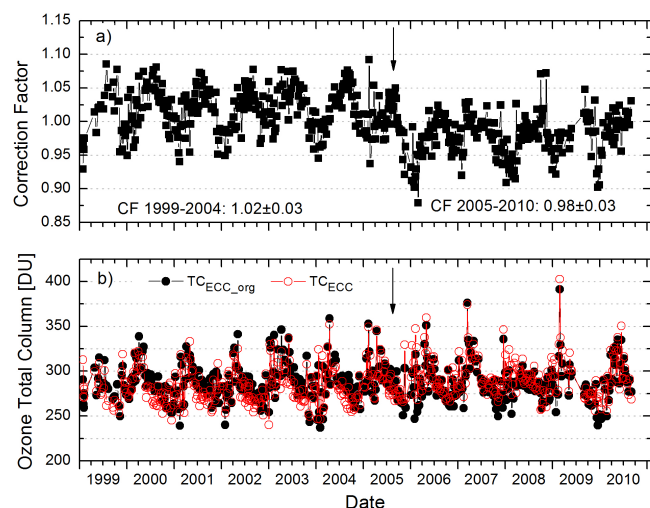
The Brewer total ozone columns can be used to correct the ECC sonde profiles ( $x_{\text{ECC}}$ ) by:

$$x_{\text{ECC}} = \frac{\text{TC}_{\text{Brewer}}}{\text{PC}_{\text{ECC.org}} + 81.5} \cdot x_{\text{ECC.org}} = \text{CF} \cdot x_{\text{ECC.org}} \quad (1)$$

whereby CF ( $\text{CF} = \frac{\text{TC}_{\text{Brewer}}}{\text{PC}_{\text{ECC.org}} + 81.5}$ ) is the daily correction factor. Schneider et al. (2008b) illustrated that by this correction the quality of the ECC data can be significantly improved.

The CF time series is displayed in Fig. 7a. We observe a jump in 2005. In fact, the mean CF is  $1.02 \pm 0.03$  ( $\pm 1\sigma$ ) and  $0.98 \pm 0.03$  ( $\pm 1\sigma$ ) for the periods 1999–2004 and





**Fig. 7.** Time series of correction factor (CF, according to Eq. 1) at the Izaña Observatory (a) and of ozone total column amounts from ECC sondes without and with correction ( $TC_{ECC} = CF \cdot TC_{ECC,org}$ ) (b). The mean correction factors for the periods 1999–2004 and 2005–2010 are also shown. The black arrows indicate the change-point date.

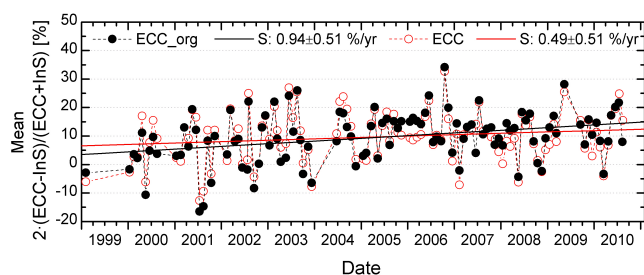
2005–2010, respectively. Such systematic change of about  $-4\%$  between the ECC records can be expected by changing the sensing solution type or ECC sonde type (Smit et al., 2007, and references therein). During 2005 the sensing solutions in the ECC sondes were substituted by a new batch, but the same manufacturer's operating procedures, and ratio of cathode sensing solutions, have been kept during whole period (1999–2010). Nonetheless, in the light of above results, we might assume that some change in the sensing solutions was introduced by using the new batch. The discontinuities in the CF time series is also identified by a rank order change point test (Lanzante, 1996; Romero et al., 2011). This non-parametric method, based on the ranks of the monthly values from a time series, is not particularly affected by gaps and outliers in the time series.

The time series of ECC ozone total columns, without and with correction ( $TC_{ECC,org}$  and  $TC_{ECC}$ , respectively), is displayed in Fig. 7b. The corrected ECC records benefit from the synergies of the two techniques: the high vertical resolution of the ECC sondes and the high precision of the Brewer measurements.

#### 4.2 ECC sonde vs. surface

At the Izaña Observatory the tropospheric surface ozone is also monitored by photometric in situ analysers. This surface dataset can also be used for documenting the long-term consistency of the ECC sonde profiles at Izaña's altitude.

