



# A tropopause-related climatological a priori profile for IASI-SOFRID ozone retrievals: improvements and validation

Brice Barret<sup>1</sup>, Emanuele Emili<sup>2</sup>, and Eric Le Flochmoen<sup>1</sup>

<sup>1</sup>Laboratoire d'Aérodynamique/Omp, Université de Toulouse, Toulouse, France

<sup>2</sup>CECI, Université de Toulouse, CERFACS, CNRS, Toulouse, France

**Correspondence:** Brice Barret (brice.barret@aero.obs-mip.fr)

Received: 9 January 2020 – Discussion started: 28 February 2020

Revised: 15 July 2020 – Accepted: 2 August 2020 – Published: 6 October 2020

**Abstract.** The MetOp/Infrared Atmospheric Sounding Interferometer (IASI) instruments have provided data for operational meteorology and document atmospheric composition since 2007. IASI ozone ( $O_3$ ) data have been used extensively to characterize the seasonal and interannual variabilities and the evolution of tropospheric  $O_3$  at the global scale. Software for a Fast Retrieval of IASI Data (SOFRID) is a fast retrieval algorithm that provides IASI  $O_3$  profiles for the whole IASI period. Until now, SOFRID  $O_3$  retrievals (v1.5 and v1.6) were performed with a single a priori profile, which resulted in important biases and probably a too-low variability. For the first time, we have implemented a comprehensive dynamical a priori profile for spaceborne  $O_3$  retrievals which takes the pixel location, time and tropopause height into account for SOFRID- $O_3$  v3.5 retrievals. In the present study, we validate SOFRID- $O_3$  v1.6 and v3.5 with electrochemical concentration cell (ECC) ozonesonde profiles from the global World Ozone and Ultraviolet Radiation Data Centre (WOUDC) database for the 2008–2017 period. Our validation is based on a thorough statistical analysis using Taylor diagrams. Furthermore, we compare our retrievals with ozonesonde profiles both smoothed by the IASI averaging kernels and raw. This methodology is essential to evaluate the inherent usefulness of the retrievals to assess  $O_3$  variability and trends. The use of a dynamical a priori profile largely improves the retrievals concerning two main aspects: (i) it corrects high biases for low-tropospheric  $O_3$  regions such as the Southern Hemisphere, and (ii) it increases the retrieved  $O_3$  variability, leading to a better agreement with ozonesonde data. Concerning upper troposphere–lower stratosphere (UTLS) and stratospheric  $O_3$ , the improvements are less important and the biases are very similar

for both versions. The SOFRID tropospheric ozone columns (TOCs) display no significant drifts ( $< 2.5\%$ ) for the Northern Hemisphere and significant negative ones (9.5% for v1.6 and 4.3% for v3.5) for the Southern Hemisphere. We have compared our validation results to those of the Fast Optimal Retrievals on Layers for IASI (FORLI) retrieval software from the literature for smoothed ozonesonde data only. This comparison highlights three main differences: (i) FORLI retrievals contain more theoretical information about tropospheric  $O_3$  than SOFRID; (ii) root mean square differences (RMSDs) are smaller and correlation coefficients are higher for SOFRID than for FORLI; (iii) in the Northern Hemisphere, the 2010 jump detected in FORLI TOCs is not present in SOFRID.

## 1 Introduction

Ozone ( $O_3$ ) in the stratosphere protects life from solar UV radiation. Close to the surface,  $O_3$  is an oxidative pollutant harmful for human health through irritation of respiratory tracts (Brunekreef and Holgate, 2002) and for vegetation through deposition on leaves that leads to the reduction of plant growth (Ainsworth et al., 2012). Tropospheric  $O_3$  is also a powerful greenhouse gas whose increase during the 20th century has significantly contributed to global warming (Shindell et al., 2006). The radiative forcing of  $O_3$  is particularly important in the tropical upper troposphere–lower stratosphere (UTLS) (Chen et al., 2007).

It is therefore important to document the evolution of  $O_3$  in these different layers independently. There is clear evidence from satellite databases that upper stratospheric  $O_3$  has in-

creased since 1997 following the ban of chlorofluorocarbons (CFCs) by the Montreal Protocol (Ball et al., 2018). Nevertheless, the total column  $O_3$  has been stable since 1998. According to Ball et al. (2018), this contradiction is due to the fact that lower stratospheric  $O_3$  is declining and compensates both stratospheric and tropospheric  $O_3$  increase. Based on Ozone Monitoring Instrument/Microwave Limb Sounder (OMI/MLS) tropospheric ozone columns (TOCs), they state that TOC is globally increasing. OMI/MLS data for the 2005–2016 period are indeed documenting global positive TOC trends with particularly large increases over Asia (Ziemke et al., 2019). Based on 10 years of retrievals with the Fast Optimal Retrievals on Layers for Infrared Atmospheric Sounding Interferometer (IASI)  $O_3$  (FORLI- $O_3$ ) software, Wespes et al. (2018) document a decrease in tropospheric  $O_3$  levels in the Northern Hemisphere (NH). Another IASI tropospheric  $O_3$  product (Karlsruhe Optimized and Precise Radiative transfer Algorithm Fit; KOPRAFIT- $O_3$ ) displays a TOC decrease over continental China (Dufour et al., 2018). In their exhaustive work on TOC evolution, Gaudel et al. (2018) clearly highlight the contradiction between global increase (OMI/MLS and other UV-vis products) on the one hand and global decrease (IASI) on the other hand. They also show that the different satellite products agree on a TOC increase over Asia. Among the two global IASI TOC datasets used in Gaudel et al. (2018), FORLI- $O_3$  indicates a significant global decrease and  $O_3$  retrievals with the Software for a Fast Retrieval of IASI Data (SOFRID) indicate a slightly weaker and less significant one. Two versions of FORLI- $O_3$  have been validated by Boynard et al. (2016) (v20141022) and Boynard et al. (2018) (v20151001). They both document a jump in the  $O_3$  retrievals in 2010 but this does not hinder the fact that TOCs are decreasing according to Wespes et al. (2017). It has to be noted that both validation studies compare IASI retrievals to ozonesonde profiles smoothed by the retrieval averaging kernels. Such a comparison enables the detection of abnormal biases, variability or drifts in the retrievals but does not document the ability of FORLI- $O_3$  to reproduce real  $O_3$  levels and variabilities. SOFRID- $O_3$  has only been validated at the beginning of the IASI period on a very short time period (Barret et al., 2011) and on a longer time period together with FORLI- $O_3$  and KOPRAFIT- $O_3$  (Dufour et al., 2012). Furthermore, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) L2 atmospheric temperature products retrieved from IASI and used for FORLI (v20141022 and v20151001) and for SOFRID- $O_3$  v1.5 retrievals are not stable in time (Boynard et al., 2018). Therefore, we have reprocessed the whole IASI database using ECMWF operational analyses for temperature and humidity to produce SOFRID- $O_3$  v1.6. SOFRID- $O_3$  has been shown to overestimate low tropospheric ozone over the Southern Hemisphere (SH) (Dufour et al., 2012; Emili et al., 2014, 2019). Emili et al. (2014) have hypothesized that this overestimation was due to the use of a single a priori profile biased towards NH

midlatitudes  $O_3$ . In order to verify this hypothesis and to improve our  $O_3$  retrievals, we have developed a new version of SOFRID- $O_3$  (v3.5), with a dynamical a priori profile based on a global  $O_3$  climatology (Sofieva et al., 2014).

The aim of the present paper is to validate both of the latest SOFRID- $O_3$  products (v1.6 and v3.5) for the whole IASI period (2008–2017) in order to infer their ability to reproduce tropospheric  $O_3$  levels and variability on seasonal to decadal timescales. The validation is based on  $O_3$  profiles from ozonesondes retrieved from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) database. In Sect. 2, we describe the characteristics and differences of SOFRID- $O_3$  v1.6 and v3.5 retrievals. Section 3 is dedicated to the description of the validation methodology based on comparisons between smoothed and raw ozonesonde data, and we provide our validation results in Sect. 4. Based on Boynard et al. (2018), we also compare our results to FORLI- $O_3$  (Sect. 5) before concluding the paper in Sect. 6.

## 2 IASI SOFRID- $O_3$ retrievals

IASI is a spaceborne thermal infrared nadir spectrometer. IASI has a moderate spectral resolution combined with a high signal-to-noise ratio and a 12 km footprint at nadir (Clerbaux et al., 2009). Thanks to its large across-track scanning ( $\sim 2200$  km), IASI revisits each scene twice daily around 09:30 LT solar time in the morning and in the evening. Three IASI instruments have been launched on the MetOp meteorological platforms (MetOp-A in 2006, MetOp-B in 2012 and MetOp-C in 2018). Here, we present results based on  $O_3$  retrievals from 10 years of MetOp-A/IASI data. We will present results based on the morning overpass data only as they are known to provide more information than nighttime data. Furthermore, it facilitates the comparison to other validation studies (Boynard et al., 2018) also based on morning data.

The SOFRID software first described in Barret et al. (2011) is based on the RTTOV (Radiative Transfer for TIROS Operational Vertical Sounder) operational radiative transfer code (Saunders et al., 1999; Matricardi et al., 2004) combined with the 1D-Var software (Pavelin et al., 2008), both developed within the framework of EUMETSAT Numerical Weather Prediction Satellite Applications Facility (NWP-SAF). The  $O_3$  profiles are retrieved from the 980–1100  $\text{cm}^{-1}$  spectral window encompassing the 9.6  $\mu\text{m}$   $O_3$  absorption band. Only cloud-free or weakly contaminated pixels are processed. Pixels with Advanced Very High Resolution Radiometer (AVHRR)-derived fractional cloud cover larger than 25 % are excluded. We also use a test based on brightness temperatures at 11 and 12  $\mu\text{m}$  when AVHRR cloud cover is not available as described in Barret et al. (2011). The two SOFRID- $O_3$  versions that are validated and compared in the present paper have significant differences that are described below.

## 2.1 Single a priori profile: v1.6

SOFRID-O3 v1.6 is almost similar to v1.5 described in Barret et al. (2011). It is based on RTTOV v9.3 (Saunders et al., 1999). In RTTOV, the optical depths are expressed as a linear combination of profile-dependent predictors that are functions of temperature, absorber amount, pressure and viewing angle. In RTTOV v9.3, the regression coefficients are derived from computations with the line-by-line radiative transfer model v11.6 (LBLRTM; Clough et al., 2005) on 43 atmospheric levels using the HITRAN2004 spectroscopic database (Rothman and Jacquemart, 2005). The single difference is that v1.6 uses temperature and humidity profiles from ECMWF operational analyses for the RTTOV simulations and v1.5 was using IASI L2 products delivered by EUMETSAT. The change has been operated for availability problems and mostly because the EUMETSAT L2 products are not homogeneous over the whole 2008–2017 period, which could result in retrieval inconsistencies (Boynard et al., 2018). We use 6-hourly ECMWF analyses which are provided on 91 (137) vertical levels until (after) 24 June 2013 from the ground up to 0.02 hPa on a  $0.25^\circ \times 0.25^\circ$  horizontal grid. The ECMWF temperature and humidity profiles are interpolated to the time and location of the target IASI pixel with a 3-D linear interpolation scheme.

O<sub>3</sub> concentrations are retrieved on the 43 RTTOV levels with the NWP-SAF 1D-Var algorithm (Pavelin et al., 2008) based on the optimal estimation method (OEM) (Rodgers, 2000). The OEM is a Bayesian method where the incomplete information provided by the measurement is complemented by a priori information which is supposed to represent the best knowledge of the state vector at the moment of the measurement. In our case, the state vector is the O<sub>3</sub> profile. For both v1.5 and v1.6, we use a single O<sub>3</sub> a priori profile which is based on 2 years (2008–2009) of WOUDC and Measurement of Ozone on Airbus In-Service Aircraft – In-Service Aircraft for a Global Observing System (MOZAIC-IAGOS) profiles completed to the top of the RTTOV v9.3 model (0.1 hPa) by MLS-averaged profiles (see Barret et al., 2011 for details).

## 2.2 Dynamical a priori profile: v3.5

As v1.6, SOFRID-O3 v3.5 uses interpolated temperature and humidity profiles from ECMWF analyses. It is based on the more recent RTTOV (v11.1) (Hocking et al., 2015), where regression coefficients are derived from LBLRTM v12.2 computations on 101 vertical levels with the HITRAN2008 spectroscopic database (Rothman et al., 2009). The second and more important one is that it uses dynamical a priori profiles from TpO<sub>3</sub>, the O<sub>3</sub> profile tropopause-based climatology of Sofieva et al. (2014). This climatology is based on ozone profiles resulting from merging ozonesonde data in the troposphere and SAGE II v6.2 data (Wang et al., 2006) in the stratosphere. The ozonesonde profiles (36 000) extracted

from the Binary Database of Profiles (BDBP) come from 136 stations for the period 1980 to 2006 (Hassler et al., 2008). For each merged ozonesonde–SAGE II profile, the tropopause was computed according to the World Meteorological Organization (WMO) definition of the lapse-rate tropopause (WMO, 1957). For each month, the ozone profiles are gathered according to 10° latitude bins and 1 km tropopause intervals, and the corresponding averaged profiles together with their 1 $\sigma$  variabilities are computed and provided. Variable a priori profiles have already been used for satellite sensor retrievals. For instance, Tropospheric Emission Spectrometer (TES) O<sub>3</sub> retrievals used monthly mean profiles from the Model for Ozone and Related chemical Tracers (MOZART) chemistry–transport model (CTM) averaged over a 10° latitude  $\times$  60° longitude grid (Bowman et al., 2006). OMI O<sub>3</sub> a priori profiles are based on a monthly and latitude-dependent ozone profile climatology (McPeters et al., 2007) derived from ozonesonde and satellite data (Liu et al., 2010). Nevertheless, the use of an a priori profile simply based on the geographical location of the satellite pixel does not allow taking the atmospheric dynamics into account. For instance, at a midlatitude location, the O<sub>3</sub> profile can be typical of midlatitudes on one day and polar (low tropopause) or tropical (high tropopause) a few days later depending on the global atmospheric dynamics (position of the polar or subtropical jets, anticyclones). The use of a tropopause-dependent climatology allows us to take the atmospheric dynamics into account and provides a more accurate a priori O<sub>3</sub> profile. This technique was once used for O<sub>3</sub> total column retrievals from Fourier-transform infrared spectroscopy (FTIR) spectra at the Jungfraujoch station (De Maziere et al., 1999). It was shown that the retrieved O<sub>3</sub> columns were largely improved when the tropopause was taken into account in the choice of the a priori profile. In a first attempt to take the tropopause into account for satellite retrievals, Sellitto et al. (2013) have implemented two a priori profiles in the KOPRAFIT-O3 retrieval algorithm to basically discriminate the tropics (tropopause higher than 14 km) from other latitudes. Dufour et al. (2015) have slightly improved the approach with a set of three a priori profiles for high latitudes (tropopause lower than 10 km), midlatitudes (tropopause between 10 and 14 km) and the tropics (tropopause higher than 14 km). Eremenko et al. (2019) have tested a set of profiles for retrievals on a synthetic database. In SOFRID-O3 v3.5, we compute the tropopause using the WMO lapse-rate definition from the ECMWF interpolated temperature profiles. The a priori profile is then picked up from the TpO<sub>3</sub> climatology according to month, latitude and tropopause height.

