

Quality Assessment of the AC SAF GOME-2 gridded ozone profile data records

Olaf N.E. Tuinder^{1,*}, Andy W. Delcloo^{2,3,*}, and Peggy Achtert^{4,5,*}

¹Royal Netherlands Meteorological Institute, Utrechtseweg 297, De Bilt, The Netherlands

²Royal Meteorological Institute of Belgium, Observations Department, ACM²-team, Ringlaan 3, B-1080 Ukkel, Belgium

³Ghent University, Department of Physics and Astronomy, Krijgslaan 281/S9, B-9000, Ghent, Belgium

⁴Deutscher Wetterdienst, Hohenpeissenberg, D-82383, Germany

⁵Leipzig Institute for Meteorology, Leipzig University, Stephanstraße 3, 04103 Leipzig, Germany

*These authors contributed equally to this work.

Correspondence: Olaf N.E. Tuinder (olaf.tuinder@knmi.nl)

Abstract. One of the objectives of the Atmospheric Composition SAF ([AC SAF](#)) is to produce satellite-derived monthly mean data records that are valuable for operational, scientific, and other applications. One of these data records is the gridded global GOME-2A/B/C ozone profile data set presented in this paper. This data record covers the period 2007–2024 and consists of ozone partial columns on a $0.25^\circ \times 0.25^\circ$ grid with 40 vertical layers, with the associated (averaged) averaging kernel

5 ~~(AK)~~ and the a priori needed to use the data in other applications. This paper presents the GOME-2 instrument, the (level-2) ozone profile retrieval method and the subsequent gridding procedure used to generate the level-3 product. We discuss the methodology for averaging AKs in latitude bands and demonstrate that the principal structural features are preserved. We provide a description of the balloon sounding, lidar, FTIR and microwave radiometers validation data and methods, and then perform a quality assessment of the gridded ozone profile product through comparison with these independent data sources

10 for the tropics, mid-latitude and polar latitude bands ~~and in three~~ [for four](#) vertical regions: the Troposphere, [the UTLS](#) and the Lower and Upper Stratosphere. Detailed analyses of absolute and relative differences are provided for each region and height range. The results demonstrate a high level of consistency across the three GOME-2 instruments (with GOME-2A used only up to 2018); ~~underscoring the reliability of the GOME-2 constellation for long-term ozone monitoring~~. In the troposphere, all three sensors tend to slightly overestimate ozone, with absolute differences ~~(AD)~~ of roughly +1 to +3 [Dobson Units \(DU\)](#)

15 in mid-latitudes and up to +7.5 DU in the tropics for Metop-C. In [the more variable UTLS altitude region, the absolute differences range between -2.5 and 5 DU](#). In the lower stratosphere, all sensors show a small negative bias, typically between -3 and -7 DU ~~(RD)~~ [relative difference](#) $\approx -3\%$ to -7% , corresponding to a modest underestimation of ozone concentrations. In the upper stratosphere, biases are minimal across all sensors, with absolute difference values, close to zero (-0.1 to -0.4 DU) and ~~extremely low variability~~ [\(STDEV \$\approx 0.1-0.2\$ \) low variability](#). These findings underscore the reliability of the GOME-2

20 constellation for long-term ozone analyses and the potential for merged multi-sensor time series without significant inter-calibration artifacts suitable for climate and atmospheric research.

1 Introduction

Monitoring atmospheric composition is essential due to several human-induced changes in the atmosphere, such as global warming, stratospheric ozone depletion, increased ultraviolet radiation, and air pollution (IPCC, 2021; WMO, 2018; Burrows et al., 2011). In addition, such monitoring enables timely responses to natural hazards and provides the scientific basis for evaluating the effectiveness of international agreements, such as the 1987 Montreal Protocol and its amendments (WMO, 2018).

The Montreal Protocol, implemented in 1989, has been instrumental in reducing the atmospheric concentrations of ozone-depleting substances. As a result, the ozone layer is on a path to recovery (Chippenfield and Bekki, 2024; Solomon et al., 2016). Recent assessments indicate that total column ozone is expected to return to 1980 levels by around 2040 for the near-global average (60°S–60°N), around 2045 for the Arctic, and around 2066 for the Antarctic (WMO, 2022).

However, recovery is not uniform across all regions and altitudes. While ozone in the upper stratosphere has increased significantly — by approximately 1–3% per decade since 2000 (Steinbrecht et al., 2017) — this increase is most pronounced in the mid-latitudes of both hemispheres (Godin-Beekmann et al., 2022). In contrast, the lower stratosphere, particularly in the tropics and mid-latitudes, has not shown clear signs of recovery (Bognar et al., 2022). The increase in upper stratospheric ozone is primarily mainly driven by two factors: the decline in stratospheric chlorine loading following the implementation of the Montreal Protocol, and the cooling of the upper stratosphere, which slows the ozone-destroying catalytic cycles and enhances ozone production via ter-molecular reactions the production of ozone through three-body molecular recombination (Steinbrecht et al., 2025). This vertical and regional discrepancy highlights the importance of sustained, detailed observations of ozone profiles (WMO, 2018) in the stratosphere (WMO, 2018). In addition, tropospheric ozone is of particular importance due to its adverse impacts on air quality, human health, and ecosystems, as well as its contribution to radiative forcing and climate change (IPCC, 2021; Monks et al., 2015). Tropospheric ozone is a secondary pollutant formed through complex photochemical reactions involving nitrogen oxides and volatile organic compounds, and its concentrations are strongly influenced by anthropogenic emissions and large-scale transport processes (Donzelli and Suarez-Varela, 2024; Monks et al., 2015). Satellite-based measurements are central to these global global ozone monitoring efforts, providing consistent, long-term, and spatially comprehensive data records of key atmospheric constituents. Ozone remains a particularly critical species in this context, as stratospheric ozone protects life on Earth from harmful ultraviolet radiation from the Sun, while tropospheric ozone acts as a pollutant and greenhouse gas (Burrows et al., 2011; Xu, 2024). In particular, the Global Ozone Monitoring Experiment (GOME) and its successor GOME-2, flown on board the Metop satellite series, have provided continuous global ozone observations for more than two decades, forming a key component of long-term ozone monitoring and climate data records (Munro et al., 2016; Callies et al., 2000).