In order to evaluate possible drifts in the ECC sonde dataset, we examine the time series of the differences in the monthly means between the ECC and the photometric



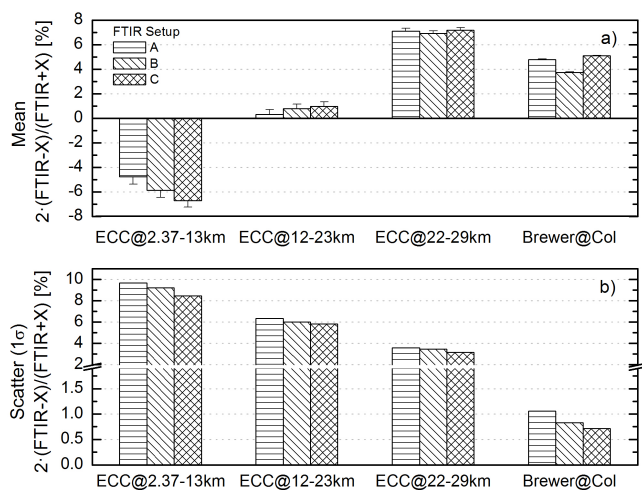
**Fig. 8.** Monthly mean time series of the relative differences [%] between the ozone VMR from the ECC sonde data (corrected, ECC, and not corrected by the daily CF,  $ECC_{org}$ ) at Izaña's altitude and surface data ( $N = 112$ ). The solid lines are the linear regression lines of the least square fits. The slopes and the 95 % confidence ranges ( $\pm 2\sigma$ ) are shown in the legend ( $S$ ).

surface data (Fig. 8). To do so the photometric surface data were averaged 30 min around the ECC measurements. For the ECC data corrected according to Eq. (1), the drift, i.e., the trend in the differences is not significant ( $0.49 \pm 0.51\% \text{ yr}^{-1}$ ). On average, we find a mean bias and a scatter ( $1\sigma$ ) of 10 % and 15 %, respectively, between the two techniques. Note that, if the ECC profiles are not corrected, we observed a significant drift in the difference between the ECC and the photometric surface data ( $0.94 \pm 0.51\% \text{ yr}^{-1}$ ), confirming the importance of correcting the ECC time series before using it for long-term studies.

Both the Brewer and the photometric surface measurements identify a jump in the ECC time series. This jump would significantly affect the estimated ozone trends as well as the inter-comparison with another technique. Therefore, in the subsequent analysis we will exclusively apply the corrected ECC sonde time series (corrected according to Eq. 1).

### 5 Validation of FTIR ozone profiles

The corrected ECC sonde time series can be used for evaluating the quality of the FTIR ozone profiles. To do so we have to consider that the vertical resolution of the ECC sonde and the FTIR profiles are rather different. Therefore, we only compare the layers that are sufficiently well detectable by the ground-based FTIR system, i.e., the ozone partial columns for the layers 2.37–13 km, 12–23 km and 22–29 km (the DOFS for all these layers is typically larger than one). Since the ECC sonde profiles are not smoothed by the FTIR's avks (e.g., Schneider et al., 2008b), the advantage of the improved vertical resolution observed for the more sophisticated retrieval setups can directly be validated. Even more important, not smoothing the ECC sonde data with the FTIR's avks assures that both the FTIR and the ECC time series remain completely independent and, for instance, the ECC sonde trends become not influenced by possible trends of the FTIR's avks (see Fig. 3). Here we compare



**Fig. 9.** Relative differences [%] between the ozone partial columns (2.37–13 km, 12–23 km and 22–29 km) calculated from the ECC sonde and from the different FTIR setups: (a) mean values (error bars indicate the standard error of the mean) and (b) scatter or standard deviation ( $1\sigma$ ). It is also shown the differences between the FTIR and Brewer ozone columns.  $X$  means the ECC or Brewer data.

two completely independent datasets, which would not be the case if we smooth the ECC data by the FTIR's avks.

The straightforward comparison of ozone partial columns from FTIR and corrected ECC sondes is rather satisfactory, as shown in Fig. 9. We find a mean bias between  $-8\%$  and  $-4\%$  in troposphere (2.37–13 km), about  $1\%$  in the tropopause region (12–23 km) and about  $7\%$  in middle stratosphere (22–29 km) for all retrieval setups, while the scatter is between  $6\%$  and  $9\%$  for the two first layers and only of  $3\%$  in middle stratosphere (see Fig. 9a and b). These results rather agree with the expected uncertainty for the ECC sondes given by Smit et al. (2007) ( $\pm 5\text{--}10\%$ ) and with our FTIR theoretical error estimation (Sect. 3.3). Furthermore, they are in good agreement with other inter-comparison studies using ECC sonde data (Calisesi et al., 2005; Nair et al., 2011 and references therein). It illustrates the good agreement between the ECC sonde and FTIR techniques. As example, Fig. 10 shows the direct comparison between the coincident FTIR (setup C) and ECC sonde partial columns from 1999 to 2010 ( $N = 262$ ).