## 2.3 Information content and retrieval error

A remote sensing instrument is not equally sensitive to the different atmospheric layers. Its vertical sensitivity depends on its instrumental characteristics and on local parameters.

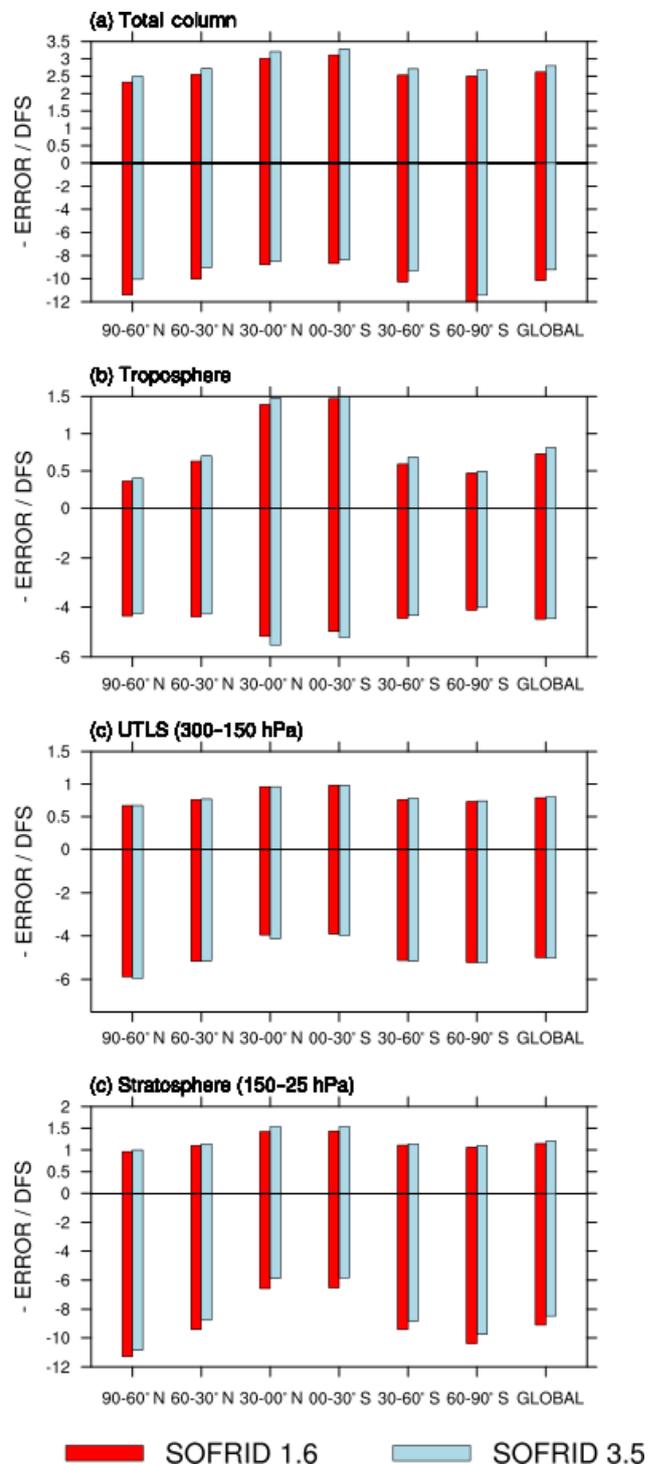
In the case of a thermal infrared nadir sounder such as IASI, surface parameters such as surface emissivity, surface temperature, thermal contrast between the surface and the first atmospheric layer are key parameters to determine the vertical sensitivity, especially in the lower troposphere (Barret et al., 2005; Boynard et al., 2016). The vertical sensitivity of a remote sensing instrument is characterized by the so-called averaging kernel (AK) matrix. For each retrieval layer, the retrieved quantity is the result of the convolution of the whole real profile by the corresponding averaging kernel (row of the AK matrix) plus a contribution from the a priori profile ( $x_a$ ) and a noise ( $\epsilon$ ) contribution (see Eq. 1).

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

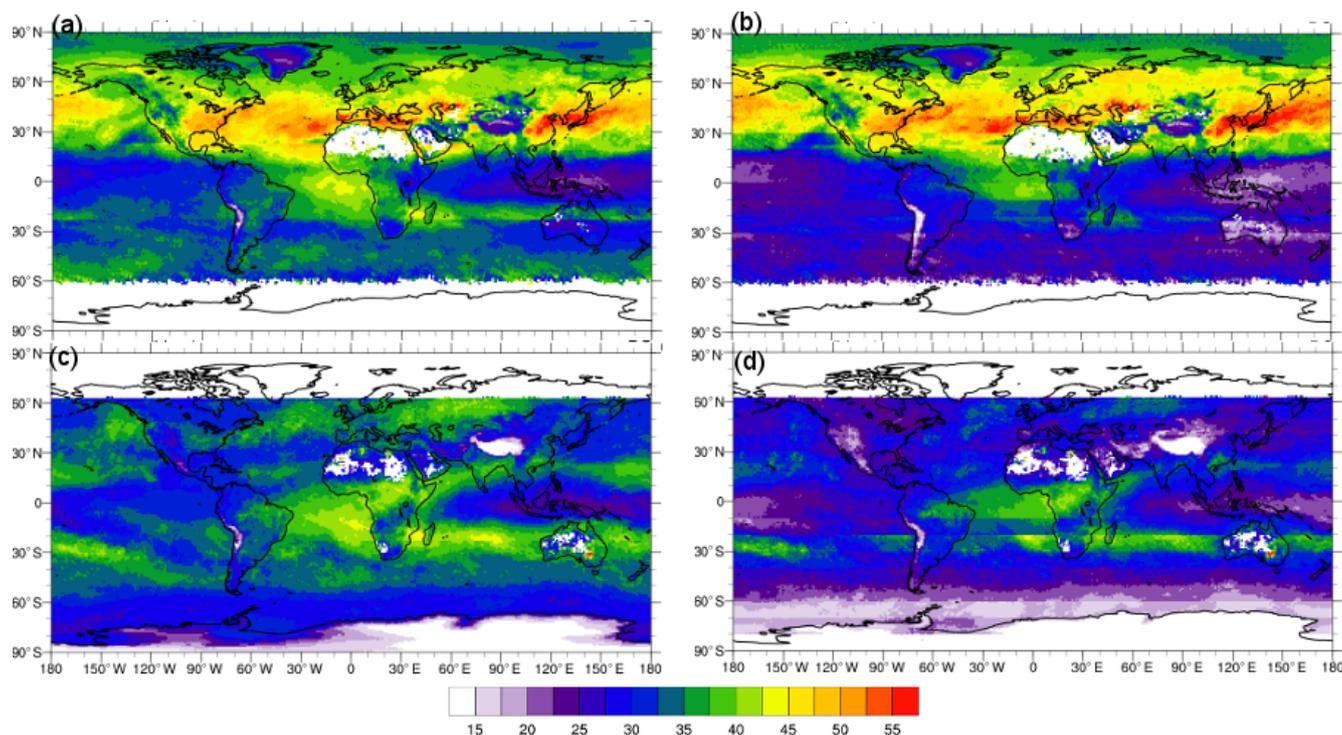
In an ideal case, the AK matrix ( $A$ ) would be the identity matrix ( $I$ ) and real ( $x$ ) and retrieved ( $\hat{x}$ ) profiles would be identical within the noise level ( $\epsilon$ ) contribution.  $G$  is the gain matrix that represents the sensitivity of the retrieval to the measurement. In a real case, the AKs are bell-shaped functions which peak at an altitude that could be different from the nominal altitude and whose width gives an indication of the retrieval vertical resolution.

The degree of freedom for signal (DFS) of a retrieval describing the number of independent pieces of information provided by the measurement is the trace of the AK matrix (Rodgers, 2000). We have divided the atmosphere in five layers which are described in Table 1. The troposphere-2 layer has been selected for comparison with Boynard et al. (2018), who did not compute a tropopause-based TOC for their validation (see Sect. 5). The DFS corresponding to these different layers is displayed in Fig. 1 for v1.6 and v3.5 averaged over the validation dataset. The total DFS ranges from 2.4 to 3.3 for v3.5 and is about 0.2 lower for v1.6. The DFS values for the troposphere (WMO lapse rate), UTLS and stratosphere are almost identical for both versions. The tropospheric DFS is the lowest (0.3–0.5) at high latitudes where surface temperature, thermal contrast and tropopause height are the lowest and the highest in the tropics (about 1.5) where surface temperature and tropopause height are the highest. At midlatitudes, the tropospheric DFS is about 0.6. Therefore, except in the tropics, SOFRID retrievals provide less than one independent piece of information in the troposphere. In the UTLS (stratosphere), the DFS values range from 0.7 to 1 (from 0.9 to 1.5), which means that SOFRID provides around one independent piece of information in these layers.

The retrieval error is the sum of the measurement and smoothing errors (Rodgers, 2000). Uncertainties in auxiliary parameters (temperature and humidity profiles, surface properties, etc.) are also responsible for errors. Coheur et al. (2005) and Barret et al. (2005) have shown that in the case of  $O_3$  and  $CO$  retrievals from thermal infrared satellite sensors, the dominant source of errors was the smoothing error. The retrieval errors for SOFRID- $O_3$  v1.6 and v3.5 are displayed in Fig. 1. Here, v1.6 displays slightly larger errors than v3.5 but has the same behavior. For the total and stratospheric



**Figure 1.** DFS and (–) retrieval errors in Dobson units (DU) for SOFRID- $O_3$  v1.6 (red) and v3.5 (light blue) retrievals for (a) total column, (b) troposphere, (c) UTLS (300–150 hPa) and (d) stratosphere (150–25 hPa).



**Figure 2.** TOC distributions in DU for (a) July 2017 v1.6, (b) July 2017 v3.5, (c) December 2017 v1.6 and (d) December 2017 v3.5.

**Table 1.** Atmospheric layers for comparisons.

Layer	Lower boundary	Upper boundary
Troposphere-1	Ground	Tropopause
Troposphere-2	Ground	300 hPa
Lower troposphere	Ground	550 hPa
UTLS	300 hPa	150 hPa
Stratosphere	150 hPa	25 hPa

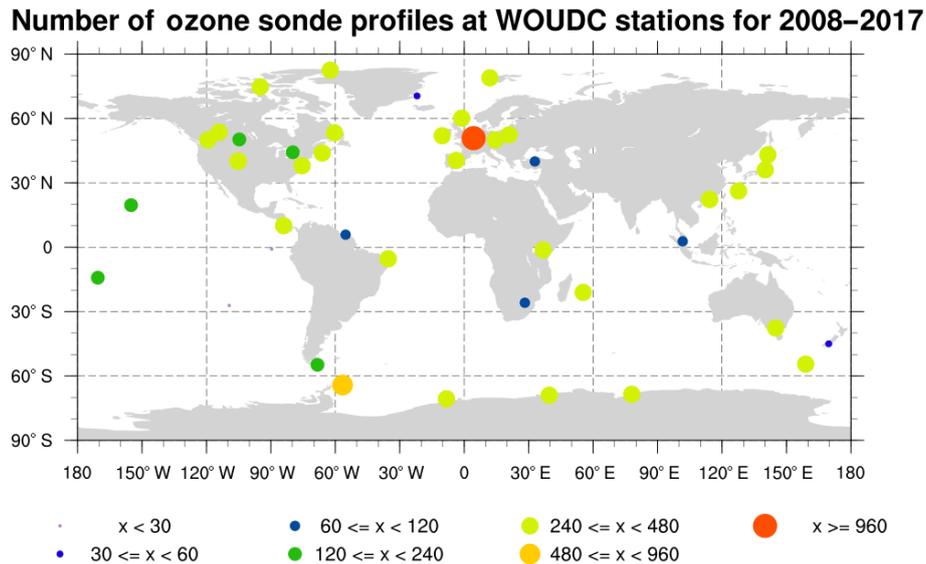
columns, the errors decrease from high latitudes (9–12 DU) to the tropics (6–8 DU). The behavior of UTLS errors is similar with lower values (4 to 6 DU). For the TOC, errors are larger in the tropics (5 DU) than at middle and high latitudes (4 DU). This is due to the fact that the tropopause height is higher in the tropics, resulting in a larger a priori variability. The impact of the increased variability exceeds the one of the increased information content, resulting in a larger smoothing error.

## 2.4 Global distributions of tropospheric ozone columns

The global distributions of TOC from SOFRID v1.6 and v3.5 for July and December 2017 are displayed in Fig. 2. The global TOC structures are similar for both versions. They both clearly show the highest TOC over the NH midlatitudes in summer with a large export region over the northern Pacific off the Chinese coast and the summertime TOC max-

imum over the Eastern Mediterranean already documented with the Global Ozone Monitoring Experiment-2 (GOME-2) sensor (Richards et al., 2013). The tropical Wave-one pattern (Thompson et al., 2003; Sauvage et al., 2006) with the highest TOC over the tropical Atlantic and the lowest one over the South Pacific Convergence Zone (SPCZ) is also noticeable for both versions. Sauvage et al. (2006) have shown that the tropical Atlantic maximum was mostly a result of African and South American lightning  $\text{NO}_x$  ( $\text{LiNO}_x$ ) emissions. High TOCs are also detected during austral summer over southern Africa and the southern Indian Ocean towards Australia. According to Zhang et al. (2012), these high TOCs are mostly caused by  $\text{LiNO}_x$  emissions from central Africa with a yearly maximum in May. The clearest difference between both versions is that v3.5 produces lower TOC than v1.6 in the low tropospheric  $\text{O}_3$  regions. This is clear over the Intertropical Convergence Zone (ITCZ) and the SPCZ, over the SH for both seasons and over the NH midlatitudes in winter. We will show in the validation part of the paper that this is an important improvement of the SOFRID- $\text{O}_3$  retrievals. The agreement is better in regions of high TOC such as NH midlatitudes in summer or the tropical Atlantic.