This study focuses on the gridded ozone profile products, generated within the framework of the EUMETSAT Atmospheric Composition Satellite Application Facility (AC SAF). These data records are retrieved from measurements of the Global Ozone Monitoring Experiment-2 (GOME-2) instruments onboard the Metop-A, Metop-B, and Metop-C satellites (launched in 2006, 2012 and 2019 respectively). Together, these platforms provide a multi-decadal dataset suitable for long-term assessments.

55 High-quality ozone profile records are essential for reliable trend detection and are expected to contribute to global assessments such as the World Meteorological Organization (WMO) Scientific Assessment of Ozone Depletion. Therefore, a rigorous validation of the GOME-2 gridded ozone profile data records is required to quantify their accuracy, stability, and suitability for scientific and policy-relevant applications.

In recent years, ~~the satellite-based instrument~~ time series of ~~satellite-based instruments of~~ the GOME-type (including
60 GOME-2) have been central to ~~construct the construction of~~ long-term ~~,harmonised~~ harmonized ozone profile and tropospheric ozone datasets. Coldewey-Egbers et al. (2025) present the GOP-ECV record, spanning 1995 to 2021, which merges nadir ~~UV/visible~~ ultra violet (UV) and visible (VIS) measurements from GOME, SCIAMACHY, OMI and GOME-2 into monthly mean partial columns (19 layers, surface to 80 km) on a $5^\circ \times 5^\circ$ grid by carefully harmonizing inter-sensor differences and aligning with the GOME-type Total Ozone Essential Climate Variable (GTO-ECV) total-ozone record. Within the European
65 Space Agency (ESA) Climate Change Initiative (CCI) framework, ozone profile retrievals from GOME-2 measurements have also been developed, most notably by the Rutherford Appleton Laboratory (RAL), contributing to the ESA Ozone CCI Climate Data Record (Pope et al., 2023; Sofieva et al., 2023; Ozone CCI, 2010). Although Ozone CCI products mainly aim to construct a homogenized climate data record using multiple sensors and retrieval algorithms, the AC SAF GOME-2 gridded ozone profile products addressed in this study are generated within an operational framework with a strong focus on near-real-time
70 processing, long-term stability, and continuity for user-driven applications. Therefore, both efforts are complementary and jointly contribute to the comprehensive use of GOME-2 measurements for atmospheric ozone monitoring. Arosio et al. (2025) provide an ~~inter-comparison~~ inter-comparison study between existing tropospheric ozone-column datasets derived from combined nadir- and limb-observations, assessing consistency, bias and trend detectability. Keppens et al. (2025) further address the harmonization challenge by combining sixteen tropospheric-ozone satellite records (including GOME-2-derived products)
75 to strengthen confidence in tropospheric ozone time-series and their long-term behavior. Regional studies such as Gaudel et al. (2024) exploit satellite records (including IASI/GOME-2) and in situ profiles to quantify tropical tropospheric ozone trends and regional burdens, while Okamoto et al. (2024) apply a ~~multispectral~~ multi spectral IASI + GOME-2 approach over the tropical Atlantic together with aircraft data to discern natural (biomass burning, stratosphere-troposphere exchange) versus anthropogenic influences on tropospheric ozone variability.

80 Collectively, these works demonstrate that GOME-2 and related sensors now underpin robust ozone-profile and tropospheric-ozone time-series, enabling assessments of spatial distribution, long-term trends and attribution of ozone changes. This paper presents a comprehensive quality assessment of the monthly mean gridded GOME-2A/B/C ozone profile data records. Validation is based on intercomparisons with independent ground-based measurements: ~~ozonesondes~~ ozone sondes, lidar, ~~fourier~~ Fourier transform infrared (FTIR) and microwave radiometers (MWR). The focus is on characterizing the performance of the
85 data products in different atmospheric regions (troposphere, Upper Troposphere/Lower Stratosphere interface (UTLS) lower, and upper stratosphere) and over various latitudinal bands, using monthly averaged profiles to evaluate the long-term consistency and bias structure of the dataset.

2 Ozone profiles from GOME-2

2.1 The GOME-2 instrument

90 The Global Ozone Monitoring Experiment 2 (GOME-2) are a series of nadir viewing instruments mounted on the Metop satellite series. These instruments fly in a polar orbit around the Earth and have an equator overpass time around 09:30 local solar time. The instrument uses back scattered light from the Sun in the UV-VIS spectral region (240-790 nm). The nominal swath width is 1920 km on the Earth's surface with 24 pixels in the forward scanning direction of $80 \times 40 \text{ km}^2$ (cross track \times along track). With the (nominal) wide swath width, global coverage can be achieved in 2 days.

95 In Table 1 the measurement modes and time periods included in this study are given. The GOME-2A instrument has gone through a number of modes. In December of 2008 the wavelength separation of ~~band-1A-band-1a~~ / ~~band-1B-band-1b~~ was shifted from 300 nm to 283 nm, resulting in stratospheric information with smaller footprints. In mid-July 2013, the swath width was changed from the nominal 1920 km to 960 km, which remained until the End of Life de-orbit and shut down of the Metop-A platform in 2021. GOME-2B and GOME-2C have run in the same wide configuration since they were launched.

100 2.2 Vertical ozone profile retrieval

While the details of the ozone profile retrieval algorithm used to produce the ozone data are given in the [ATBD-Algorithm Theoretical Baseline Document \(ATBD\)](#) of the Near Real Time (NRT), Offline and Data Record Vertical Ozone Profile and Tropospheric Ozone Column Products (Tuinder et al., 2022), we will provide a description of the relevant aspects of the algorithm below.

105 The Ozone Profile Retrieval Algorithm (Opera) retrieves vertical ozone profiles from GOME-2 data using the 265 nm–330 nm (UV-VIS) spectral range from the Main Science Channels in band-1a, 1b and 2b. Because of the longer integration time, ~~the~~ band-1a is used across multiple smaller band-1b/2b measurements.

The retrieved information consists of a vertical ozone profile of 40 ozone partial column ~~density~~-values (in DU) on a fixed pressure grid from the surface to 0.001 hPa, a fitted albedo value (either surface albedo or cloud albedo) and an additional offset
110 to correct for a systematic bias in ~~band-1A-band-1a~~ (possibly a result of stray light or signal leakage) ([Tuinder et al., 2022](#)). The fixed pressure grid is adjusted to the surface height and, in case of clouds, to the cloud top height.