Part of the discrepancy observed between FTIR and ECC sonde measurements is due to the smoothing error (recall Table 4). Smoothing the ECC sonde profiles with the FTIR avks improves the agreement for all setups by reducing the scatter, on average, to  $7\%$ ,  $5\%$  and  $2\%$  for the 2.37–13 km, 12–23 km and 22–29 km layers, respectively. Other sources of discrepancy might be errors in the FTIR and/or ECC data and the observation of different air masses by the FTIR experiment, on the one hand, and ECC experiment, on the other (Schneider et al., 2008b). The systematic discrepancies might be attributable to systematic ECC and FTIR errors

caused, for instance, by uncertainties in the spectroscopic line parameters.

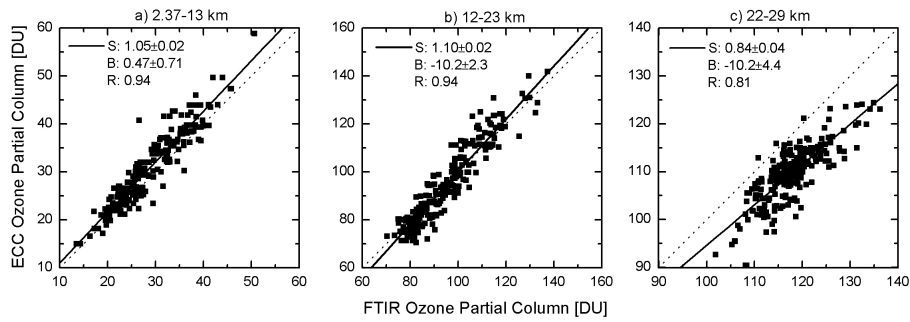
It is important to note that the scatter between FTIR and ECC concentrations decreases as the level of refinement of setups increases, especially in the troposphere and the tropopause region. Due to the reduction of the temperature error and the use of a realistic constraint, setup C provides the FTIR profiles that best agree with the ECC sonde profiles. This is in good agreement to our theoretical estimation as presented in Table 4 and Fig. 4. As addendum to the comparison to the ECC data, Fig. 9 shows the comparison to the Brewer data that have been used in the Schneider et al. (2008a) paper. We observe that the scatter between the FTIR and Brewer ozone total column amounts is significantly reduced (from  $1.06\%$  to  $0.71\%$ ) as the FTIR setups become more refined, confirming the results of the FTIR-ECC comparison. Similar results were found for the comparison between FTIR and Izaña's surface data, i.e., the best agreement is obtained for the retrieval setup C. We find a scatter ( $1\sigma$ ) of  $20\%$  for the setup C and of  $23\%$  for setup A. The scatter between the ECC sonde and Izaña's surface data is smaller ( $15\%$ , recall Sect. 4.2). The slightly increased scatter when comparing with the FTIR is due to the FTIR's smoothing error.

In order to evaluate possible artificial trend in the FTIR's DOFS time series, we analyse the time series of the differences between smoothed and unsmoothed ECC sonde data. This can be done for the 262 occasions during the 1999–2010 time period where FTIR and ECC measurements are performed in coincidence. For all setups, we observe that there are no significant trends in the 2.37–13 km and 12–23 km layers, but that there is a significant trend (at 95% of confidence) for the 22–29 km layer. For example, for setup C, the trend is of  $0.12 \pm 0.10 \text{ DU yr}^{-1}$  (i.e.,  $0.11 \pm 0.09 \text{ \% yr}^{-1}$ ), which might indicate that drifts in the avks affect the Izaña 1999–2010 middle/upper stratospheric FTIR ozone trend. These facts consolidate our applied strategy not to smooth the ECC sondes for trend comparisons. Otherwise, both instruments might show similar trends, due to the fact that there is a trend in the avks. However, we also have to consider that the data subset of ECC/FTIR coincidences is not well representative for the whole FTIR time series (the number of coincidences is rather small and most of them occur from April to October).

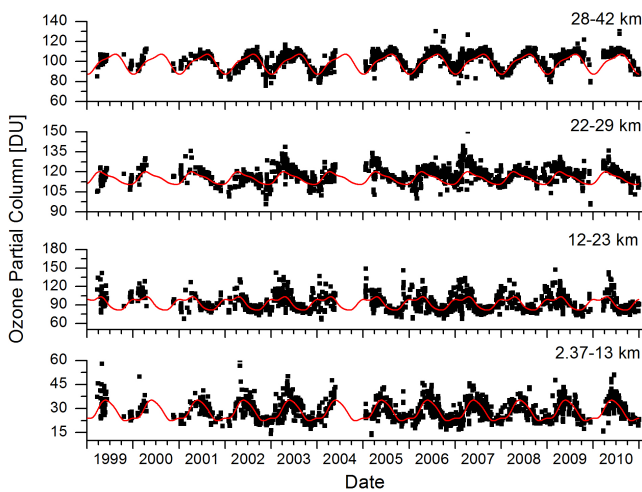
## 6 Trends and seasonality

### 6.1 Ozone trends

Obtaining observational evidence of the projected trends in the vertical distribution of ozone is a difficult task. The trends are rather small and small instrumental drifts might cause artificial trends. Therefore, it is important to apply different measurement techniques. Our study estimates the



**Fig. 10.** Comparison between the ozone partial columns from FTIR (setup C) and ECC sonde data ( $N = 262$ ): (a) 2.37–13 km, (b) 12–23 km and (c) 22–29 km. 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.