The use of a dynamical a priori profile is responsible for visible stripes along the  $10^\circ$  latitude bands. These stripes are generally indicating a discontinuity of 2.5 to 5 DU between two adjacent latitude bands with different a priori profiles. They are clearly caused by the impact of the a priori profile on the retrieval which is taken into account in the retrieval



**Figure 3.** Maps of WOUDC stations with ECC O<sub>3</sub> sonde data during the 2008–2017 period. Colors and sizes of the markers indicate the number of valid sondes at each station.

error (see Eq. 1). The latitudinal discontinuities are therefore consistent with our retrieval errors (4–5 DU) from Fig. 1. Such stripes may appear as a problem for the use of SOFRID v3.5 data for model validation. They are a minor problem for two main reasons. First, as is demonstrated in Sect. 4, the use of a dynamical a priori profile largely improves the retrieved O<sub>3</sub> profiles. Second, when model profiles are compared to SOFRID retrievals, the impact of the a priori profile is taken into account by using Eq. (1) such as in Barret et al. (2016).

### 3 Validation methodology

#### 3.1 Ozonesonde data

Ozonesonde data come from the WOUDC database (<https://www.woudc.org/>, last access: 29 september 2020). For consistency purposes, we have chosen to use data from electrochemical concentration cell (ECC) sondes only. For the 10 years (the IASI period; 2008–2017), valid comparisons were effective for about 12 000 ozonesonde profiles among the 16 000 downloaded. A map with the number of sondes used for the validation at each station over the 2008–2017 period is displayed in Fig. 3. Most (~ 7000) of the validation sondes were launched in the NH midlatitudes, with 15 stations providing more than one profile per month on average (more than 120 profiles for 10 years) mostly in western Europe and North America. For all other 30° latitude bands, the number of validation profiles ranges from 800 to 1200, with only three to four stations providing more than 120 profiles. The balloons that carry the ozonesondes often explode below 40 km. In order to complete the ozonesonde profiles in the upper stratosphere and mesosphere, we have used MLS

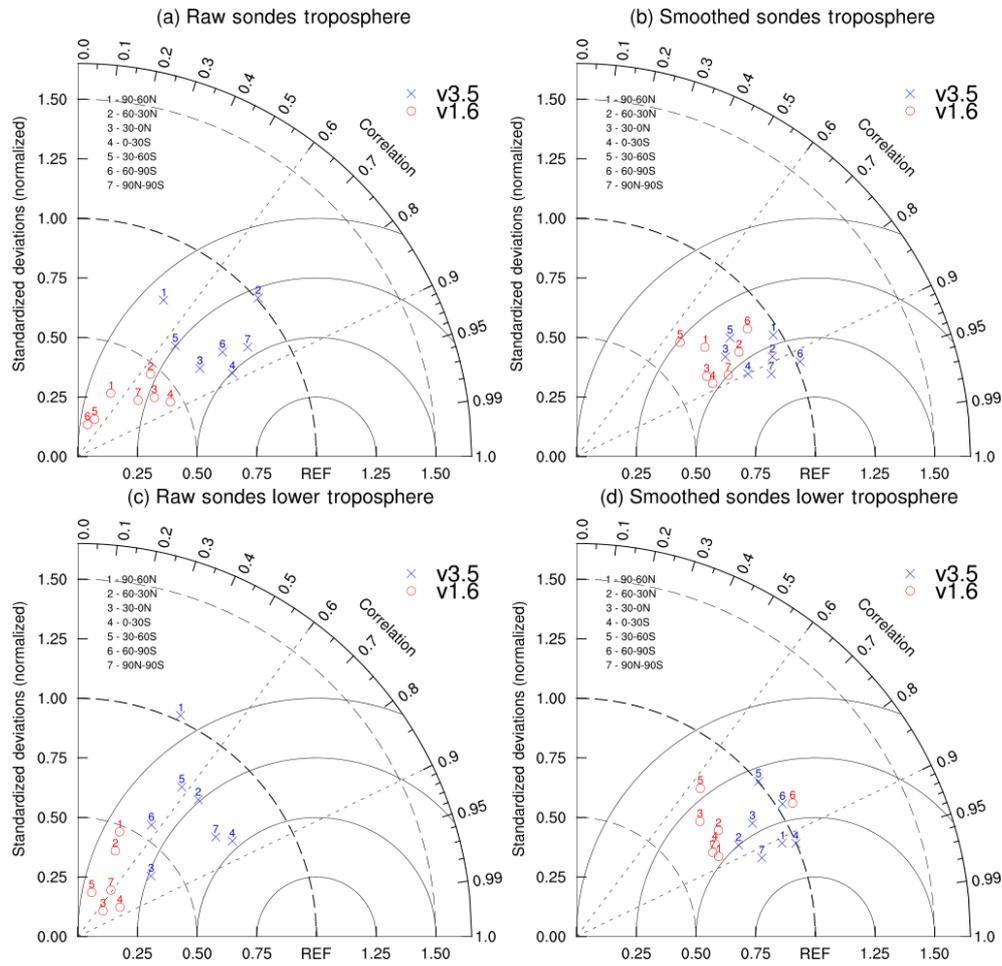
data averaged over 10 d on a 10° × 10° grid (see Barret et al., 2011 for details).

#### 3.2 Coincidence criteria

The spatiotemporal coincidence criteria are ±1° latitude, ±1° longitude and ±12 h. They are similar to those used in Barret et al. (2011), Boynard et al. (2016) (50 km ± 10 h), Boynard et al. (2018) (100 km, ±6 h), Dufour et al. (2012) (110 km, ±7 h). As we compare sondes with IASI morning data only, and since most of the sonde launches are performed in the morning, using 6 or 12 h coincidence does not introduce significant differences. We have computed statistics for nine latitude bands which are the whole globe, the two hemispheres and six 30° wide latitude bands. For each band, the monthly mean is computed if there are at least four coincident profiles within this latitude band. We first keep pixels for which convergence is achieved. Convergence is based on the value of the retrieval cost function output from the 1D-Var analysis (Jcost) which has to be positive, the value of its normalized gradient and the evolution of Jcost between the two last iterations (Havemann, 2020). We have also set an upper limit (1.0) for Jcost in order to eliminate pixels with poor-quality fits. Thirdly, only pixels with a total DFS > 2.0 are selected. Using these criteria, we have kept about  $9.0 \times 10^5$  pixels out of  $1.1 \times 10^6$ .

#### 3.3 Comparison with raw and smoothed data

To compare remote-sensed to in situ or modeled profiles, it is important to apply Eq. (1) to the in situ or simulated profile (Rodgers, 2000; Barret et al., 2002). This procedure allows



**Figure 4.** Taylor diagrams for (a, b) tropospheric columns and (c, d) lower tropospheric columns. (a, c) Raw sonde data, (b, d) smoothed sonde data. Red circles indicate v1.6; blue crosses indicate v3.5.

us to check the quality of the retrieval taking its degraded vertical resolution and sensitivity into account.

Nevertheless, in a validation objective, it is also necessary to compare the retrieved profiles to raw (not smoothed by the AKs) in situ profiles in order to perform a fully informative validation. This is of particular importance when the satellite data are used for issues such as the ozone seasonal to interannual variabilities (Wespes et al., 2017; Peiro et al., 2018) or to document the long-term tropospheric ozone tendencies (Gaudel et al., 2018; Wespes et al., 2018; Dufour et al., 2018). Indeed, the application of Eq. (1) implies the mixing of information between the different layers. Therefore, the variabilities and the drifts computed from raw and smoothed sonde data may be different and need to be documented. Raw ozone sonde data have been compared to IASI retrievals in few studies at the beginning of the IASI period (Barret et al., 2011; Dufour et al., 2012) but have been disregarded in more recent validation work (Boynard et al., 2016, 2018). The importance of raw data validation regarding sea-

sonal and interannual variabilities and trends analyses will be highlighted in detail in Sect. 4.

### 3.4 Taylor diagram

In order to validate remote sensing with reference in situ observations, we need to determine how well they are able to reproduce the same behavior. There are four statistical indicators that have to be computed: (i) the absolute difference or bias which documents the accuracy, (ii) the root mean square of the differences (RMSDs) which tell whether the bias is significant or not, (iii) the coefficient of correlation ( $R$ ) which documents the consistency and phase of the variabilities of both datasets and (iv) the ratio of the standard deviations of both datasets which documents the goodness of the amplitude of the retrieval variability. In the case of IASI  $O_3$ , the first three indicators are frequently computed (Boynard et al., 2016, 2018; Barret et al., 2011), but the last one is rarely compared (Dufour et al., 2012), which makes most validation exercises incomplete.

Based on the relationship between correlation coefficients, RMSDs and variances of the reference (validating) and test (validated) datasets, Taylor (2001) has developed the Taylor diagram initially for climate model evaluation. It displays all of these parameters (except the biases) in a more convenient and synthetic way than tables with numbers. Each experiment or observation to be validated correspond to a point placed within a quarter circle. The reference is located in the middle of the  $x$  axis (see Figs. 4, 5). The correlation coefficient between the reference and test dataset is given by the azimuthal position of the point. The RMSD is proportional to the distance between the test and the reference point. Finally, the radial distance from the origin is proportional to the variance of the experiment. We have normalized both RMSDs and standard deviations by the standard deviation of the reference to display the results from multiple experiments on a single diagram (see Taylor, 2001 for details).

## 4 Validation results

### 4.1 General statistics for tropospheric, UTLS and stratospheric partial columns

For the different latitude bands, the statistics from the comparisons between ozonesondes and SOFRID data are presented in Table 2 for the biases and corresponding RMSDs. Taylor diagrams are displayed in Fig. 4 for the TOC and lower tropospheric columns and in Fig. 5 for the UTLS and stratospheric columns.

Concerning the troposphere, the comparison between SOFRID and raw sonde clearly shows the improvement from v1.6 to v3.5 (Fig. 4a). Here, v3.5 displays a larger variability in better agreement with the raw sonde data with a ratio between SOFRID and sonde variances ranging from 0.62 to 1.01. For v1.6, this ratio ranges from 0.15 to 0.45. The RMSDs of the SOFRID versus raw sonde data are lower and the correlation coefficients larger for v3.5 than for v1.6. Tropospheric biases are smaller than 10% with the noticeable exception of midlatitudes and high latitudes of the SH for v1.6 and raw sonde data with significant biases of 29% and 55%, respectively (Table 2). This problem of SOFRID v1.6 retrievals in the SH had already been diagnosed by Dufour et al. (2012) and by Emili et al. (2014). The use of a dynamical a priori profile in v3.5 allows us to reduce these large biases to almost zero.

As expected, when the sonde profiles are smoothed with SOFRID AKs (Fig. 4b and d), the agreement between sonde data and SOFRID retrievals is better. The retrieval variabilities are closer to the sonde variabilities, the RMSDs are smaller, and the correlation coefficients are higher. It is also noticeable that differences between both retrieval versions are less important and that the improvement of v3.5 relative to v1.6 is less evident. Furthermore, the large v1.6 biases in the SH troposphere at midlatitudes and high latitudes are re-

duced below 10% when the impact of the a priori profile is taken into account with Eq. (1), hiding the problem.

The lower tropospheric retrieved columns agree less with raw sonde data with degraded correlation coefficients and larger RMSDs (Fig. 4c) compared to the TOCs. For raw sonde data comparisons, the lower tropospheric variability is better for v3.5 than for v1.6. When the sondes are smoothed, the statistics are much better and similar to the TOC results (Fig. 4d). The added value of lower tropospheric columns relative to TOCs is therefore not obvious for SOFRID-O3.

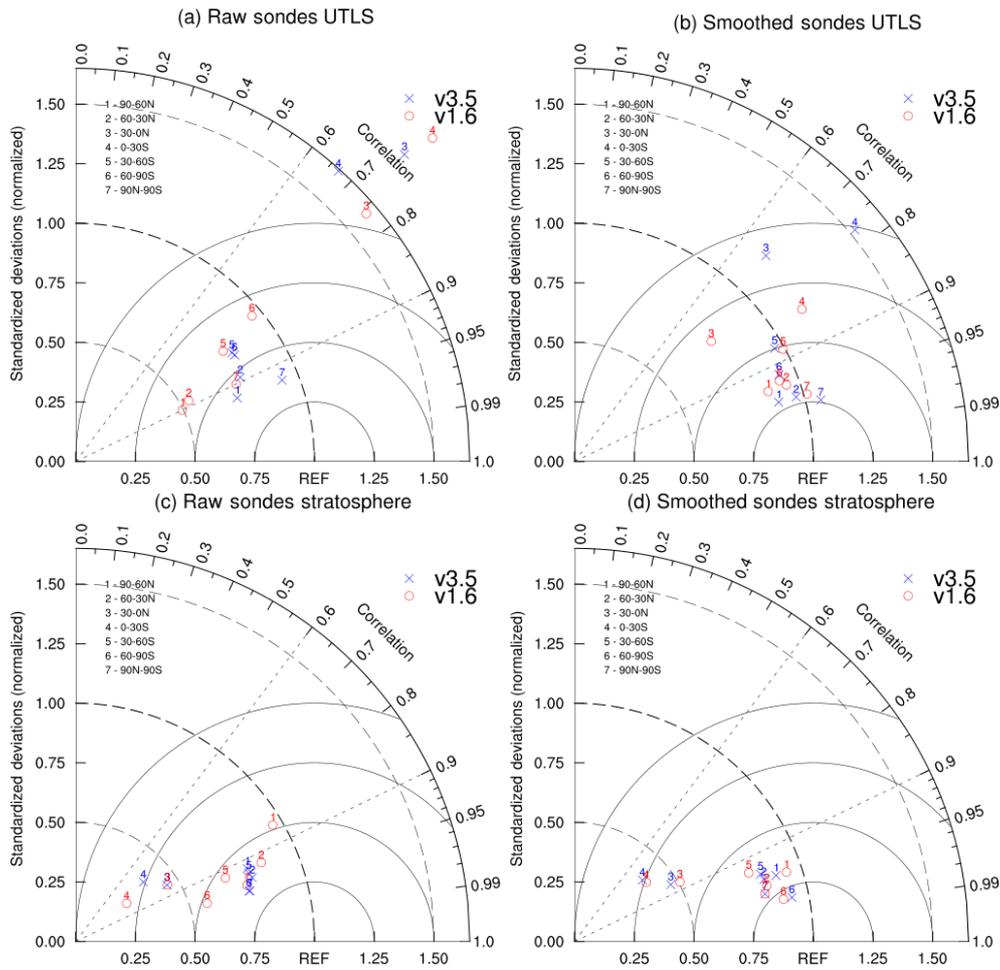
In the UTLS, both v1.6 and v3.5 are in good agreement with raw sonde data (Fig. 5a) and the differences between both versions are much lower than for the tropospheric columns. Correlation coefficients range from 0.67 to 0.93, and the ratios between retrieved and raw sonde variances range from 0.5 to 1.0 at midlatitudes and high latitudes. For the northern and southern tropical latitudes, the correlations coefficients range from 0.6 to 0.75, and the variance ratios are between 1.6 and 2.1, highlighting a too-high variability retrieved in the tropical UTLS. In the UTLS, biases are positive (5% to 18%) at high latitudes and midlatitudes, and negative (−3% to −21%) at tropical latitudes and not significant because of large RMSDs.