Instrument	Swath	Period	Remarks
GOME-2A	Wide	2007-01-01 – 2008-12-10	original 1a/1b setting
"	Wide	2008-12-10 – 2013-07-15	1a/1b shift wide swath
"	Small	2013-07-15 – 2021-11-15	small swath until End Of Life
GOME-2B	Wide	2013-01-01 – 2024-12-31	currently flying
GOME-2C	Wide	2019-01-01 – 2024-12-31	currently flying

Table 1. Measurement modes for GOME-2A/B/C

The LIDORT-A radiative transfer model (Spurr et al., 2008) is used in combination with a polarization correction. For the inversion we use Optimal Estimation (Rodgers, 2000), iteratively reaching a converged state of the atmosphere.

Based on the GOME-2 measurement alone, the ozone profile is under-defined, and therefore an a priori ozone profile climatology is used (resolved in latitude bands and months) (McPeters and Labow, 2012).

The (initial) surface albedo comes from the GOME-2 version 3 surface ~~DLER~~ Directional Lambertian-Equivalent Reflectivity (DLER) database (see Tilstra et al. (2017, 2021)). The initial cloud albedo is set to 0.8.

Spectral spikes, such as those caused by highly energetic particles common over the South Atlantic Anomaly region, are removed by starting at 290 nm and filtering out all peaks below that wavelength above 4σ (where σ is the reflectance error), relative to the higher (valid) reflectance measurement. If there are more than 30 peaks removed, the remaining spectrum below the last valid reference measurement is removed completely because ~~the reference measurement~~ it is no longer representative.

2.2.1 Degradation correction

The GOME-2 instrument degrades over time, i.e. the signal level measured by the detector decreases with time, either through effects in the optical light path (including the scan mirror) or the sensitivity of the photon detector itself is affected (or both) (EUMETSAT, 2022). One complicating factor is that the measured radiance from the Earth degrades in a different way than the irradiance measured from the Sun, so they do not fully cancel out when the reflectance is calculated. In order to compensate for the long term instrument degradation, correction polynomials are fitted based on the ratio of daily averaged measured and calculated reflectance. The latter were generated by providing the forward model with an ozone profile from the climatology scaled with expected ozone columns from an assimilated ~~ozone field~~ total ozone column field (Eskes et al., 2003) from <https://temis.nl/protocols/O3global.php>.

2.3 Averaging Kernels

An averaging kernel (AK) ~~matrix~~ of an optimal estimation inversion is a ~~linear representation of~~ matrix containing the weights of the linear effect of a change in the fitted parameters on the retrieved value at a specific position in the state vector (Rodgers et al., 1990). In other words, it represents how changes in the ozone partial column at various layers along the vertical model atmosphere affect the outcome of a particular retrieved layer. In an optimal retrieval, the AK would be close to the identity matrix (peaking near the nominal layer), but in real world situations vertical resolution of a retrieved profile is wider, and can span multiple layers or even the whole atmosphere.

The retrieved ozone profile depends on the a priori, the true ozone profile and the averaging kernel according to the following relationship (Equation 1):

$$X_{retrieved} = X_{a\ priori} + AK(X_{true} - X_{a\ priori}) \quad (1)$$

where X denotes a state vector (i.e.: the list of values: the ozone profile, albedo and offset), and the subscripts indicate the type of vector. Unfortunately, inverting the averaging kernel ~~can not be inverted (due to singularity)~~ is not possible due to the matrix being singular (and also the pseudo inverse method does not produce sensible results), so the X_{true} cannot be calculated

on its own; it is always dependent on the other components and the relationship described above. Therefore a reference data set with a high vertical resolution can be easily compared to a satellite data set, not the other way around. In section 3.3 we will revisit the topic of averaging kernels.

3 The gridding from level-2 to level-3

3.1 The gridding method

The gridded level-3 ozone profile product is created based on a grid definition, which sets the name of the variable, the name of the source parameter, and the number of latitudes, longitudes and vertical layers in which to divide the horizontal and the vertical dimension. The default horizontal resolution is $0.25^\circ \times 0.25^\circ$ with 40 vertical layers from the surface to 0.001 hPa.

The ground footprint of each retrieved vertical ozone profile is divided into equal footprint-sub-cells (4 along track \times 8 across track) ~~which~~, creating 10×10 km sub-cells which centers are projected (binned) onto the fixed grid. At the end of the data binning, the averaged value, the standard deviation and other variables are calculated.

3.2 Calculated parameters

In the gridded ozone profile product, we provide the arithmetic mean, and the population standard deviation associated with the arithmetic mean, ~~the weighted mean (weighted by the error values of the parameter, if available) and the standard error of the weighted mean.~~

The arithmetic mean value in a grid cell is calculated as follows:

$$\text{Arithmetic Mean} = \frac{\sum_{n=1}^N x_i}{N} \quad (2)$$

where x_i is the individual data point and N is the number of data points. In order to calculate the arithmetic mean, the number of data points in a grid cell and the sum of the data values in that grid cell are stored. This enables daily grids to be merged into monthly grids and monthly grids into yearly grids, without loss of information. This also applies to re-gridding onto coarser horizontal resolutions which can also be calculated without loss of information.

The population standard deviation σ in a grid cell is calculated as follows:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{N}} \quad (3)$$

$$\sigma = \sqrt{\left(\frac{1}{N} \sum_{n=1}^N x_i^2 \right) - \left(\frac{1}{N} \sum_{n=1}^N x_i \right)^2} \quad (4)$$

$$\sigma = \frac{\sqrt{N \left(\sum_{n=1}^N x_i^2 \right) - \left(\sum_{n=1}^N x_i \right)^2}}{N} \quad (5)$$

In order to calculate the standard deviation, the sum of the squared data values ($\sum_{n=1}^N x_i^2$) needs to be stored, in addition to the sum of the values (already stored) ~~)-~~

If the weight of the weighted mean is taken as $1/\text{variance} = 1/(\text{error}_i^2)$, then the weighted mean and the standard error can be written as:-

$$\text{Weighted Mean} = \frac{\sum_{n=1}^N (x_i/(\text{error}_i)^2)}{\sum_{n=1}^N (1/(\text{error}_i)^2)}$$

$$\text{Standard Error} = \frac{1}{\sum_{n=1}^N (1/(\text{error}_i)^2)}$$

175 If one keeps track of two sums:-

$$\text{SumValDivSqErr} = \sum_{n=1}^N (x_i/(\text{error}_i)^2)$$

$$\text{SumOneDivSqErr} = \sum_{n=1}^N (1/(\text{error}_i)^2)$$

then the weighted mean and the standard error can be written as:-

$$\text{Weighted Mean} = \frac{\text{SumValDivSqErr}}{\text{SumOneDivSqErr}}$$

180
$$\text{Standard Error} = \sqrt{\frac{1}{\text{SumOneDivSqErr}}}$$

For parameters where there is no error available a unity value is taken (i.e.: 1.0) [for the Arithmetic Mean](#).