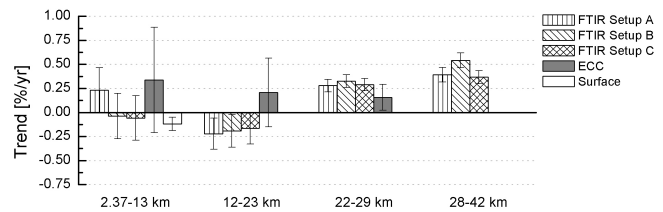
**Fig. 11.** Time series of FTIR ozone partial columns [DU] (setup C) for the different layers. Black dots: individual measurements ( $N = 1887$ ); Red lines: annual cycle according to Eq. (2).

ground-based FTIR ozone trends and examines its consistency to trends obtained from ECC sonde and surface in situ analyser datasets.

The ozone trends are estimated by using a bootstrap re-sampling method (Gardiner et al., 2008), which models the total variation in ozone by a function  $F(t)$  and allows for separating the annual cycles from possible long-term trends:

$$F(t) = f_o + f_{\text{trend}} t + \sum_{i=1}^p [a_i \cos(\omega_i t) + b_i \sin(\omega_i t)] \quad (2)$$

where  $t$  is measured in years,  $f_o$  is a baseline constant and  $f_{\text{trend}}$  the linear trend in change per year. The annual cycle is modelled 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 provided the best overall results (Gardiner et al., 2008;

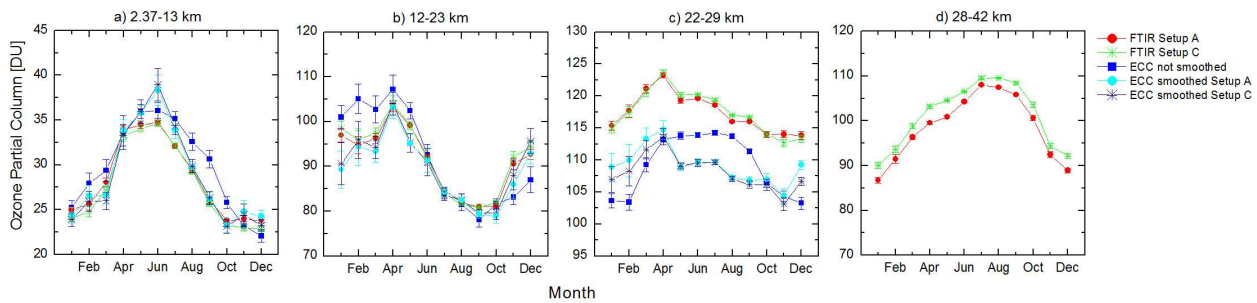


**Fig. 12.** Linear trends ( $\% \text{ yr}^{-1}$ ) between 1999 and 2010 for the FTIR ( $N = 1887$ ), the corrected ECC sonde ( $N = 515$ ) and surface in situ ( $N = 42415$ ) datasets. The error bars indicate the 95 % confidence ranges ( $\pm 2\sigma$ ).

Vigouroux et al., 2008). Figure 11 shows the FTIR ozone time series (setup C) and the fitted function  $F(t)$  for the different layers.

Although our bootstrap model does not capture atmospheric processes such as quasi-biennial oscillations (QBO) or solar cycle variations (e.g., Reinsel et al., 2002), they become a noise source in the linear trend determination and, thus, feed into the uncertainties in the determined trends (Vigouroux et al., 2008). 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).

The observed trends from the FTIR time series slightly depend on the applied retrieval setup (Fig. 12). However, these differences are not significant and lie clearly within the 95 % confidence range (see error bars in Fig. 12, note that the confidence range is given by  $2\sigma$  standard deviations of the bootstrap re-sampled distributions). For all setups, we observe a significant negative trends between 1999 and 2010 for the tropopause region, while in the middle/upper stratosphere the FTIR data document a significant stratospheric ozone increase. The corrected ECC sonde time series confirms the FTIR trends in the middle stratosphere, while for the rest of layers the ECC trends are not significant. In the troposphere, we observe no significant trends when analysing the datasets produced by the setups that include a temperature



**Fig. 13.** Annual cycle of the ozone partial columns [DU] obtained from the FTIR (setup A and C), the corrected ECC sonde data (not smoothed and smoothed by avks from the FTIR setup A and C): (a) 2.37–13 km, (b) 12–23 km, (c) 22–29 km and (d) 28–42 km. The error bars indicate the standard error of the mean value.