In the stratosphere, the agreement between raw sonde data and SOFRID retrievals is very good for the two versions as well as in all latitude bands, with correlation coefficients in the 0.75–0.98 range and variance ratios in the 0.56–0.96 range, except in the tropical bands where the retrieved variances are much lower than the ozonesonde variances (Fig. 4c). Stratospheric columns from v3.5 are in slightly better agreement (higher  $R^2$ , lower RMSDs) with ozonesonde data than v1.6. Large positive biases (10%–14%) are found at tropical latitudes for both v1.6 and v3.5 (Table 2).

Both in the UTLS and the stratosphere, the agreement is only slightly improved (larger correlation coefficients and lower RMSDs) when the sonde profiles are smoothed by the AKs (Fig. 5b and d). Smoothing of the sonde profiles does not significantly modify the UTLS and stratospheric biases. In particular, the tropical UTLS large negative biases are still present when the AKs are applied to the sonde data. The small differences between v1.6 and v3.5 on the one hand and between raw and smoothed sonde data on the other hand highlight the larger sensitivity of IASI to the UTLS and the stratosphere than to the troposphere as already discussed in Barret et al. (2011) and Dufour et al. (2012) for SOFRID v1.5.

### 4.2 Impact of the intraseasonal tropopause dependence of the a priori profile on SOFRID improvements

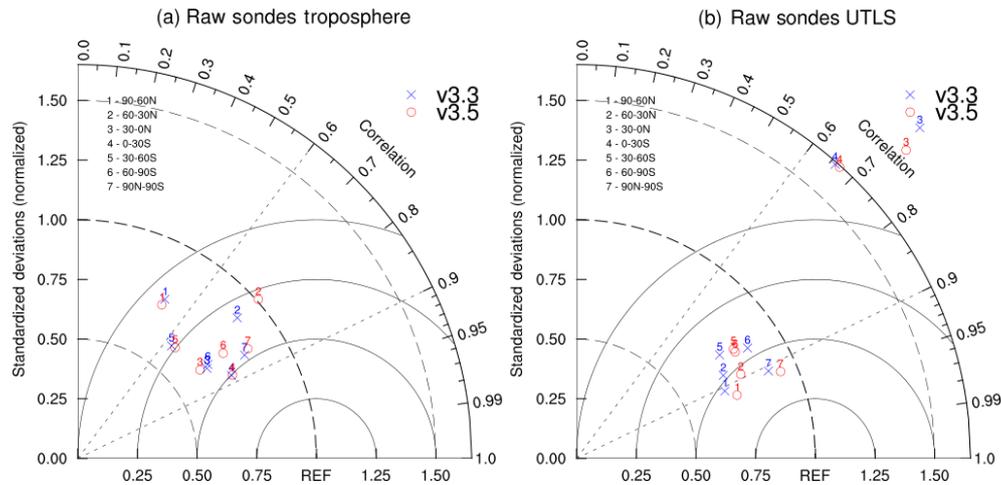
The Sofieva et al. (2014) climatology is tropopause dependent in two different ways. First, it implicitly documents the seasonal and latitudinal relationship between the tropopause and the  $O_3$  profiles with the classification of the  $O_3$  profiles by month and 10° latitude band. Second, it explicitly



**Figure 5.** Taylor diagrams for **(a, b)** UTLS columns and **(c, d)** stratospheric columns. **(a, c)** Raw sonde data, **(b, d)** smoothed sonde data. Red circles indicate v1.6; blue crosses indicate v3.5.

**Table 2.** Biases (%) between sondes and SOFRID retrievals with corresponding RMSDs (%). Values between brackets correspond to smoothed sonde data. Significant biases (bias > RMSD) are in bold text.

Latitude band	SOFRID	Troposphere	UTLS	Stratosphere
90–60° N	v1.6	6 ± 14 (0 ± 6)	6 ± 18 (10 ± 10)	7 ± 10 (4 ± 6)
	v3.5	–2 ± 14 (–1 ± 7)	10 ± 15 (11 ± 10)	1 ± 3 (3 ± 6)
60–30° N	v1.6	2 ± 15 (0 ± 8)	18 ± 27 (13 ± 16)	2 ± 8 (4 ± 6)
	v3.5	–6 ± 14 (–3 ± 9)	17 ± 27 (13 ± 17)	1 ± 7 (3 ± 6)
30–0° N	v1.6	2 ± 17 (4 ± 11)	–3 ± 30 (1 ± 37)	<b>14 ± 8 (12 ± 7)</b>
	v3.5	–3 ± 16 (0 ± 14)	–12 ± 33 (–13 ± 39)	<b>12 ± 8 (12 ± 7)</b>
0–30° S	v1.6	–2 ± 14 (–2 ± 10)	–21 ± 27 (–16 ± 25)	<b>14 ± 10 (10 ± 8)</b>
	v3.5	–8 ± 14 (–7 ± 12)	–21 ± 30 (–24 ± 25)	10 ± 11 (10 ± 11)
30–60° S	v1.6	<b>29 ± 22 (5 ± 9)</b>	11 ± 29 (13 ± 22)	1 ± 8 (4 ± 7)
	v3.5	1 ± 18 (1 ± 13)	10 ± 28 (13 ± 23)	3 ± 7 (4 ± 7)
60–90° S	v1.6	<b>55 ± 25 (7 ± 6)</b>	5 ± 22 (15 ± 13)	7 ± 12 (4 ± 7)
	v3.5	0 ± 16 (1 ± 9)	7 ± 19 (13 ± 13)	6 ± 11 (4 ± 8)



**Figure 6.** Taylor diagrams for (a) tropospheric columns and (b) UTLS columns for comparisons with raw sonde data. Red circles indicate v3.5; blue crosses indicate v3.3.

documents the intraseasonal tropopause  $O_3$  profiles' relationship with various tropopause-dependent profiles for each month and latitude band. In order to determine the impact of the intraseasonal tropopause dependence on the SOFRID retrievals, we have performed retrievals using the profile with the highest occurrence for each month and each  $10^\circ$  latitude band as a priori information. The version with a single monthly a priori profile in each  $10^\circ$  latitude band is v3.3.

The comparisons between v3.5 and v3.3 are presented in Taylor diagrams for the tropospheric and the UTLS columns in Fig. 6. In the stratosphere, the changes (not shown) are negligible. Concerning the TOC, the improvements are negligible except in the  $60\text{--}30^\circ\text{N}$  and  $60\text{--}90^\circ\text{S}$  bands where v3.5 better reproduces the variability of the TOC. In the UTLS, v3.5 gives slightly better results in terms of correlation coefficients and variability relative to v3.3 in all the latitude bands except in the  $60\text{--}90^\circ\text{S}$  one. Nevertheless, the improvement from v3.3 to v3.5 is minor compared to the overall improvement from v1.6 to v3.5 (Figs. 4 and 5). We can therefore conclude that the seasonal (monthly) and latitude dependence of the a priori  $O_3$  profile is responsible for most of SOFRID improvements.

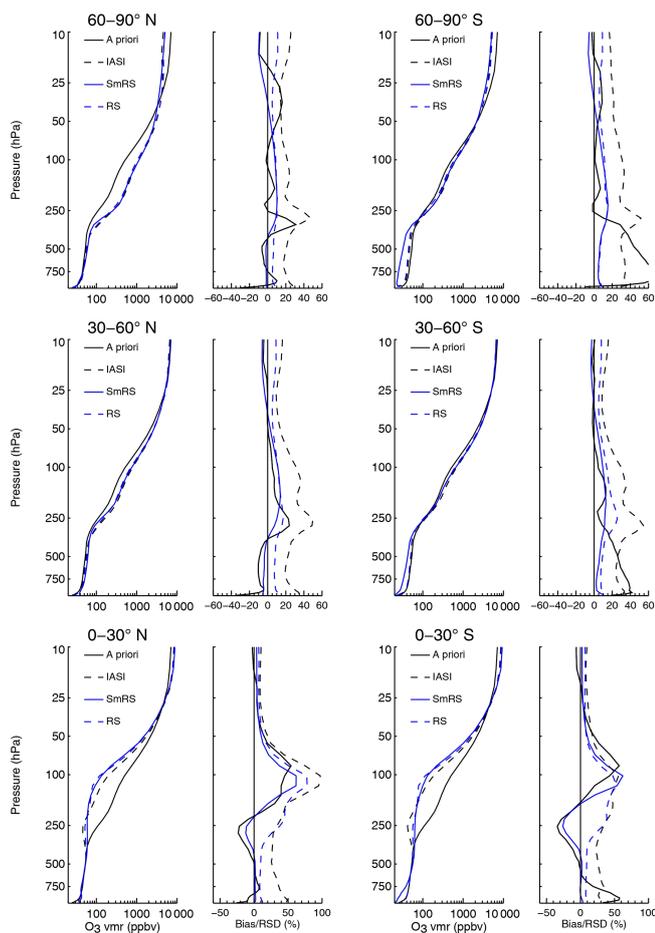
### 4.3 Vertical profiles

After comparing partial columns, it is interesting to look at complete profiles to get better insight about the discrepancies between IASI retrievals and sonde data. The annual average profiles for v1.6 and v3.5 are displayed in Figs. 7 and 8, respectively, for the different latitude bands.

In the NH, v1.6 and v3.5 show similar behavior with a large upper tropospheric positive bias at midlatitudes and high latitudes and a large oscillation from a negative bias at 250 hPa to a large positive bias at 100 hPa in the tropics. These profile features are responsible for the positive (nega-

tive) biases for the midlatitudes and high latitudes (tropics) UTLS columns and for the positive biases for the tropical stratospheric columns (see Table 2). In the SH, the large tropospheric positive biases of SOFRID relative to raw sondes (below 300 hPa in the high latitudes and midlatitudes and below 500 hPa in the tropics) present in v1.6 almost disappear in v3.5. The improvement of SOFRID accuracy in the SH extratropical troposphere is the clearest advantage of using dynamical a priori profiles. In the SH tropics, the TOC difference between v1.6 and v3.5 is not so clear (see Table 2) because the positive bias in the lower troposphere is compensated by a larger negative bias in the upper troposphere in v1.6. As already discussed from column comparisons, it is also noticeable from profile comparisons (Figs. 7 and 8) that the agreement between SOFRID retrievals and smoothed sonde profiles is better than with raw sondes. An important exception are the large UTLS oscillations in both the NH and SH tropics and for both v1.6 and v3.5. Therefore, unlike what was expected, this important discrepancy between retrievals and sonde data does not result from the use of a single a priori profile too far from the real profile. The differences between v3.5 and v1.6 are largely reduced when sondes are smoothed. For instance, the large tropospheric biases for v1.6 in the SH disappear when the smoothing is applied to the sonde profiles.

For all latitude bands, RMSD profiles display the largest values around the tropopause (below 60 % in the extratropics and up to 100 % in the NH tropics), as is expected because it is the altitude range with the largest relative variability. RMSDs between retrievals and smoothed data are generally much lower than with raw data. This is also expected since the smoothing error is the largest source of error in IASI retrievals (see Barret et al., 2011; Dufour et al., 2012). RMSDs with smoothed sondes in the troposphere are somewhat larger for v3.5 than v1.6 especially in the SH. This is an

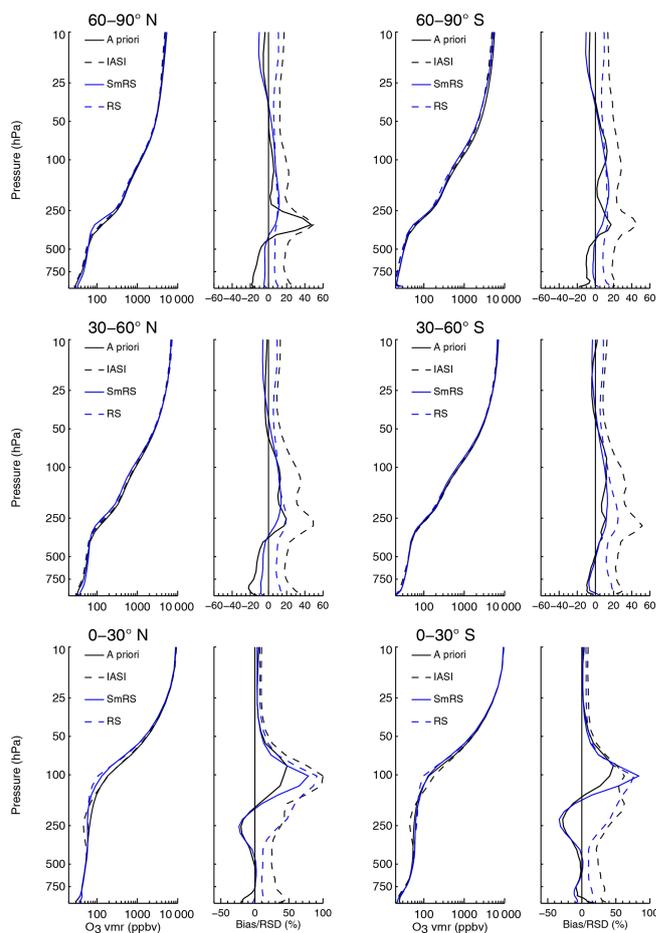


**Figure 7.** Profile comparisons between sonde and SOFRID-O3 v1.6 a priori profiles (left panels) (solid black lines), IASI (dashed black lines), smoothed (SmRS, solid blue lines) and raw (RS, dashed blue lines) sonde vertical profiles, biases (right panels) (solid lines) and RMSD (dashed lines) between IASI and raw (black lines) and smoothed (blue lines) sondes for the NH (left panels) and SH (right panels).

indication of the increased sensitivity and decreased smoothing of v3.5. This is also evident in the Taylor diagrams which show that tropospheric variabilities are larger and in better agreement with sonde data (raw and smoothed) for v3.5 (see Fig. 4).