Table 2 lists the primary and derived values that are stored in the gridded ozone profile product.

3.3 Gridded averaging kernel

As explained earlier in section 2.3, the AK of the ozone profile describes the linear effect (weights) of ozone in all atmospheric
 185 layers on a particular (nominal) retrieved layer. In the level-2 ozone data, each individual retrieved profile has 40 layers and the associated AK matrix has 40×40 elements. While it is technically possible to store gridded / averaged AKs for all $0.25^\circ \times 0.25^\circ$ grid cells, this would make a regular level-3 gridded data product 40 times larger and unwieldy to use. As an alternative we have chosen to grid / average all AKs in 18 bands of 10° latitude.

Primary values	Derived values
the number of sub-pixels seen-used	the arithmetic mean (Eq. 2)
the minimum value of all data point values	the arithmetic-mean-standard deviation (Eq. 23)
the maximum value of all data point values	the standard deviation (Eq. 3)
the sum of the data point values	the weighted-mean (Eq. ??)
the sum of the square of each of the data point values	the weighted-mean-error (Eq. ??)
SumValDivSqError (Eq. ??) SumOneDivSqError (Eq. ??)	

Table 2. Primary and derived values stored in the gridded GOME-2 ozone profile product

An example of such a gridded / averaged AK is shown in Figure 1, where for March 2019 the GOME-2B AK matrix for the
190 latitude band at $45^\circ N$ is shown on the left, and the smoothing weight curves for the nominal layers are shown in the middle. In the matrix plot, the x-axis (x_i) can be seen as the nominal retrieved layer index, and the y-axis then represents the index of the smoothing weights of the contributing layers to the particular nominal layer at index x_i . Also shown in Figure 1 on the right is the ozone profile and a priori profile at $45^\circ N/0^\circ E$ for the same month, which gives a perspective in the shape of the profile and the position of the ozone maximum in the gridded data.

195 What stands out is that the bottom four layers have very little sensitivity (the matrix values are green across the vertical). In other words: the GOME-2 retrieved ozone profile is not very sensitive in the troposphere where (on average) relatively little ozone is present. For the next few layers, there is a positive sensitivity close to the nominal (indexed) layer, and a negative or positive weight higher up. Above layer 15 (31 hPa), the positive values start a little above the diagonal, which means that the profile is more sensitive to ozone above the nominal layer. For the top eight layers there is also little sensitivity, which
200 corresponds with the upper stratosphere with diminishing ozone as one goes up.

In order to assess the the impact of these positive and negative effects indicated in the AK, we should also take into account the ozone content at those layers location. Even though the weights of the AK curves are fanning out at the top layers, their impact on the retrieved value is relatively small because there is very little ozone above layer 30-30 (1 hPa).

As mentioned above, we longitudinally average the AKs ~~into matrices-so they are~~ representative for bands of 10° latitude
205 in order to save space in the output product. While this is unconventional, a study done by Johnson et al. (2018) shows that using a representative AK results in relatively small differences in a wide range of ozone profile cases (generally $< 5 - 10\%$ across the whole profile). To explore this a little further, we present as an example the variability of the averaging kernels ~~of the eo-located level-2 profiles in Ueele in in the $40^\circ - 50^\circ N$ latitude band in~~ Figure 2. This figure shows the mean of the collocated averaging kernels (top) and the 10th, 50th and 90th percentiles (bottom row) of a full year of level-2 GOME-2A ozone profile
210 retrievals near Ueele (Belgium) in the $40^\circ - 50^\circ N$ latitude band. The percentile plots provide an indication of the spread of the AK values in the level-2 data used as a source for the level-3. While the change in the AK value along the percentiles is most prominent in the upper stratosphere (above layer 25 (3 hPa)), this is also the region with diminishing ozone and therefore has a smaller effect on the retrieved ozone. At the same time, the main structure along the diagonal remains intact across the whole range of percentiles in the altitude region where most of the ozone is present. The mean and median (i.e. the 50th percentile) of

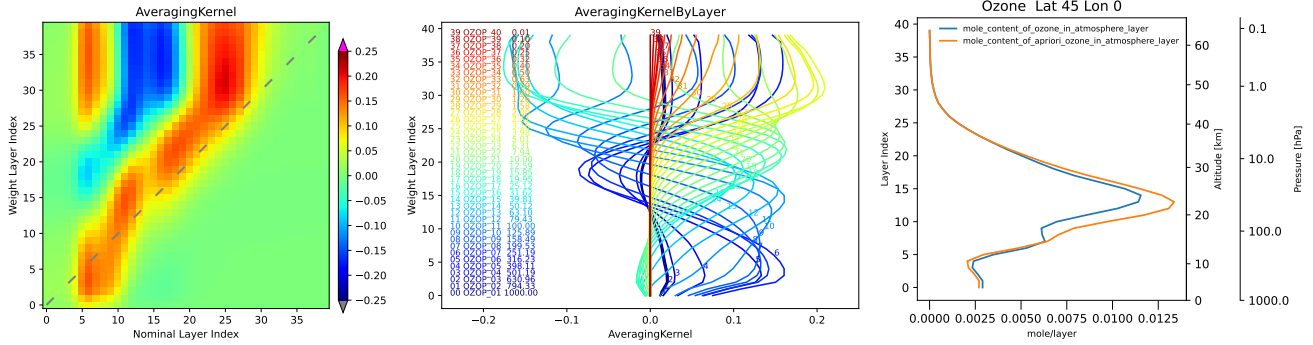


Figure 1. Example of the gridded averaging kernel, for GOME-2B at $45^\circ N$, (i.e. the $40^\circ N - 50^\circ N$ band) on March 2019. Left: the whole averaging kernel matrix; Middle: the weights-averaging kernel curves of the individual layers; Right: the ozone profile and a priori at $45^\circ N/0^\circ E$ in mole for the layer. In the left panel there is a dashed line indicating the diagonal.