retrieval (setup B and C), whereas the setup A dataset indicates a positive trend. In order to get more insight into the long-term evolution of this tropospheric layer, we have also estimated the ozone trends from the night-time surface in situ data, whose for Izaña are well representative of the free troposphere conditions. Thus, the airmass detected by the in situ technique can be compared to the tropospheric airmass remotely sensed by the FTIR system (Sepúlveda et al., 2012). We observe a significant small negative trend (see Fig. 12), which better agrees with the setups that simultaneously fit the temperature profile. In summary, the trends obtained from the setup C datasets tend to be in better agreement with the trends obtained from the ECC sonde and surface in situ datasets, although the differences to the trends estimated from setup A and B datasets are rather small.

It is interesting to mention that the significant ozone trends observed by the FTIR system in the upper troposphere/lower stratosphere (negative trend) and in the middle/upper stratosphere (positive trend) are also predicted by current climate models at the northern subtropical latitudes (Hegglin and Shepherd, 2009; Li et al., 2009). For example, Li et al. (2009) estimate that a decrease of ozone is to be expected in the lower stratosphere in the northern subtropical latitudes associated with the predicted increase in the stratospheric Brewer-Dobson circulation. An accelerated Brewer-Dobson circulation transports more ozone from tropics to the mid-high latitudes and could delay ozone recovery in the tropics and advanced ozone recovery in the extra-tropics. However, we have to be aware that a 12-yr time series might be too short for validating the models due to the large year-to-year variability of this stratospheric circulation (Weber et al., 2011). Likewise, Li et al. (2009) show that the photochemical response to strong cooling induced by increasing greenhouse gases concentrations would lead to an upper stratospheric ozone increase. In addition, the leveling off of anthropogenic halogen components in the upper stratosphere since the mid-1990s has to be considered (WMO, 2007). Similar results were reported by WMO (2011).

For the upper stratosphere the quality of the FTIR ozone time series has not been empirically validated in this work

by day-to-day intercomparisons due to the lack of respective ECC data. Previous campaign studies show good agreement between the ozone measurements obtained from FTIR and other measurement techniques in this layer, such as ground-based LIDAR and millimeter-wave radiometer (Kopp et al., 2002; Vigouroux et al., 2008). However, at these altitudes the FTIR ozone data are very sensitive to ILS uncertainties and to measurement noise. Consequently, our upper stratospheric trend estimations should be treated with care, although they are consistent for all setups and are supported by other experimental studies. For example, Steinbrecht et al. (2009) found that the upper stratospheric ozone (35–45 km) has been slightly increasing since the late 1990s at other NDACC sites at similar latitudes (Mauna Loa, 19.5° N 155.6° E, and Table Mountain, 34.5° N 117.7° E), using satellite- and ground-based LIDAR measurements. This stratospheric ozone recovery has also been documented experimentally by FTIR records in northern subtropical, middle and polar latitudes (updated from Vigouroux et al., 2008 in WMO, 2011).

## 6.2 Ozone annual cycle

The annual ozone cycles have been calculated by producing monthly averages considering the whole time series. Figure 13 shows the annual cycles for the FTIR data (setup A and C) and the corrected ECC sonde data (not smoothed and smoothed by avks from the FTIR setup A and C). We observe that the agreement between the two techniques is rather satisfactory.

The annual cycle for the troposphere (2.37–13 km) reveals a maximum in spring-summer, which indicates the importance of photochemical production of tropospheric ozone. In the tropopause region (12–23 km) we observe a maximum in winter-spring and a minimum in summer and in early autumn, which is due to the large vertical shift of the subtropical tropopause and this layer belongs to the lower stratosphere, while in summer there is a tropical tropopause and this layer has rather upper tropospheric characteristics. The transformation from mid-latitudinal to tropical characteristics is also

observed in the middle/upper stratosphere (22–29 km and 28–42 km). There we observe maximum ozone concentrations in summer-autumn when the tropical conditions prevail. In the middle stratosphere the FTIR system detects a maximum in spring that is not found in the ECC sonde annual cycle. This is due to the smoothing error: the FTIR ozone partial column in this layer contains information from the lower layers (12–23 km), where the spring maximum is due to the annual cycle in the tropopause altitude.

Note that applying different constraints and temperature retrievals do not significantly affect the ozone seasonality.

## 7 Conclusions

In this paper, we document the quality of the ozone profiles obtained from ground-based FTIR systems and discuss its application for long-term studies. We investigate three different retrieval setups: (A) an ad-hoc constraint for ozone and no temperature profile retrieval, (B) an ad-hoc constraint for ozone and a simultaneous temperature profile retrieval, and (C) an ozone constraint based on an ozone climatology (optimal estimation retrieval) and a simultaneous temperature profile retrieval.