#### 4.4 Time series of tropospheric columns

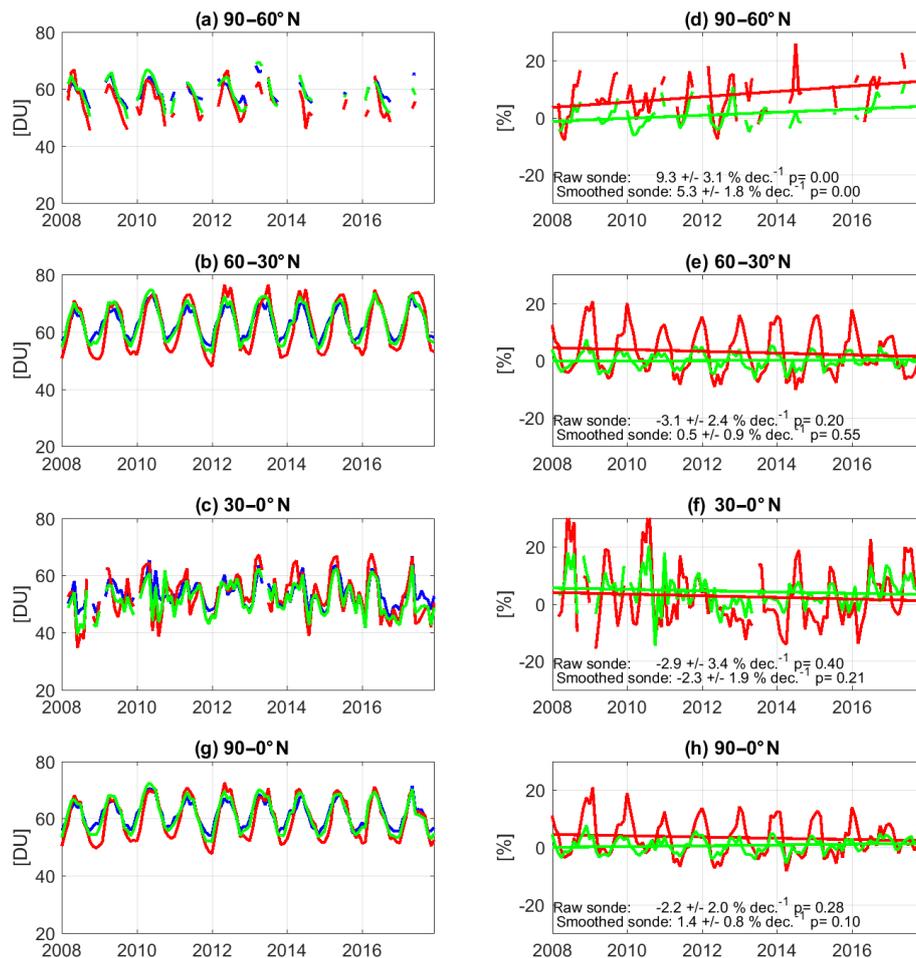
As tropospheric O<sub>3</sub> trend assessment is a major issue and one of the main topic of the TOAR (Tropospheric Ozone Assessment Report)/International Global Atmospheric Chemistry (IGAC) international initiative (Gaudel et al., 2018), we focus in this section on TOC time series. Time series are also interesting to bring insight about the general statistics discussed in the previous sections and to identify possible drifts of the data.



**Figure 8.** Same as Fig. 7 for SOFRID-O3 v3.5.

The time series of IASI and sonde monthly TOCs are presented in Fig. 9 (10) for v1.6 and in Fig. 11 (12) for v3.5 for the Northern (Southern) Hemisphere. We present both raw and smoothed sonde data to highlight the impact of smoothing upon the agreement between IASI and sondes. This impact is particularly obvious for SOFRID v1.6 at midlatitudes. At northern midlatitudes, the bias between SOFRID v1.6 and raw sonde TOCs displays large seasonal variations from  $-5\%$  to  $-10\%$  in summer and  $10\%$  to  $20\%$  in winter, resulting in a negligible  $2\% \pm 15\%$  average bias (Table 2). When sonde data are smoothed by IASI AKs, the sonde variability is largely reduced. Bias varies from  $5\%$  in winter to  $-5\%$  in summer.

For southern midlatitudes, as already highlighted by Dufour et al. (2012) and Emili et al. (2014), SOFRID TOCs are significantly biased high ( $29\% \pm 22\%$ ) relative to raw sonde data (Table 2). This was explained by the fact that the single a priori profile used in v1.6 is biased towards northern midlatitude O<sub>3</sub> (Emili et al., 2014). When the sonde data are smoothed by IASI AKs, the agreement is much better and the bias becomes non-significant ( $5\% \pm 9\%$ ) as a result of taking the a priori contribution into account (Eq. 1). The largest sig-



**Figure 9.** Time series of SOFRID-O3 v1.6 TOCs (DU) in the Northern Hemisphere for (a) 90–60° N, (b) 60–30° N, (c) 30–0° N and (d) 90–0° N. Blue lines indicate IASI retrievals, red lines indicate raw sonde data, and green lines indicate smoothed sonde data. Differences (%) between IASI and sonde data for (e) 90–60° N, (f) 60–30° N, (g) 30–0° N and (h) 90–0° N. Red lines indicate differences with raw sonde data and green lines indicate differences with smoothed sonde data.

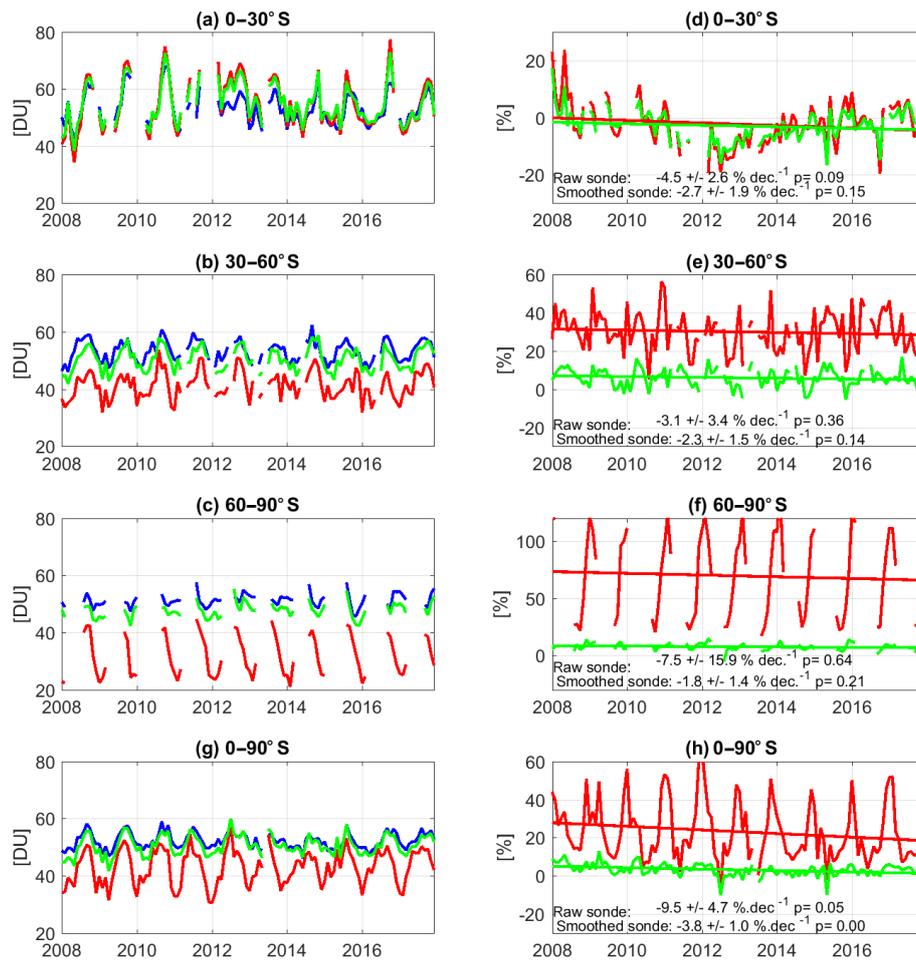
nificant bias ( $56\% \pm 25\%$ ) is found in the SH high latitudes for v1.6 TOCs (Table 2) with large seasonal variations from 20% in winter to 120% in summer. The large bias variabilities at midlatitudes and especially high latitudes of the SH result from the very low seasonal variability of the retrieved columns (see Fig. 4a).

For v3.5, the use of a dynamical a priori profile clearly improves the retrievals at midlatitudes. At northern midlatitudes, the seasonal bias variation is reduced to  $-10\%$ – $0\%$  and the average bias remains small ( $-6\% \pm 14\%$ ). When smoothing is applied, the seasonal variability almost disappears and the bias is only  $-3\% \pm 9\%$ . At southern midlatitudes, the agreement is very good and very similar for raw and smoothed sonde data, with no real seasonal signature detectable and an average bias close to 0%.

At tropical latitudes, the situation is quite different. First, the seasonal variability is not so notable and regular, and the difference between raw and smoothed sondes is lower than at

midlatitudes. Furthermore, the behavior of v1.6 and v3.5 is close even though v3.5 is in better agreement with sonde data (see Sect. 4.1). In the southern tropics, there is a noticeable variation of bias between 2011 and 2014, with large negative biases of  $-10\%$  and  $-15\%$  for 2008–2010 and 2015–2017 with biases of 0 and  $-5\%$  for v1.6 and v3.5, respectively. As such a bias variation is not detected for other latitude bands, we assume that it may be linked to a gap in sonde data for the 2011–2014 period. A closer look at SH tropics' ECC sonde data shows that only two stations (Réunion and Nairobi) provide data regularly (30–50 profiles per year) over the period. For the Pago Pago Pacific station, data are available only from 2014 to 2016 and starting from 2012 for Irene in South Africa. For the Natal Atlantic station, more than 25 profiles are available during the 2008–2010 and 2014–2017 periods and almost none during the 2011–2013 period.

One issue that was raised in TOAR (Gaudel et al., 2018) was the different trends computed from different satellite



**Figure 10.** Same as Fig. 9 for SOFRID-O3 v1.6 in the Southern Hemisphere.

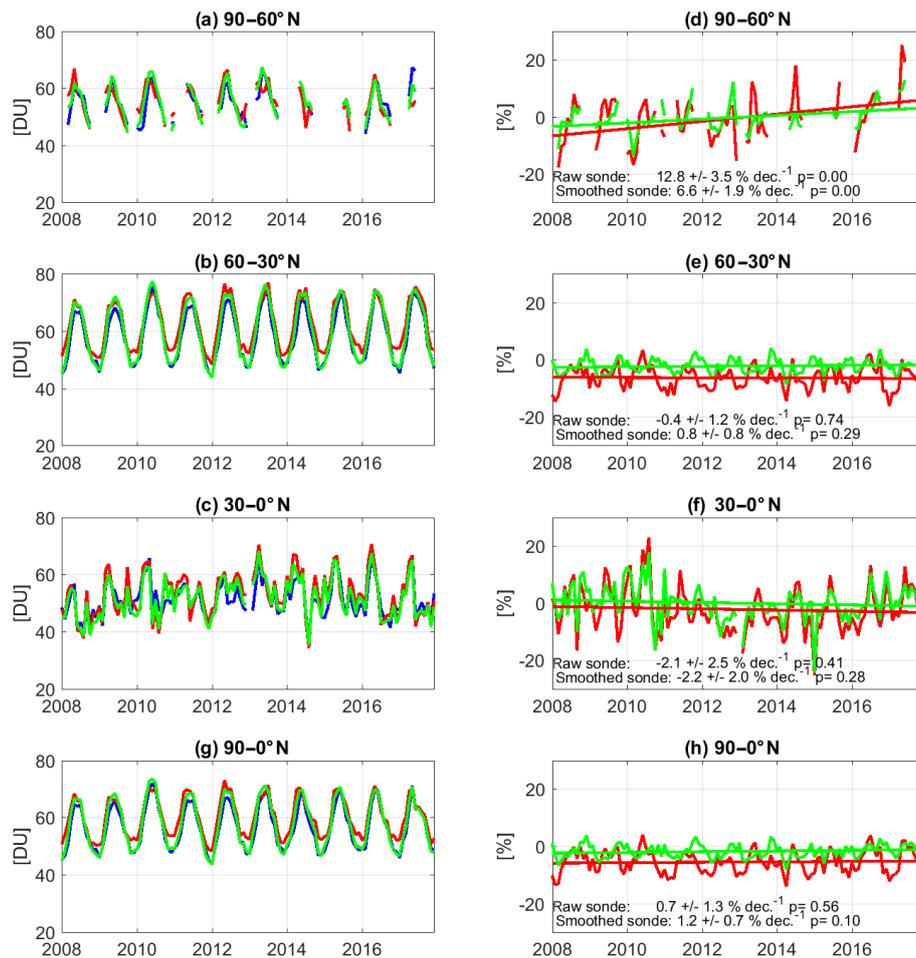
products. UV–visible satellite sensors produce positive tropospheric O<sub>3</sub> burden trends in both hemispheres, while trends from IASI products are negative. It has to be noted that in Gaudel et al. (2018), negative O<sub>3</sub> burden trends from SOFRID v1.5 in the Northern Hemisphere and Southern Hemisphere, and for the whole Earth are, respectively, one-fourth, one-half and one-third smaller than FORLI's. The drifts computed from the SOFRID–sonde differences are displayed in Figs. 9 to 12.

At high northern latitudes, for both v1.6 and v3.5, the drifts are large ( $> 9 \text{ \% decade}^{-1}$  and  $> 5 \text{ \% decade}^{-1}$  for raw and smoothed data, respectively) and significant at the 95 % level. For midlatitudes and tropical latitudes, drifts are between  $0.9 \text{ \% decade}^{-1}$  and  $-3.4 \text{ \% decade}^{-1}$  but are not significant. The NH midlatitude drift with raw sonde data is reduced from  $-3.1$  with v1.6 to  $-0.4 \text{ \% decade}^{-1}$  with v3.5. For the whole NH, the drifts are not significant and decreases from  $-2.2$  with v1.6 to  $0.7 \text{ \% decade}^{-1}$  with v3.5 for raw sonde data. They are  $< 1.5 \text{ \% decade}^{-1} \pm 0.8 \text{ \% decade}^{-1}$  and hardly significant ( $p > 0.10$ ) for smoothed sonde data.