215 the averaging kernels are very similar, so in the L3-level-3 gridded product we provide the mean because of ease of calculation. In the Appendix 1, the 10, 50 and 90 percentile are shown for all 18 latitude bands for 2009 as a reference.

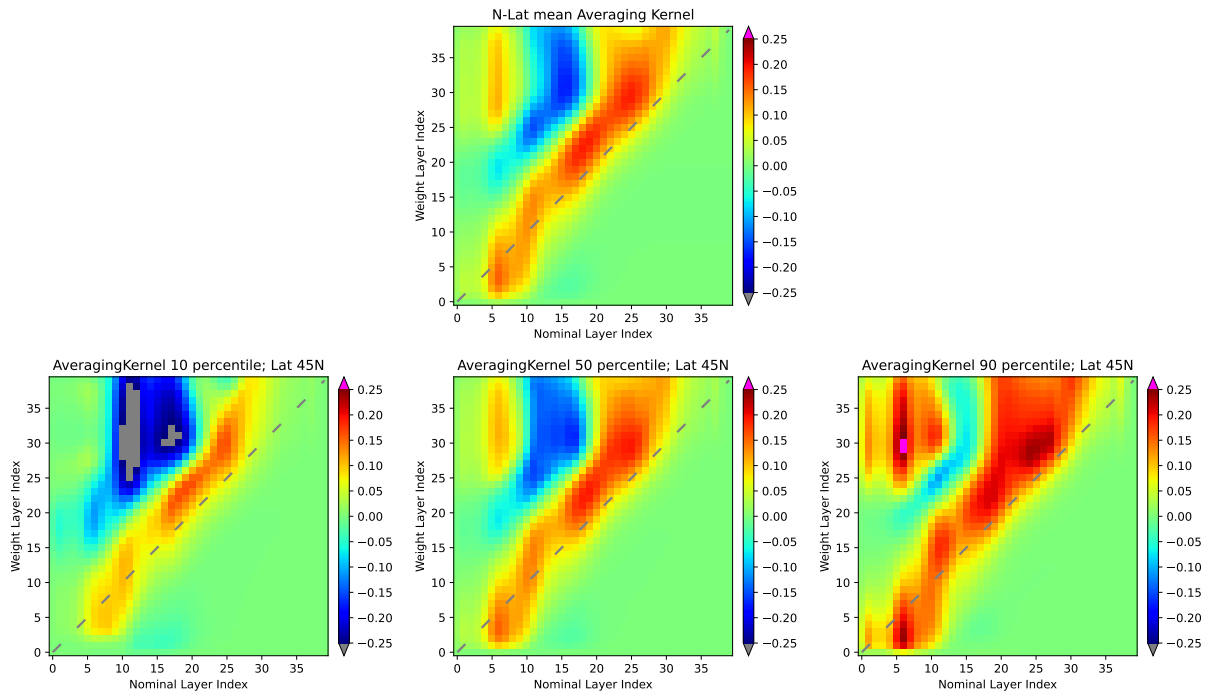


Figure 2. Averaged averaging kernels of GOME-2A ozone profile retrievals near-Uccle (Belgium) in the $40^\circ - 50^\circ N$ latitude band in 2009: the mean (top), and the 10th, 50th and 90th percentile (bottom row). The gray dashed lines represent the diagonal.

3.4 Time series of parameters

Figure 3 shows a time series of the global arithmetic mean of some of the parameters in the [L3-level-3](#) data set, which give a general impression of its behaviour over the mission period. It shows the atmosphere mole content of ozone (i.e.: the full atmosphere), the atmosphere mole content of the lower troposphere (surface to 500 hPa), the number of spectral detector pixels used in the (level-2) retrieval and the degrees of freedom of the resulting profile ([the DFS is the trace of the averaging kernel](#)). Due to specific validity of the spectral ranges of each of the GOME-2 instruments, the number of spectral pixels (NMeasurements) is also different. The degradation of the GOME-2 instrument contributes to a reduction of valid spectral information going into the level-2 retrieval over time. In the case of GOME-2A, the degradation (and quality control of the spectrum) causes an increasingly steep ~~drop-off~~[drop-off](#) of the number of valid spectral information of GOME-2A after 2018. The effect of this can be seen in an increasing deviation of the total ozone column compared to the other instruments, and especially in the troposphere. The degradation in general also has an effect on the degrees of freedom of the retrieval. [The DFS of GOME-2A has a discontinuity around September 2009, when an instrument de-contamination attempt caused a permanent degradation of the transmission and sensing of light through the instrument and detector.](#)

Even though the retrieved level-2 and gridded level-3 data is available to users, we recommend to not use the ozone data from GOME-2A after 2018 and restrict our validation of this instrument therefore up to the end of that year.

4 Validation Data and Methods

The quality assessment of the AC SAF GOME-2A/B/C gridded ozone profile products was carried out using independent ground-based observations. For the lower stratosphere, [UTLS](#) and troposphere, comparisons were made with ozonesonde measurements. Ozonesondes are balloon-borne instruments capable of measuring vertical ozone profiles from the surface up to approximately 30 km, with a much higher vertical resolution (~ 180m) than satellite sensors (WMO, 2018). In addition, ozonesondes generally provide better precision and accuracy in the lower stratosphere and troposphere. They can also be launched under nearly all meteorological conditions and at any time of day. The precision of ozonesonde measurements varies with altitude and depends on the type of instrument (Stauffer et al., 2022). For this validation study, only the ozonesonde station of Hohenpeissenberg still uses Brewer-Mast ozonesondes. For the Japanese stations ~~we can report that~~ KC sondes were operational until December 2009. The other stations used (Figure 4) in this validation study are all using ECC ozonesondes. Intercomparison campaigns and quality assurance studies have demonstrated systematic differences in precision between electrochemical concentration cell (ECC) and Brewer–Mast (BM) ozonesondes across the troposphere and stratosphere. In the troposphere (surface to ~10 km; ~1000–300 hPa), ECC sondes exhibit high precision, typically within $\pm 3\text{--}4\%$ [Stubi et al. \(2008\); Smit et al., \(2007\); WMO \(2014\)](#) ([Stubi et al., 2008; Smit et al., 2007; WMO, 2014](#)). In the lower to mid-stratosphere (10–30 km; ~300–10 hPa), ECC sondes maintain excellent precision of approximately $\pm 3\%$ across different manufacturers and cathode sensing solutions, while BM sondes show moderate precision of $\pm 3\text{--}5\%$ [Stubi et al. \(2008\); WMO \(2014\)](#). In the upper stratosphere (30–35 km; ~10–1 hPa), ECC sondes continue to perform well, with uncertainties slightly increasing to ~3–4%, with best performance reported around 40 hPa near the ozone maximum (Smit et al., 2007; WMO,