Our theoretical error assessment reveals that the measurement noise and the uncertainties in the ILS and the applied temperature profile (for setup A) are the leading error sources. In particular, the retrieved middle/upper stratospheric ozone amounts are strongly affected by ILS and temperature uncertainties. We show that the temperature error can be significantly reduced by performing a simultaneous temperature profile retrieval. Moreover, the ad-hoc constraint retrievals offer more DOFS in the middle/upper stratosphere than the optimal estimation retrieval. At lower altitudes it is vice versa. An optimal estimation retrieval is supposed to interpret the measured spectra in a best possible manner. Consequently, the ad-hoc constraint retrievals might misinterpret ozone variability (over-/under-estimate variability at higher/lower altitudes).

For an empirical quality assessment, we use a coincident ECC sonde ozone profile dataset as reference, whose quality, in turn, has been checked, independently from the FTIR data, by a comparison to Brewer total column and surface in situ measurements. During the 12-yr period of 1999–2010, the agreement between the vertical ozone distribution obtained by the FTIR and the ECC sondes is very satisfactory. We show empirically that the FTIR system is well able to capture the day-to-day ozone variability in the troposphere, tropopause region and middle stratosphere. Furthermore, both techniques reveal very similar annual seasonality. For the ozone retrieval setup that applies a constraint based on an ozone climatology and includes a simultaneous temperature profile retrieval we observe a slightly better agreement than for the other setups. These observations confirm our theoretical quality assessment.

Regarding ozone trends, we estimate the trends for the 1999–2010 time period for the ECC, surface in situ analyzers, and FTIR datasets. In the middle stratosphere we observe a significant positive trend (95 % confidence interval) for the ECC and for all three FTIR datasets (the FTIR also reveals a significant positive trend above 30 km, where there are no ECC data available). In the upper troposphere/lower stratosphere region the FTIR observes a significant negative trend (95 % confidence interval), which cannot be confirmed by the ECC dataset. At these altitudes ozone amounts are very variable and the trend estimates are rather uncertain. This is especially true for the ECC trend estimates, since there is only one ECC observation per week. FTIR observations are made more frequently (several times per week), leading to smaller uncertainties in the trend estimates. In the troposphere we observe no significant trend neither in the ECC nor in the FTIR datasets. The surface in situ data reveal a significant negative trend, which is interestingly also indicated by the two FTIR retrieval setups that apply a simultaneous temperature retrieval.

A main reason for this satisfactory agreement is the fact that we take a lot of care in documenting the ILS (see Fig. 5), thereby avoiding artificial trends due to drifts in the ILS. A regular ILS monitoring, applying low pressure gas cell measurements, is very important for FTIR trend studies of stratospheric absorbers. Furthermore, we think that a simultaneous temperature retrieval is important, since it can significantly reduce the risk of artificial trends caused by possible drifts in the temperature uncertainty, thereby theoretically increasing the reliability of the FTIR trends. In our study, we observe that the temperature retrieval modifies the estimated trends. Using a realistic constraint instead of an ad-hoc constraint is important for reproducing the large day-to-day variability (see comparisons in Sect. 5), and it also slightly affects the estimated trends. Finally, one should consider the temporal evolution of the DOFS when using remote-sensing data for trend studies. For example, if there is a bias in the remote-sensing data, this bias will very likely decrease with decreasing DOFS, thereby giving rise of an artificial trend.

In summary, we think that correctly estimating the small expected ozone trends is a very difficult task for any measurement technique. In this context, ozone super-sites like the Izaña Observatory, that concentrate numerous measurement techniques, are important. They allow for intercomparing the techniques, thereby documenting the long-term consistency of the different profile datasets. Small trends that are consistently detected by different and independently working measurement techniques are rather reliable, whereas a small trend detected by an individual technique implicates the risk to be artificial and caused by drifts in error sources or instrumental properties.

*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 Program (FP7/2007-2013) under grant agreement no. 284421 (see Article II. 30. of the Grant Agreement). We are grateful to the Goddard Space Flight Center for providing the temperature and pressure profiles of the National Centers for Environmental Prediction. M. Schneider is supported by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013)/ERC Grant agreement no. 256961 and E. Sepúlveda enjoys a pre-doctoral fellowship from the Spanish Ministry of Education. Finally, the authors would like to thank Carlos Marrero and Pedro M. Romero for their useful comments.