In the SH tropics, drifts are  $\sim -5 \text{ \% decade}^{-1}$  and  $\sim -3 \text{ \% decade}^{-1}$  for raw and smoothed sonde data, respectively, and only significant for v3.5 compared to raw data. These drifts are linked to the large negative biases of the 2011–2014 period resulting from missing data (see above). For v1.6, a large but non-significant drift ( $-8 \text{ \%}$ ) also occurs at high latitudes, which is largely reduced for v3.5. For the whole SH, we found a significant negative drift (relative to raw sonde data) of  $-9.5 \text{ \% decade}^{-1} \pm 4.7 \text{ \% decade}^{-1}$  for v1.6 which is reduced to  $-4.3 \text{ \% decade}^{-1} \pm 1.4 \text{ \% decade}^{-1}$  and becomes non-significant for v3.5.

## 5 Comparison with IASI-FORLI

Two versions of IASI O<sub>3</sub> retrievals with the FORLI software have been validated by Boynard et al. (2016) and Boynard et al. (2018) (B18). Part of their validation results are based on the same data as the present study, namely ECC ozone sondes from the WOUDC database between 2008 and 2014 for Boynard et al. (2016) and 2008 and 2016 for B18. As they document the latest FORLI version (v20151001) on a



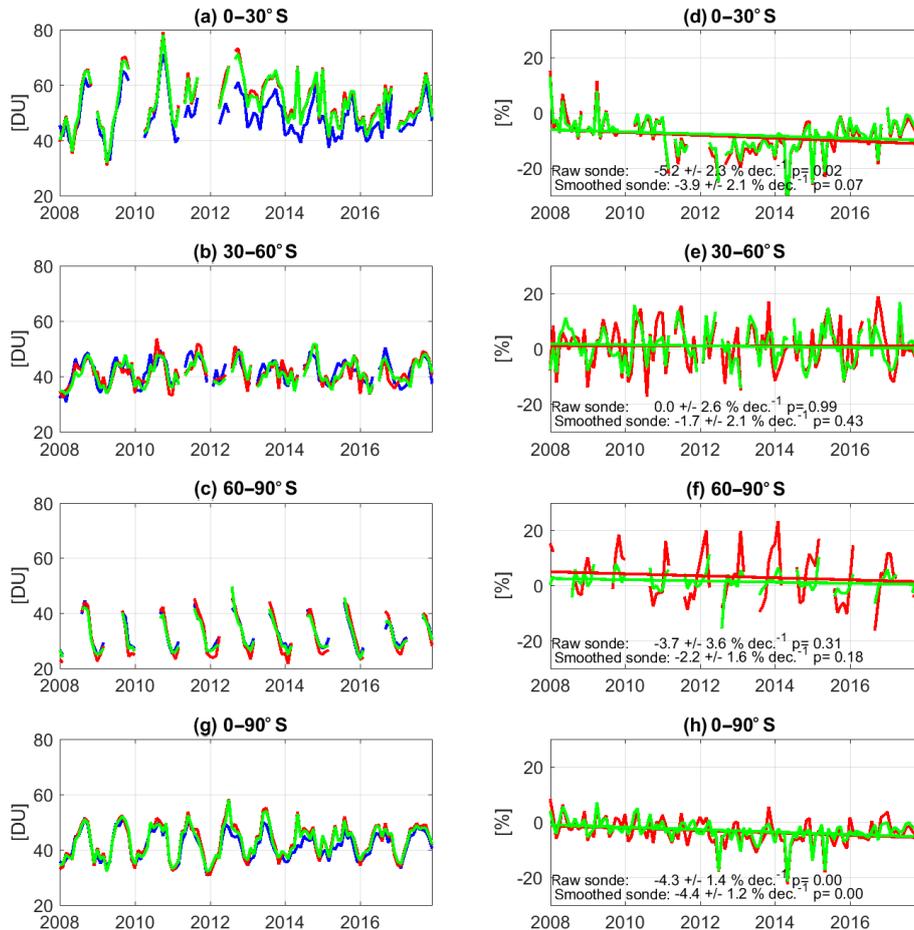
**Figure 11.** Same as Fig. 9 for SOFRID-O3 v3.5 in the Northern Hemisphere.

longer time period, we will focus on our comparison with B18. They have used a comparable number (11 600) of ozone sonde profiles as in the present study, and their comparison methodology is close to the one we have used (spatiotemporal coincidence criteria set to 100 km and  $\pm 6$  h). We have collected the coefficients of determination ( $R^2$ ), the biases, the RMSDs, the DFS of the retrievals and the slopes of the linear fit between the smoothed sondes and retrievals from B18.

There are some limitations to the comparison between the validation of our SOFRID retrievals and the FORLI validation from B18. We are comparing our data with literature results which do not provide the same information as we do. For instance, B18 do not document the sonde and IASI variabilities, and it is therefore not possible to draw their data in Taylor diagrams. B18 have also limited their comparisons to smoothed sonde data. Another limitation is that FORLI and SOFRID use their own quality flags to filter the data. In order to document the impact of the pixel selection on SOFRID validation, we have performed the comparison with sonde data using modified quality flags. The cloud filtering threshold is the clearest source of difference between the pixel se-

lection of both algorithms. We have therefore lowered the upper limit of the AVHRR cloud fraction cover to 13 %, which is the threshold used by B18, resulting in a loss of 5 % of the treated pixels. The Jcost threshold has been decreased from 1.0 to 0.15 with a 6 % decrease of the selected retrieved profiles. Finally, the DFS lower value has been set to 1.75, increasing the number of selected retrievals by 2 %. These threshold modifications resulted in negligible changes of the general statistics (bias, RMSD,  $R$ ) for the three atmospheric layers (troposphere, UTLS and stratosphere) and the different latitude bands that are presented in this section. These statistics, based on large amounts of data, are therefore not hindered by pixel selection differences.

In Fig. 13, we have drawn DFS from SOFRID v1.6 and v3.5 and from FORLI for the layers selected by B18 (1013–300, 300–150 and 150–25 hPa). Figure 14 displays the coefficients of determination ( $R^2$ ) and the slopes ( $b$ ) from linear relationships fitted between IASI retrievals and smoothed sonde data. Biases and RMSDs are shown for the three retrievals in Fig. 15. Finally, Fig. 16 documents the drifts be-



**Figure 12.** Same as Fig. 9 for SOFRID-O3 v3.5 in the Southern Hemisphere.

tween sondes and SOFRID retrievals for the whole NH for the surface–300 hPa layer to be comparable to B18.

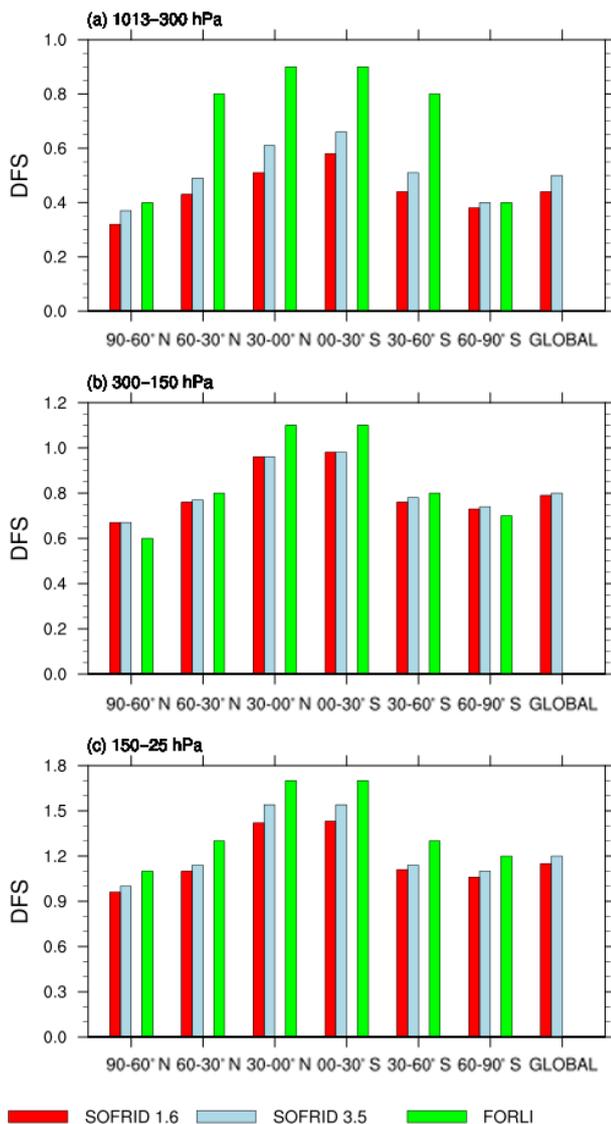
In the three atmospheric layers, the information content is larger with FORLI than with SOFRID v1.6 and v3.5 (Fig. 13). This is particularly visible for the midlatitudes and tropics in the troposphere with DFS of 0.8 to 0.9 for FORLI and DFS of only 0.4 to 0.6 for SOFRID. This probably results from the retrieval noise level which is lower for FORLI than for SOFRID (Dufour et al., 2012). At high latitudes, the DFS values are low and closer for both algorithms, and the increase from high to midlatitudes is therefore much larger for FORLI than for SOFRID. As both algorithms use a single retrieval noise and a priori covariance matrix and similar surface and atmospheric temperatures, the reason for such a difference is unclear. In the UTLS and stratosphere, the same increases of DFS from high latitudes to the tropics are visible for the three products. The difference in information content between retrievals is less pronounced in the UTLS and in the stratosphere than in the troposphere.

The RMSDs (see Fig. 15) are generally larger for FORLI than for SOFRID. In the troposphere, RMSDs reach 18 % for FORLI and are below 10 % for both SOFRID v1.6 and

v3.5. In the UTLS, RMSDs are larger than in the other layers due to the lower absolute columns. For SOFRID, UTLS RMSDs are in the range of 10 %–30 % and 20 %–45 % for FORLI. For both SOFRID and FORLI, the highest RMSDs are in the tropics where the 150–300 hPa columns are the lowest. In the stratosphere, FORLI's RMSDs are also systematically larger than SOFRID's. The differences are the largest at high latitudes with FORLI RMSDs 3 to 4 times larger than SOFRID's.

The  $R^2$  differences (Fig. 14) are partly related to the RMSD differences. Generally, SOFRID has larger  $R^2$  than FORLI. As for the RMSDs, the differences between both algorithms are the largest at high latitudes (especially in the Southern Hemisphere where  $R^2 < 0.4$  for FORLI products) in the three layers. In the troposphere, the coefficients of determination are comparable for both algorithms in the tropical bands, and SOFRID v3.5 gives higher  $R^2$  than SOFRID v1.6. The differences between retrieval versions are generally lower and can even be reversed in the UTLS and in the stratosphere.

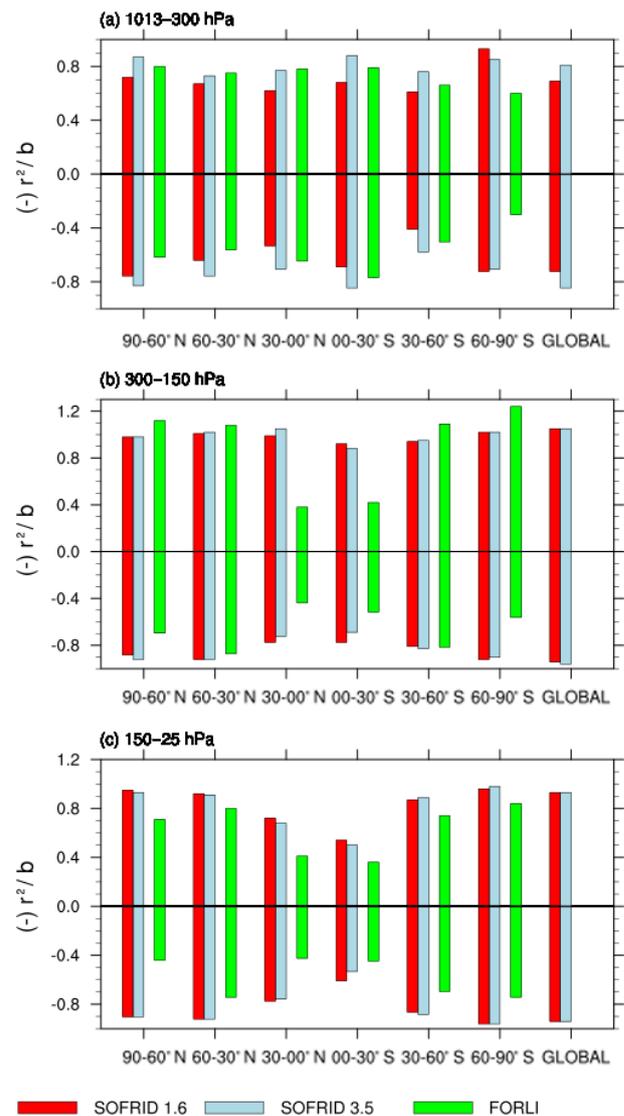
The slopes of the linear fits between retrievals and sonde data provide complementary information to the  $R^2$  coeffi-



**Figure 13.** DFS of IASI SOFRID-O3 v1.6 (red), SOFRID-O3 v3.5 (light blue) and FORLI-O3 (green) retrievals in the different latitude bands for (a) 1013–300 hPa, (b) 300–150 hPa and (c) 150–25 hPa. FORLI data are taken from Boynard et al. (2018).

cients. A slope smaller than 1 indicates that the retrieved variability is too low compared to the reference data, and conversely, a slope larger than 1 indicates an overestimation of the variability. In the troposphere, SOFRID v1.6 and v3.5 and FORLI have similar slopes except in the 60–90° S band where FORLI has a significantly lower slope than SOFRID (Fig. 14).