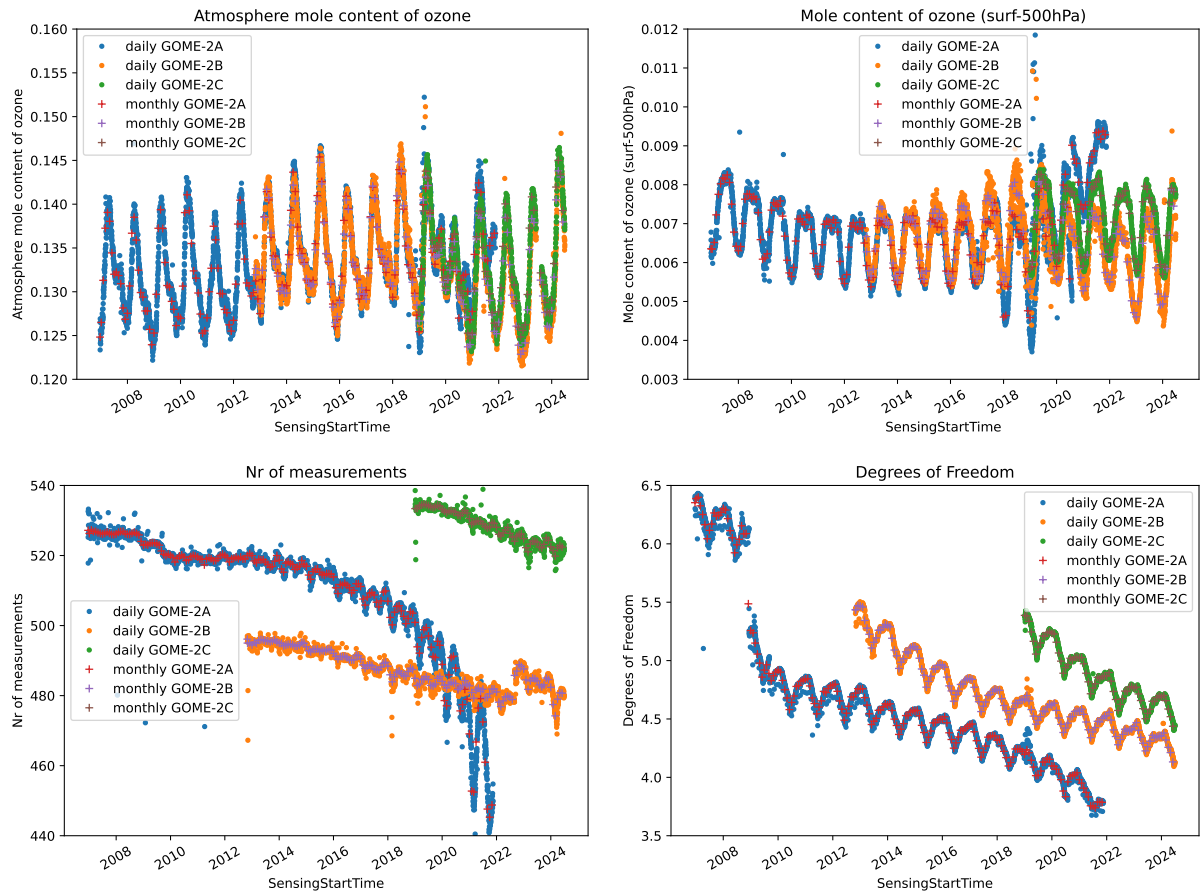


Figure 3. Time series of the global arithmetic mean of some parameters from GOME-2A/B/C jointly for the full available data length in the data record. Both the daily global mean is shown as well as the monthly mean. Top left: mole content of ozone (full vertical column), top right: mole content of ozone from the surface to 500hPa500hPa, bottom left: number of spectral detector pixels used in the retrieval, bottom right: number of degrees of freedom in the retrieval.

Table 3. Typical precision and vertical resolution of lidar and MWR systems, microwave radiometer (see Steinbrecht et al. (2006) MWR), and FTIR ozone profile measurements as used for satellite validation (Steinbrecht et al., 2006; Vigouroux et al., 2015).

Altitude	Lidar		MWR		FTIR	
<u>Altitude</u> (km)	<u>Precision</u> Precision-(%)	<u>Resolution</u> Resolution-(km)	<u>Precision</u> Precision-(%)	<u>Resolution</u> Resolution-(km)	<u>Precision</u> (%)	<u>Effective resolution</u> (km)
15	5	1.4	~	~	~	~
20	5	1.2	3	10		
25	3	1.0	3	10		
30	3	1.8	3	10	5–6	8–15
35	3	4.2	3	14		
40	5	7.2	3	14	~	~
45	15	8.6	3	20	~	~
50	55	8.6	3	20	~	~

250 2014). Recent intercomparison work (Smit et al., 2024) indicates that for ECC ozonesondes the relative uncertainty remains in the order of 5–10% throughout much of the profile, but rises at pressures below ~ 13 hPa (~ 30 – ~ 32 km altitude).

For the upper stratosphere (15–60 km), measurements from lidar, FTIR and MWR were used. These instruments extend vertical coverage beyond the range of ozonesondes and provide valuable overlap between 15 km and 30 km. However, global coverage remains limited: only a few long-term lidar (approx. < 5) and MWR (approx. < 4) stations provide regular observations, often with significant delays in data availability. No long-term upper stratospheric observations are available from the southern polar regions.

Lidar systems, typically operating with the Differential Absorption Lidar (DIAL) technique (Godin et al., 1998), provide accurate ozone profiles in clear sky, nighttime conditions of up to about 50 km (Leblanc and McDermid, 2000; Steinbrecht et al., 2006). Usually, 5 to 8 profiles are acquired per month, each requiring several hours of integration depending on the system and the atmospheric conditions. The natural output format of lidar profiles is ozone number density versus geometric altitude.