Edited by: M. Weber

## References

- Anderson, J. G., Wilmouth, D. M., Smith, J. B., and Sayres, D. S.: UV Dosage Levels in Summer: Increased Risk of Ozone Loss from Convectively Injected Water Vapor, *Science*, 835–839, 2012.
- Barret, B., De Mazière, M., and Demoulin, P.: Retrieval and characterization of ozone profiles from solar infrared spectra at the Jungfraujoch, *J. Geophys. Res.*, 107, 4788–4803, 2002.
- Calisesi, Y., Soebijanta, V., and van Oss, R.: Regridding of remote soundings: Formulation and application to ozone profile comparison, *J. Geophys. Res.*, 110, D23306, doi:10.1029/2005JD006122, 2005.
- Gardiner, T., Forbes, A., de Mazière, M., Vigouroux, C., Mahieu, E., Demoulin, P., Velasco, V., Notholt, J., Blumenstock, T., Hase, F., Kramer, I., Sussmann, R., Stremme, W., Mellqvist, J., Strandberg, A., Ellingsen, K., and Gauss, M.: Trend analysis of greenhouse gases over Europe measured by a network of ground-based remote FTIR instruments, *Atmos. Chem. Phys.*, 8, 6719–6727, doi:10.5194/acp-8-6719-2008, 2008.
- Hase, F., Blumenstock, T., and Paton-Walsh, C.: Analysis of the instrumental line shape of high-resolution Fourier transform IR spectrometers with gas cell measurements and new retrieval software, *Appl. Optics*, 38, 3417–3422, 1999.
- Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Höfner, M., Jones, N. B., Rinsland, C. P., and Wood, S. W.: Intercomparison of retrieval codes used for the analysis of high-resolution ground-based FTIR measurements, *J. Quant. Spectrosc. Ra.*, 87, 25–52, 2004.
- Hegglin, M. I. and Shepherd, T. G.: Large climate-induced changes in ultraviolet index and stratosphere-to troposphere ozone flux, *Nat. Geosci.*, 2, 687–691, doi:10.1038/NNGEO604, 2009.
- Johnson, D., Jucks, K., Traub, W., and Chance, K.: Isotopic composition of stratospheric ozone, *J. Geophys. Res.*, 105, 9025–9031, 2000.
- Kopp, G., Berg, H., Blumenstock, T., Fischer, H., Hase, F., Hochschild, G., Höpfner, M., Kouker, W., Reedmann, T., Ruhnke, R., Raffalsky, U., and Kondo, Y.: Evolution of ozone and ozone-related species over Kiruna during the SOLVE/THESEO 2000 campaign retrieved from ground-based millimeter-wave and infrared observations, *J. Geophys. Res.*, 107, 8308, doi:10.1029/2001JD001064, 2002.
- Kurylo, M. J. and Zander, R.: The NDSC-Its status after 10 years of operation, in: Proceedings of XIX Quadrennial Ozone Symposium, Hokkaido University, Sapporo, Japan, 167–168, 2000.
- Lanzante, J. R.: Resistant, robust and non-parametric techniques or the analysis of climate data: theory and examples, including applications to historical radiosonde station data, *Int. J. Climatol.*, 16, 1197–1226, 1996.
- Li, F., Stolarski, R. S., and Newman, P. A.: Stratospheric ozone in the post-CFC era, *Atmos. Chem. Phys.*, 9, 2207–2213, doi:10.5194/acp-9-2207-2009, 2009.
- Mauersberger, K., Lämmerzahl, P., and Krankowsky, D.: Stratospheric ozone isotope enrichments-revisited, *Geophys. Res. Lett.*, 28, 3155–3158, 2001.
- Nair, P. J., Godin-Beekmann, S., Pazmiño, A., Hauchecorne, A., Ancellet, G., Petropavlovskikh, I., Flynn, L. E., and Froidevaux, L.: Coherence of long-term stratospheric ozone vertical distribution time series used for the study of ozone recovery at a northern mid-latitude station, *Atmos. Chem. Phys.*, 11, 4957–4975, doi:10.5194/acp-11-4957-2011, 2011.
- Redondas, A. and Cede, A.: Brewer algorithm sensitivity analysis, in: Sauna Workshop, Puerto de la Cruz, Tenerife, 2006.
- Reinsel, G. C., Weatherhead, E. C., Tiao, G. C., Miller, A. J., Nagatani, R. M., Wuebbles, D. J., and Flynn, L. E.: On detection of turnaround and recovery in trend for ozone, *J. Geophys. Res.*, 107, 4078–4090, 2002.
- Rodgers, C.: *Inverse Methods for Atmospheric Sounding: Theory and Praxis*, World Scientific Publishing Co., Singapore, 2000.
- Romero, P. M., Marrero, C. L., Alonso-Pérez, S., Cuevas, E., and Afonso, S.: Una climatología del agua precipitable en la región subtropical sobre la Isla de Tenerife, vol. NTD CIAI-6, NIPO: 281-12-007-5, Servicio de Publicaciones de la AEMET, Madrid, Spain, 2011.
- Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Fally, S., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W., Mandin, J.-Y., Massie, S. T., Mikhailenko, S. N., Miller, C. E., Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., A. Perrini, A. P.-C., Rinsland, C. P., Rotger, M., Šimečková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C., and VanderAuwera, J.: The HITRAN 2008 Molecular Spectroscopic Database, *J. Quant. Spectrosc. Ra.*, 110, 533–572, 2009.
- Schneider, M. and Hase, F.: Technical Note: Recipe for monitoring of total ozone with a precision of around 1 DU applying mid-infrared solar absorption spectra, *Atmos. Chem. Phys.*, 8, 63–71, doi:10.5194/acp-8-63-2008, 2008.
- Schneider, M., Hase, F., and Blumenstock, T.: Ground-based remote sensing of HDO/H<sub>2</sub>O ratio profiles: introduction and validation of an innovative retrieval approach, *Atmos. Chem. Phys.*, 6, 4705–4722, doi:10.5194/acp-6-4705-2006, 2006.
- Schneider, M., Redondas, A., Hase, F., Guirado, C., Blumenstock, T., and Cuevas, E.: Comparison of ground-based Brewer and FTIR total column O<sub>3</sub> monitoring techniques, *Atmos. Chem. Phys.*, 8, 5535–5550, doi:10.5194/acp-8-5535-2008, 2008a.