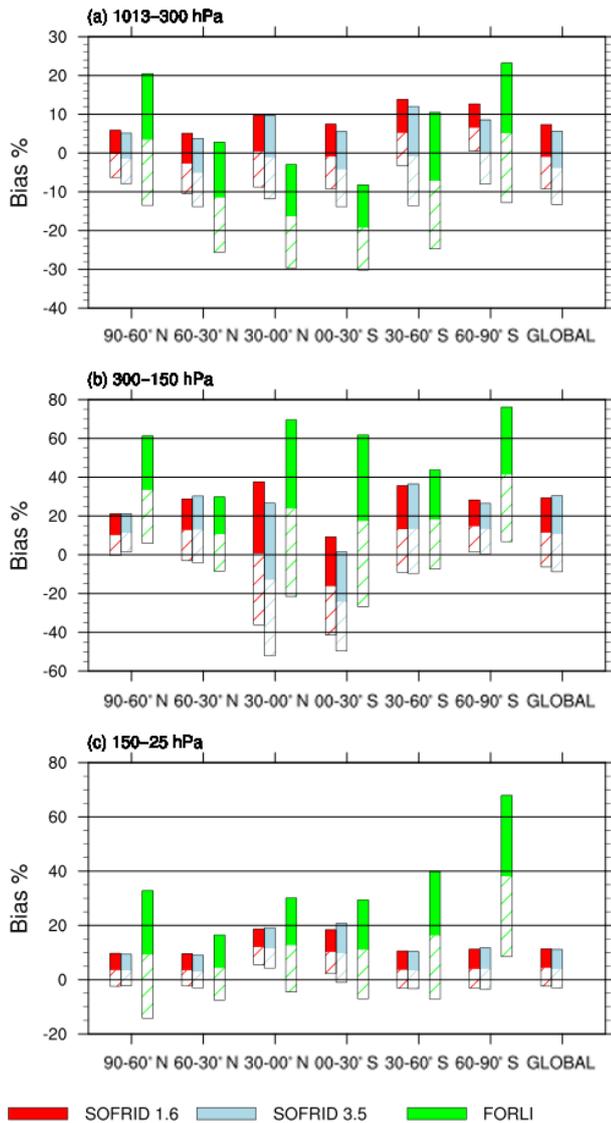
In the troposphere, FORLI products present systematic negative biases from 7 % to 20 % except in the polar regions. Concerning SOFRID, the tropospheric biases are within  $\pm 6\%$  (comparable to TOC biases in Table 2). The results are largely different when the raw sonde data are considered with very large biases in the Southern Hemisphere with SOFRID



**Figure 14.**  $b$  indicates slopes of the linear regression (positive values), and  $(-)$   $R^2$  indicates coefficients of determination (negative values) between IASI retrievals and sonde data. Red indicates SOFRID-O3 v1.6, light blue indicates SOFRID-O3 v3.5, and green indicates FORLI-O3 (from Boynard et al., 2018).

v1.6, as discussed in Sect. 4.4. In the UTLS, SOFRID and FORLI biases are significantly positive except in the tropics, and more specifically in the SH tropics, where SOFRID columns are negatively biased by  $\sim 20\%$ , as discussed in Sect. 4.1 (Table 2). In the stratosphere, both SOFRID and FORLI products are positively biased. The largest differences between both retrieval algorithms are found in the extratropical southern latitudes with FORLI biases larger than SOFRID. In the 60–90° S latitude band, FORLI biases reach about 40 % against about 5 % for SOFRID.

From the perspective of a better quantification of tropospheric O<sub>3</sub> evolution and of the TOAR results (Gaudel et al., 2018), it is also important to compare the drifts be-



**Figure 15.** Biases and  $\pm$ RMSDs (bars) of the differences between IASI retrievals and sonde data. Red indicates SOFRID-O3 v1.6, light blue indicates SOFRID-O3 v3.5, and green indicates FORLI-O3 (from Boynard et al., 2018).

tween sonde and retrievals. B18 present and discuss the drift between FORLI and sonde data for different layers in the whole NH. The SOFRID NH tropospheric drifts discussed in Sect. 4.4 are smaller and opposite in sign to the significant  $-8.6\% \text{ decade}^{-1} \pm 3.4\% \text{ decade}^{-1}$  drift between FORLI and smoothed sonde data in the NH troposphere presented in B18. As B18 computed a surface–300 hPa column instead of a tropospheric column, we have computed the drifts based on the same layer (see Fig. 16). Drifts for surface–300 hPa columns are slightly (0.1 % to 0.4 %) smaller than for TOCs and are not significant in both cases. The comparison of the NH drift with B18 is therefore not dependent on the tropospheric layer definition. For v1.6

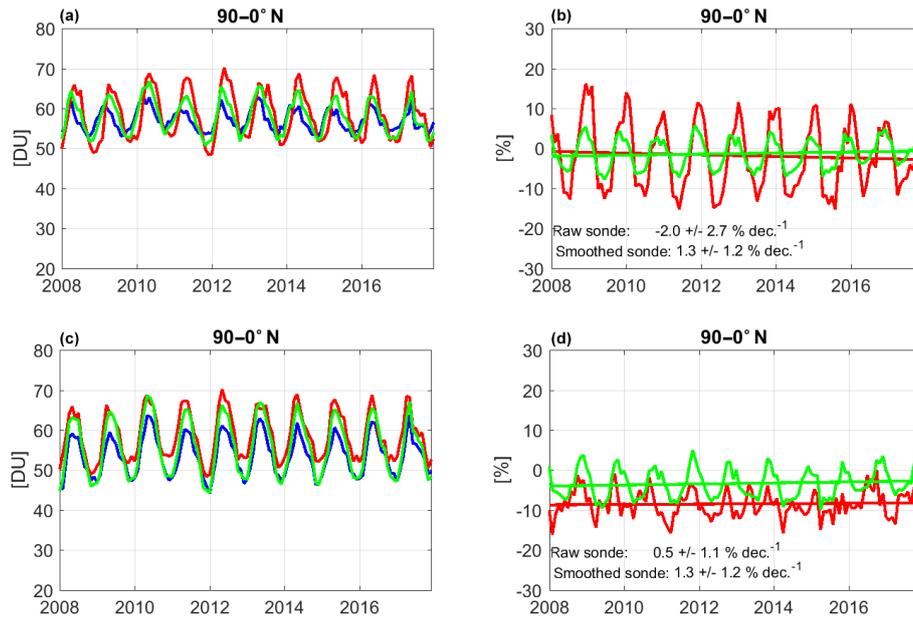
and v3.5, compared to raw and smoothed sonde data, the surface–300 hPa column drifts range from  $-2.0\% \text{ decade}^{-1}$  to  $1.3\% \text{ decade}^{-1}$  (see Fig. 16), values which are much smaller than in B18. Nevertheless, the NH tropospheric drift from FORLI is attributed to an abrupt change or jump detected in 2010 (Boynard et al., 2018; Wespes et al., 2018). Indeed, the drift strongly decreases after the jump and it becomes even non-significant for most of the stations over the periods before or after the jump, separately (Wespes et al., 2018). The discontinuity is suspected to result from updates in level-2 temperature data from EUMETSAT used as inputs into FORLI (Wespes et al., 2019). The absence of a jump and the small drift in SOFRID v1.6 (Fig. 9h) and v3.5 (Fig. 11h) NH tropospheric data are therefore probably linked to the use of temperature profiles from ECMWF analyses instead of EUMETSAT L2 products.

## 6 Conclusions

This study aimed at assessing the quality of two different versions of SOFRID-O3 at the global scale and over the 10-year IASI period using ozonesondes from the WOUDC. SOFRID-O3 v1.6 retrievals are based on a single a priori profile like most other global IASI O<sub>3</sub> retrievals (Barret et al., 2011; Dufour et al., 2012; Boynard et al., 2016, 2018). In v3.5, the a priori profile is dynamically selected from an O<sub>3</sub> profile climatology (Sofieva et al., 2014) based on latitude, season and the tropopause height. Other satellite O<sub>3</sub> retrievals use a priori profiles from climatologies but they are chosen based on geographical and temporal criteria only (Bowman et al., 2006; Liu et al., 2010). Dufour et al. (2015) use three different a priori profiles picked up according to three broad tropopause height classes to represent high, middle and tropical latitudes. To our knowledge, it is the first time that the tropopause height is used in such a comprehensive way for the choice of the a priori profile for spaceborne O<sub>3</sub> retrievals.

The general statistics (Taylor diagrams) of the comparisons between ozonesonde and SOFRID have highlighted the large improvements brought by v3.5 especially in the troposphere. The use of a tropopause-based a priori profile generally reduces the RMSDs and increases the correlation coefficients and the amplitude of the retrieved variability. The high TOC biases of v1.6 relative to low O<sub>3</sub> are also corrected with v3.5. This is of particular importance in the SH extratropics where the very large biases almost disappear. In the NH, lower TOCs are retrieved in winter, leading to a better seasonal cycle. A sensitivity test demonstrated that these SOFRID improvements are dominated by the seasonal and latitude dependence of the a priori profile.

In the UTLS and stratosphere, the improvements are less important. In particular, both versions are impacted by positive biases for the UTLS (18 % at NH midlatitudes) and stratospheric ( $< 7\%$ ) columns at extratropical latitudes that were already discussed in Dufour et al. (2012). In the trop-



**Figure 16.** Time series of SOFRID-O3 (a) v1.6 and (c) v3.5 surface–300 hPa columns for the Northern Hemisphere (0–90° N). Blue lines indicate IASI retrievals, red lines indicate raw sonde data, and green lines indicate smoothed sonde data. Differences between IASI and sonde data for (b) v1.6 and (d) v3.5. Red lines indicate raw sonde data and green lines indicate smoothed sonde data.

ics, large profile oscillations around the tropopause result in negative biases in the UTLS (21 % in the SH) and positive biases (< 14 %) in the stratospheric columns.

Concerning the TOC drifts, we have shown that there were no significant differences between v1.6 and v3.5. There are no significant drifts except at high northern latitudes (increase of 9 % decade<sup>-1</sup>–13 % decade<sup>-1</sup>) and at southern tropical latitudes (decrease of 4 % decade<sup>-1</sup>–5 % decade<sup>-1</sup>). For southern tropics, the apparent decrease is probably linked to a sampling weakness at different stations which makes the time series inhomogeneous.

Our study has also demonstrated the importance of making comparisons with both raw and smoothed in situ data. Comparisons only with smoothed data could lead to the conclusion that the satellite data are better than they really are. For instance, the high bias for low TOC with v1.6 is almost completely corrected when smoothing is applied. The real improvement of v3.5 relative to v1.6 is only sizable when we compare SOFRID retrievals with raw sonde data.

Finally, we have compared our validation results to the latest (v20151001) FORLI-O3 retrieval validation. The comparison had to be limited because the variability of FORLI-O3 retrievals and ozonesonde data was not provided in Boynard et al. (2018), which prevented us to draw Taylor diagrams. Furthermore, in Boynard et al. (2018), the FORLI-O3 data are compared to smoothed sonde data only. FORLI produces larger RMSDs than SOFRID especially in the stratosphere at high latitudes. The coefficients of determination ( $R^2$ ) are consequently lower for FORLI columns than for SOFRID. Tropospheric biases are significantly larger for

FORLI (7 %–20 %) than for SOFRID (< 6 %). Finally, no significant tropospheric O<sub>3</sub> drift is detected for both versions of SOFRID-O3 in the NH. The difference with FORLI which is impacted by a significant TOC jump in 2010 (Boynard et al., 2018; Wespes et al., 2018) is likely linked to the use of different temperature profiles for the radiative transfer calculations (ECMWF analyses for SOFRID and EUMETSAT L2 for FORLI).

**Data availability.** The SOFRID-O3 data are freely available on the IASI-SOFRID website (<http://thredds.sedoo.fr/iasi-sofrid-o3-co/>, last access: 29 September 2020; SEDOO, 2014). IASI L1c data have been downloaded from the Ether French atmospheric database (<http://ether.ipsl.jussieu.fr>, last access: 29 September 2020; Ether, ). The research with IASI was conducted with some financial support from the Centre national d'études spatiales (CNES) (TOSCA–IASI project). The ozonesonde data used in this study were provided by the WOUDC (<https://doi.org/10.14287/10000008>, WOUDC database, 2020).

**Author contributions.** BB performed the validation of SOFRID-O3 data and wrote the paper. EE initiated and contributed to the development of SOFRID-O3 v3.5. ELF was in charge of the SOFRID retrieval operations.

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

*Acknowledgements.* The authors thank those responsible for the WOUDC measurements and archives for making the ozonesonde data available.

*Financial support.* This research has been supported by the “Centre National d’Etudes Spatiales” (CNES) within the framework of the IASI project.

*Review statement.* This paper was edited by Mark Weber and reviewed by two anonymous referees.