MWRs retrieve ozone profiles by measuring the pressure-broadened shape of a thermal emission line (Lobsiger et al., 1984; Parrish et al., 1988). Using an optimal estimation technique (Rodgers et al., 1990), these instruments deliver ozone number density profiles typically between 20 km and 60 km altitude. In contrast to lidar, MWRs operate during the day and are less affected by weather conditions. On average, 20 profiles are acquired per month, with integration times between 30 minutes and 5 hours, depending on the instrument (Boyd et al., 2007; Hocke et al., 2007).

An FTIR spectrometer measures ozone profiles between 3 and 42 km by analyzing the interaction of infrared radiation with the atmosphere. The instrument records high-resolution infrared spectra. Ozone absorbs certain wavelengths of infrared light due to its molecular structure. The FTIR spectrometer detects this absorption and converts the information into an ozone spectrum-profile using the Fourier transform method. This conversion allows ozone concentrations to be represented as a function of altitude. Note, FTIR ozone profile retrievals provide independent information typically limited to 2 to 4 degrees

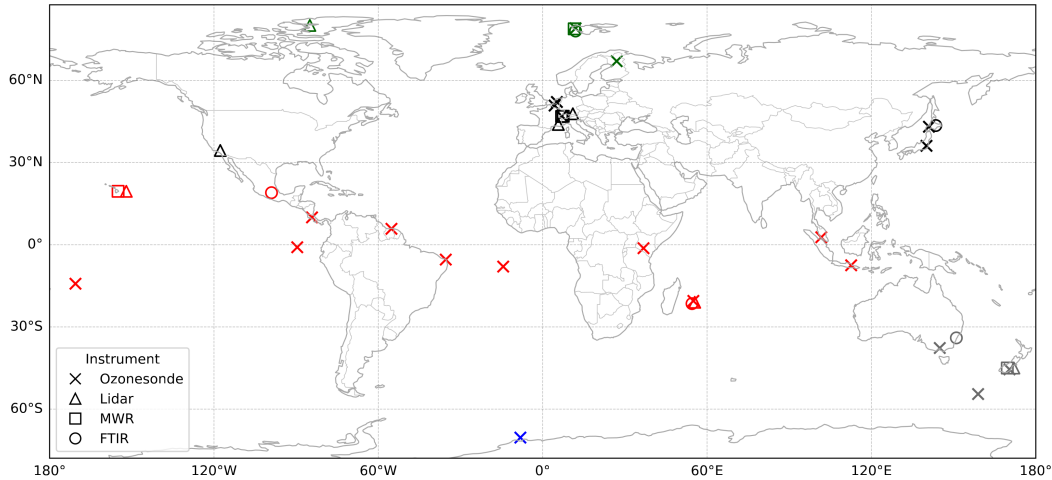


Figure 4. Overview of the stations used for tropospheric and stratospheric quality assessment. Ozone sonde stations are indicated by crosses, lidar stations by triangles, FTIR by circles, and MWR stations by squares. The latitudinal classification is: northern polar stations (green, $67^{\circ}N - 90^{\circ}N$), northern mid-latitude stations (black, $30^{\circ}N - 67^{\circ}N$), tropical stations (red, $30^{\circ}N - 30^{\circ}S$), southern mid-latitude stations (grey, $30^{\circ}S - 70^{\circ}S$), and southern polar stations (blue, $70^{\circ}S - 90^{\circ}S$).

[of freedom, corresponding to an effective vertical resolution of about 8–15 km in the stratosphere.](#) The total error range is from 1.5 to 2.5% between 250 and 3000300 DU ([Garcia et al. \(2022\)](#)) ([Vigouroux et al., 2015](#); [Garcia et al., 2022](#)) without a simultaneous temperature fit, but around 1.5% when a temperature retrieval is included in the data retrieval.

4.1 Ground-based Dataset Description

275 Ground-based ozone sonde observations used in this study are available from the World Ozone and Ultraviolet Data Center (WOUDC, <http://www.woudc.org>) and the NILU Atmospheric Database for Interactive Retrieval (NADIR) hosted by the Norwegian Institute for Air Research (NILU, <http://www.nilu.no/nadir/>). Ground-based lidar, [FTIR](#) and MWR ozone profiles were obtained from the Network for the Detection of Atmospheric Composition Change (NDACC, <http://www.ndsc.ncep.noaa.gov/>). All NDACC lidar, FTIR, and MWR instruments undergo a standardized evaluation process and extensive quality control
280 (Keckhut et al., 2004).

Figure 4 provides an overview of the stations used for this validation. Ozone sonde stations are indicated by crosses, lidar stations by triangles, FTIR stations by circles, and MWR stations by squares. The latitudinal classification is as follows: northern polar stations (green, $67^{\circ}N - 90^{\circ}N$), northern mid-latitude stations (black, $30^{\circ}N - 67^{\circ}N$), tropical stations (red, $30^{\circ}N - 30^{\circ}S$), southern mid-latitude stations (gray, $30^{\circ}S - 70^{\circ}S$) and southern polar stations (blue, $70^{\circ}S - 90^{\circ}S$). For each
285 station, we computed a monthly mean ~~ozone profile~~ [ground-based ozone profile](#), and compared it to the monthly mean level-3 ozone data ~~;~~ obtained with a horizontal distance smaller than 100 km around the location of the ground-based stations.

4.2 Pre-processing of ground-based ozone profile Data and Application of Satellite Averaging Kernels

The ground-based ozone lidar, FTIR, and MWR profiles were pre-processed and harmonized with the satellite retrieval grid in order to enable a consistent comparison. The lidar, [FTIR, and MWR](#) data files contain monthly mean ozone number densities (cm⁻³) between approximately 15 and 50 km. The lidar, [FTIR, and MWR](#) profiles were converted into partial ozone columns expressed in Dobson Units (DU) per layer according to:

$$\text{DU}_{\text{layer}} = \frac{n(z) \Delta z_{\text{cm}}}{2.687 \times 10^{16}}, \quad (6)$$

where $n(z)$ is the ozone number density (in cm⁻³) and Δz_{cm} is the layer thickness in centimeters.