- Schneider, M., Hase, F., Blumenstock, T., Redondas, A., and Cuevas, E.: Quality assessment of O<sub>3</sub> profiles measured by a state-of-the-art ground-based FTIR observing system, *Atmos. Chem. Phys.*, 8, 5579–5588, doi:10.5194/acp-8-5579-2008, 2008b.
- Sepúlveda, E., Schneider, M., Hase, F., García, O. E., Gomez-Pelaez, A., Dohe, S., Blumenstock, T., and Guerra, J. C.: Long-term validation of tropospheric column-averaged CH<sub>4</sub> mole fractions obtained by mid-infrared ground-based FTIR spectrometry, *Atmos. Meas. Tech.*, 5, 1425–1441, doi:10.5194/amt-5-1425-2012, 2012.
- Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J., Northam, T., and J. C. Witte, A. M. T., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozone sondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), *J. Geophys. Res.*, 112, D19306, doi:10.1029/2006JD007308, 2007.
- Steinbrecht, W., Claude, H., Schönnenborn, F., McDermid, I. S., Leblanc, T., Godin-Beekmann, S., Keckhut, P., Hauchecorne, J. A. E. V. G., Swart, D. P. J., Bodeker, G. E., Parrish, A., Boyd, I. S., Kämpfer, N., Hocke, K., Stolarski, R. S., Frith, S. M., Thomason, L. W., Remsberg, E. E., Savigny, C. V., Rozanov, A., and Burrows, J. P.: Ozone and temperature trends in the upper stratosphere at five stations of the Network for the Detection of Atmospheric Composition Change, *Int. J. Remote Sens.*, 30, 3875–3886, 2009.
- Viatte, C., Schneider, M., Redondas, A., Hase, F., Eremenko, M., Chelin, P., Flaud, J.-M., Blumenstock, T., and Orphal, J.: Comparison of ground-based FTIR and Brewer O<sub>3</sub> total column with data from two different IASI algorithms and from OMI and GOME-2 satellite instruments, *Atmos. Meas. Tech.*, 4, 535–546, doi:10.5194/amt-4-535-2011, 2011.
- Vigouroux, C., De Mazière, M., Demoulin, P., Servais, C., Hase, F., Blumenstock, T., Kramer, I., Schneider, M., Mellqvist, J., Strandberg, A., Velasco, V., Notholt, J., Sussmann, R., Stremme, W., Rockmann, A., Gardiner, T., Coleman, M., and Woods, P.: Evaluation of tropospheric and stratospheric ozone trends over Western Europe from ground-based FTIR network observations, *Atmos. Chem. Phys.*, 8, 6865–6886, doi:10.5194/acp-8-6865-2008, 2008.
- Weatherhead, E. and Andersen, S.: The search for signs of recovery of the ozone layer, *Nature*, 441, 39–45, 2006.
- Weber, M., Dikty, S., Burrows, J. P., Garny, H., Dameris, M., Kubin, A., Abalichin, J., and Langematz, U.: The Brewer-Dobson circulation and total ozone from seasonal to decadal time scales, *Atmos. Chem. Phys.*, 11, 11221–11235, doi:10.5194/acp-11-11221-2011, 2011.
- WMO: Scientific Assessment of Ozone Depletion: 2006, Global Ozone Research and Monitoring Project-Report No. 50, World Meteorological Organization, Geneva, Switzerland, 2007.
- WMO: Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project-Report No. 2, World Meteorological Organization, Geneva, Switzerland, 2011.
- Zellweger, C., Klausen, J., Buchmann, B., and Scheel, H.-E.: System and Performance Audit of Surface Ozone, Carbon Monoxide, Methane and Nitrous Oxide at the Global GAW Station Izaña, Tech. rep., WCC-EMPA Report 09/1, 2009.