## References

- Ainsworth, E., Yendrek, C., Sitch, S., Collins, W., and Emberson, L.: The Effects of Tropospheric Ozone on Net Primary Productivity and Implications for Climate Change, *Annu. Rev. Plant Biol.*, 63, 637–661, <https://doi.org/10.1146/annurev-arplant-042110-103829>, 2012.
- Ball, W. T., Alsing, J., Mortlock, D. J., Staehelin, J., Haigh, J. D., Peter, T., Tummon, F., Stübi, R., Stenke, A., Anderson, J., Bourassa, A., Davis, S. M., Degenstein, D., Frith, S., Froidevaux, L., Roth, C., Sofieva, V., Wang, R., Wild, J., Yu, P., Ziemke, J. R., and Rozanov, E. V.: Evidence for a continuous decline in lower stratospheric ozone offsetting ozone layer recovery, *Atmos. Chem. Phys.*, 18, 1379–1394, <https://doi.org/10.5194/acp-18-1379-2018>, 2018.
- Barret, B., De Maziere, M., and Demoulin, P.: Retrieval and characterization of ozone profiles from solar infrared spectra at the Jungfraujoch, *J. Geophys. Res.-Atmos.*, 107, 4788, <https://doi.org/10.1029/2001JD001298>, 2002.
- Barret, B., Turquety, S., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Bey, I., Auvray, M., and Coheur, P.-F.: Global carbon monoxide vertical distributions from spaceborne high-resolution FTIR nadir measurements, *Atmos. Chem. Phys.*, 5, 2901–2914, <https://doi.org/10.5194/acp-5-2901-2005>, 2005.
- Barret, B., Le Flochmoen, E., Sauvage, B., Pavelin, E., Matricardi, M., and Cammas, J. P.: The detection of post-monsoon tropospheric ozone variability over south Asia using IASI data, *Atmos. Chem. Phys.*, 11, 9533–9548, <https://doi.org/10.5194/acp-11-9533-2011>, 2011.
- Barret, B., Sauvage, B., Bennouna, Y., and Le Flochmoen, E.: Upper-tropospheric CO and O<sub>3</sub> budget during the Asian summer monsoon, *Atmos. Chem. Phys.*, 16, 9129–9147, <https://doi.org/10.5194/acp-16-9129-2016>, 2016.
- Bowman, K. W., Rodgers, C. D., Kulawik, S. S., Worden, J., Sarkissian, E., Osterman, G., Steck, T., Lou, M., Eldering, A., Shephard, M., Worden, H., Lampel, M., Clough, S., Brown, P., Rinsland, C., Gunson, M., and Beer, R.: Tropospheric emission spectrometer: Retrieval method and error analysis, *IEEE T. Geosci. Remote*, 44, 1297–1307, 2006.
- Boynard, A., Hurtmans, D., Koukouli, M. E., Goutail, F., Bureau, J., Safieddine, S., Lerot, C., Hadji-Lazaro, J., Wespes, C., Pommereau, J.-P., Pazmino, A., Zyrichidou, I., Balis, D., Barbe, A., Mikhailenko, S. N., Loyola, D., Valks, P., Van Roozendaal, M., Coheur, P.-F., and Clerbaux, C.: Seven years of IASI ozone retrievals from FORLI: validation with independent total column and vertical profile measurements, *Atmos. Meas. Tech.*, 9, 4327–4353, <https://doi.org/10.5194/amt-9-4327-2016>, 2016.
- Boynard, A., Hurtmans, D., Garane, K., Goutail, F., Hadji-Lazaro, J., Koukouli, M. E., Wespes, C., Vigouroux, C., Keppens, A., Pommereau, J.-P., Pazmino, A., Balis, D., Loyola, D., Valks, P., Sussmann, R., Smale, D., Coheur, P.-F., and Clerbaux, C.: Validation of the IASI FORLI/EUMETSAT ozone products using satellite (GOME-2), ground-based (Brewer–Dobson, SAOZ, FTIR) and ozonesonde measurements, *Atmos. Meas. Tech.*, 11, 5125–5152, <https://doi.org/10.5194/amt-11-5125-2018>, 2018.
- Brunekreef, B. and Holgate, S. T.: Air pollution and health, *The Lancet*, 360, 1233–1242, 2002.
- Chen, W., Liao, H., and Seinfeld, J.: Future climate impacts of direct radiative forcing of anthropogenic aerosols, tropospheric ozone, and long-lived greenhouse gases, *J. Geophys. Res.-Atmos.*, 112, D14209, <https://doi.org/10.1029/2006jd008051>, 2007.
- Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, *Atmos. Chem. Phys.*, 9, 6041–6054, <https://doi.org/10.5194/acp-9-6041-2009>, 2009.
- Clough, S., Shephard, M., Mlawer, E., Delamere, J., Iacono, M., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: A summary of the AER codes, *J. Quant. Spectrosc. Rad. Transf.*, 91, 233–244, 2005.
- Coheur, P., Barret, B., Turquety, S., Hurtmans, D., Hadji-Lazaro, J., and Clerbaux, C.: Retrieval and characterization of ozone vertical profiles from a thermal infrared nadir sounder, *J. Geophys. Res.*, 110, D24303, <https://doi.org/10.1029/2005JD005845>, 2005.
- De Maziere, M., Hennen, O., Van Roozendaal, M., Demoulin, P., and De Backer, H.: Daily ozone vertical profile model built on geophysical grounds, for column retrieval from atmospheric high-resolution infrared spectra, *J. Geophys. Res.-Atmos.*, 104, 23855–23869, <https://doi.org/10.1029/1999jd900347>, 1999.
- Dufour, G., Eremenko, M., Griesfeller, A., Barret, B., LeFlochmoën, E., Clerbaux, C., Hadji-Lazaro, J., Coheur, P.-F., and Hurtmans, D.: Validation of three different scientific ozone products retrieved from IASI spectra using ozonesondes, *Atmos. Meas. Tech.*, 5, 611–630, <https://doi.org/10.5194/amt-5-611-2012>, 2012.
- Dufour, G., Eremenko, M., Cuesta, J., Doche, C., Foret, G., Beekmann, M., Cheiney, A., Wang, Y., Cai, Z., Liu, Y., Takigawa, M., Kanaya, Y., and Flaud, J.-M.: Springtime daily variations in lower-tropospheric ozone over east Asia: the role of cyclonic activity and pollution as observed from space with IASI, *Atmos. Chem. Phys.*, 15, 10839–10856, <https://doi.org/10.5194/acp-15-10839-2015>, 2015.
- Dufour, G., Eremenko, M., Beekmann, M., Cuesta, J., Foret, G., Fortems-Cheiney, A., Lachâtre, M., Lin, W., Liu, Y., Xu, X., and Zhang, Y.: Lower tropospheric ozone over the North China Plain: variability and trends revealed by IASI satellite observations for 2008–2016, *Atmos. Chem. Phys.*, 18, 16439–16459, <https://doi.org/10.5194/acp-18-16439-2018>, 2018.
- Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, A., El Amraoui, L., Pannekoucke, O., and Cariolle, D.: Combined assimilation of IASI and MLS observations to constrain tropospheric and stratospheric ozone in a global

- chemical transport model, *Atmos. Chem. Phys.*, 14, 177–198, <https://doi.org/10.5194/acp-14-177-2014>, 2014.
- Emili, E., Barret, B., Le Flochmoën, E., and Cariolle, D.: Comparison between the assimilation of IASI Level 2 ozone retrievals and Level 1 radiances in a chemical transport model, *Atmos. Meas. Tech.*, 12, 3963–3984, <https://doi.org/10.5194/amt-12-3963-2019>, 2019.
- Eremenko, M., Sgheri, L., Ridolfi, M., Cuesta, J., Costantino, L., Sellitto, P., and Dufour, G.: Tropospheric ozone retrieval from thermal infrared nadir satellite measurements: Towards more adaptability of the constraint using a self-adapting regularization, *J. Quant. Spectrosc. Ra. Transf.*, 238, UNSP 106577, <https://doi.org/10.1016/j.jqsrt.2019.106577>, 2019.
- Ether: AERIS, Données et services pour l'atmosphère, <https://www.aeris-data.fr/>, last access: 29 September 2020.
- Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P. F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojje, F., Foret, G., Garcia, O., Granados-Munoz, M. J., Hannigan, J. W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoen, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, *Elem. Sci. Anthropol.*, 6, 39, <https://doi.org/10.1525/elementa.291>, 2018.
- Hassler, B., Bodeker, G. E., and Dameris, M.: Technical Note: A new global database of trace gases and aerosols from multiple sources of high vertical resolution measurements, *Atmos. Chem. Phys.*, 8, 5403–5421, <https://doi.org/10.5194/acp-8-5403-2008>, 2008.
- Havemann, S., NWPSAF 1D-Var User Manual Software Version 1.2, Report NWPSAF-MO-UD-032, EUMETSAT Numerical Weather Prediction Satellite Applications Facility, Met Office, Exeter, U.K., available at: <https://nwp-saf.eumetsat.int/site/>, last access: 29 September 2020.
- Hocking, J., Rayer, P., Rundle, D., Saunders, R., Matricardi, M., Geer, A., Brunel, P., and Vidot, J.: RTTOV v11 Users Guide, NWP SAF, 2015.
- Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, *Atmos. Chem. Phys.*, 10, 2521–2537, <https://doi.org/10.5194/acp-10-2521-2010>, 2010.
- Matricardi, M., Chevallier, F., Kelly, G., and Thepaut, J. N.: An improved general fast radiative transfer model for the assimilation of radiance observations, *Q. J. Roy. Meteorol. Soc.*, 130, 153–173, 2004.
- McPeters, R. D., Labow, G. J., and Logan, J. A.: Ozone climatological profiles for satellite retrieval algorithms, *J. Geophys. Res.-Atmos.*, 112, D05308, <https://doi.org/10.1029/2005jd006823>, 2007.
- Pavelin, E., English, S., and Eyre, J.: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction, *Q. J. Roy. Meteorol. Soc.*, 134, 737–749, 2008.
- Peiro, H., Emili, E., Cariolle, D., Barret, B., and Le Flochmoën, E.: Multi-year assimilation of IASI and MLS ozone retrievals: variability of tropospheric ozone over the tropics in response to ENSO, *Atmos. Chem. Phys.*, 18, 6939–6958, <https://doi.org/10.5194/acp-18-6939-2018>, 2018.
- Richards, N. A. D., Arnold, S. R., Chipperfield, M. P., Miles, G., Rap, A., Siddans, R., Monks, S. A., and Hollaway, M. J.: The Mediterranean summertime ozone maximum: global emission sensitivities and radiative impacts, *Atmos. Chem. Phys.*, 13, 2331–2345, <https://doi.org/10.5194/acp-13-2331-2013>, 2013.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics – Vol. 2, World Scientific, Singapore, New Jersey, London, Hong Kong, 238 pp., 2000.
- Rothman, L. S. and Jacquemart, D., and Barbe, A. E. A.: The HITRAN 2004 molecular spectroscopic database, *J. Quant. Spectrosc. Ra. Transf.*, 96, 139–204, 2005.
- Rothman, L. S., Gordon, I. E., and Barbe, A. E. A.: The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Ra. Transf.*, 110, 533–572, <https://doi.org/10.1016/j.jqsrt.2009.02.013>, 2009.
- Saunders, R., Matricardi, M., and Brunel, P.: An improved fast radiative transfer model for assimilation of satellite radiance observations, *Q. J. Roy. Meteorol. Soc.*, 125, 1407–1425, 1999.
- Sauvage, B., Thouret, V., Thompson, A. M., Witte, J. C., Cammas, J. P., Nedelec, P., and Athier, G.: Enhanced view of the “tropical Atlantic ozone paradox” and “zonal wave one” from the in situ MOZAIC and SHADOZ data, *J. Geophys. Res.-Atmos.*, 111, D01301, <https://doi.org/10.1029/2005jd006241>, 2006.
- SEDOO: IASI-SOFRID database, available at: <http://thredds.sedoo.fr/iasi-sofrid-o3-co/> (last access: 29 September 2020), 2014.
- Sellitto, P., Dufour, G., Eremenko, M., Cuesta, J., Dauphin, P., Forêt, G., Gaubert, B., Beekmann, M., Peuch, V.-H., and Flaud, J.-M.: Analysis of the potential of one possible instrumental configuration of the next generation of IASI instruments to monitor lower tropospheric ozone, *Atmos. Meas. Tech.*, 6, 621–635, <https://doi.org/10.5194/amt-6-621-2013>, 2013.
- Shindell, D. T., Faluvegi, G., and Stevenson, D. S. E. A.: Multi-model simulations of carbon monoxide: Comparison with observations and projected near-future changes, *J. Geophys. Res.-Atmos.*, 111, D19306, <https://doi.org/10.1029/2006jd007100>, 2006.
- Sofieva, V. F., Tamminen, J., Kyrölä, E., Mielonen, T., Veefkind, P., Hassler, B., and Bodeker, G. E.: A novel tropopause-related climatology of ozone profiles, *Atmos. Chem. Phys.*, 14, 283–299, <https://doi.org/10.5194/acp-14-283-2014>, 2014.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.-Atmos.*, 106, 7183–7192, <https://doi.org/10.1029/2000jd900719>, 2001.
- Thompson, A. M., Witte, J. C., Oltmans, S. J., Schmidlin, F. J., Logan, J. A., Fujiwara, M., Kirchhoff, V. W. J. H., Posny, F., Coetzee, G. J. R., Hoegger, B., Kawakami, S., Ogawa, T., Fortuin, J. P. F., and Kelder, H. M.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology 2. Tropospheric variability and the zonal wave-one, *J. Geophys.*

- Res.-Atmos., 108, 8241, <https://doi.org/10.1029/2002JD002241>, 2003.
- Wang, P. H., Cunnold, D. M., Trepte, C. R., Wang, H. J., Jing, P., Fishman, J., Brackett, V. G., Zawodney, J. M., and Bodeker, G. E.: Ozone variability in the midlatitude upper troposphere and lower stratosphere diagnosed from a monthly SAGE II climatology relative to the tropopause, *J. Geophys. Res.-Atmos.*, 111, D21304, <https://doi.org/10.1029/2005jd006108>, 2006.
- Wespes, C., Hurtmans, D., Clerbaux, C., and Coheur, P. F.: O-3 variability in the troposphere as observed by IASI over 2008–2016: Contribution of atmospheric chemistry and dynamics, *J. Geophys. Res.-Atmos.*, 122, 2429–2451, <https://doi.org/10.1002/2016jd025875>, 2017.
- Wespes, C., Hurtmans, D., Clerbaux, C., Boynard, A., and Coheur, P.-F.: Decrease in tropospheric O<sub>3</sub> levels in the Northern Hemisphere observed by IASI, *Atmos. Chem. Phys.*, 18, 6867–6885, <https://doi.org/10.5194/acp-18-6867-2018>, 2018.
- Wespes, C., Hurtmans, D., Chabrillat, S., Ronsmans, G., Clerbaux, C., and Coheur, P.-F.: Is the recovery of stratospheric O<sub>3</sub> speeding up in the Southern Hemisphere? An evaluation from the first IASI decadal record (2008–2017), *Atmos. Chem. Phys.*, 19, 14031–14056, <https://doi.org/10.5194/acp-19-14031-2019>, 2019.
- WMO: WMO: Meteorology – A three-dimensional science: Second session of the Commission for Aerology, *WMO Bulletin*, 4, 134–138, 1957.
- WOUDC database: Dataset Information: OzoneSonde, Government of Canada, <https://doi.org/10.14287/10000008>, 2020.
- Zhang, L., Li, Q. B., Murray, L. T., Luo, M., Liu, H., Jiang, J. H., Mao, Y., Chen, D., Gao, M., and Livesey, N.: A tropospheric ozone maximum over the equatorial Southern Indian Ocean, *Atmos. Chem. Phys.*, 12, 4279–4296, <https://doi.org/10.5194/acp-12-4279-2012>, 2012.
- Ziemke, J. R., Oman, L. D., Strode, S. A., Douglass, A. R., Olsen, M. A., McPeters, R. D., Bhartia, P. K., Froidevaux, L., Labow, G. J., Witte, J. C., Thompson, A. M., Haffner, D. P., Kramarova, N. A., Frith, S. M., Huang, L.-K., Jaross, G. R., Seftor, C. J., Deland, M. T., and Taylor, S. L.: Trends in global tropospheric ozone inferred from a composite record of TOMS/OMI/MLS/OMPS satellite measurements and the MERRA-2 GMI simulation, *Atmos. Chem. Phys.*, 19, 3257–3269, <https://doi.org/10.5194/acp-19-3257-2019>, 2019.