For each month, the corresponding satellite data (retrieved ozone, a priori ozone, averaging kernel, and standard deviation) were extracted from the level-3 NetCDF ozone profile data record files. The satellite quantities, given as ozone column amounts in mol m⁻², were converted into Dobson Units by using $1 \text{ DU} = 4.462 \times 10^{-4} \text{ mol m}^{-2}$. Ozonesonde data, already available in DU, were integrated over the GOME-2 pressure layers. The vertical coordinates of the satellite layers, originally defined in pressure, were converted to geometric altitude using the U.S. Standard Atmosphere (1976).

The averaging kernels (AKs) provided with the satellite product were applied to the ground-based ~~profiles following (Equation 1) in the Dobson Unit space.~~ [ozone profiles from lidar, MRW, and ozonesonde following the standard smoothing approach described by Eq. \(1\). Specifically, the smoothed ground-based profile \$\hat{x}_{\text{gb}}\$ was calculated following Rodgers \(2000\) as \$\hat{x}_{\text{gb}} = x_{\text{a}} + \mathbf{A}\(x_{\text{gb}} - x_{\text{a}}\)\$ where \$x_{\text{gb}}\$ denotes the original ground-based ozone profile, \$x_{\text{a}}\$ the satellite a priori profile, and \$\mathbf{A}\$ the satellite averaging kernel. This operation ensures that the ground-based profiles are smoothed to the vertical sensitivity and smoothing characteristics of the satellite retrieval.](#) All ozone profiles (satellite retrieved, a priori, ground-based, and ~~AK-retrieved-smoothed~~ ground-based) were subsequently interpolated and re-binned onto a uniform 1-km vertical grid up to 60 km altitude. The re-binning of partial columns from the irregular satellite layers to the uniform grid conserved the total amount of ozone in DU by weighting each layer according to the geometric overlap between the native and target bins. The described processing ensures that the ground-based profiles are expressed on the same vertical grid and smoothing characteristics as the satellite retrievals, thus allowing a direct and consistent validation of the satellite ozone profiles. [FTIR ozone profiles are themselves retrieval products characterized by their own averaging kernels and a priori profiles. To ensure a consistent comparison with the level-3 satellite retrievals, a bidirectional averaging-kernel approach is applied following Rodgers and Connov \(2003\). In this framework, the satellite averaging kernels are applied in the Dobson Unit \(DU\) domain to the FTIR profiles while explicitly accounting for the FTIR a priori state. This procedure minimizes inconsistencies that arise from differences in vertical resolution and retrieval characteristics between the two data sets. Such treatment is consistent with established methodology for intercomparison of retrieval products with differing vertical sensitivities \(Calisesi et al., 2005\). A comparable strategy has been applied in previous satellite validation studies involving FTIR data, for example, in the validation of IASI ozone products using GOME-2 and ground-based observations \(Boynard et al., 2018\), where the role of averaging kernels and a priori information in profile comparisons is explicitly addressed. For high-vertical-resolution reference measurements \(lidar and ozonesondes\), only the satellite averaging kernels are applied, as these measurements can be considered close approximations of the true atmospheric state.](#)

For the validation the satellite data as well as the ground-based data were integrated over several altitude bands (e.g. 0–8 km, 8–14 km, 14–30 km, 30–40 km, and 40–50 km for mid-latitude levels). These band-integrated ozone columns were used to obtain time series for the ground-based, satellite, and ~~AK-retrieved~~ smoothed ground-based datasets.

5 Validation of the GOME-2 Metop-A/B/C level-3 Ozone Profile Dataset

325 This section describes the validation of the reprocessed GOME-2A/B/C level-3 (gridded) ozone profile products. The validation periods are as follows: January 2007 to December 2018 for GOME-2A, December 2012 to December 2024 for GOME-2B, and January 2019 to December 2024 for GOME-2C.

The validation is structured around the troposphere, UTLS, lower stratosphere (up to 30 km altitude), and the upper stratosphere. ~~Table 4 defines the vertical ranges used for the troposphere and the lower and upper stratosphere~~ in different latitude bands (see Table 4):

Table 4. Definition of the vertical ranges for the ~~lower and upper stratosphere in different~~ latitude belts.

Region	Troposphere	<u>UTLS</u>	Lower Stratosphere	Upper Stratosphere
Polar Region	0-6 km	<u>6-12</u> km	12–30 km	30–50 km
Mid-Latitudes	0-8 km	<u>8-14</u> km	14–30 km	30–50 km
Tropical Region	0-12 km	<u>12-18</u> km	18–30 km	30–50 km

For the ~~troposphere-troposphere~~, UTLS and lower stratosphere, comparisons are made against ozonesonde measurements. For the upper stratosphere, lidar, FTIR, and microwave radiometer (MWR) data are used as reference. Relative differences are calculated using:

$$RD = \frac{\text{GOME-2 L3} - \text{retrieved ground-based}}{\text{retrieved ground-based}} \frac{O_{3,L3} - O_{3,AK}}{O_{3,AK}} \times 100 \quad (7)$$

335 where $O_{3,L3}$ indicates the ozone values in the gridded product, and $O_{3,AK}$ indicates the ground-based reference ozone profile (either lidar, MWR, FTIR or balloon based) that is smoothed with the averaging kernel.

Table 5 summarizes the ~~differences between L3 gridded ozone profiles and the validation statistics for the different latitude belts, based on all available~~ ground-based reference ~~data~~ stations shown in Figure 4. The statistics include contributions from lidar, microwave radiometer (MWR), FTIR, and ozonesonde observations, depending on data availability in the respective altitude ranges. For the upper stratospheric validation in the tropics, the analysis was limited to the period up to the end of 2022 due to the reduced availability of ground-based observations after 2022. The relative difference is generally below 15% in the troposphere, except in the tropical troposphere, and both the lower and upper stratosphere for all latitude bands. For the southern polar stations, there was no information available for the upper stratospheric part. A station-to-station consistency analysis was performed using the mid-latitude stations as a representative example to assess the robustness of the validation results across independent reference data sets.

Table 5. GOME-2A/B/C gridded validation statistics (Absolute Difference [AD], Relative Difference [RD], and standard deviation [STDEV]) for the troposphere, UTLS-zone, lower and upper stratosphere across latitude regions: Units: AD in DU, RD and STDEV in %.

Region	Troposphere			UTLS-zone			Lower Stratosphere			Upper Stratosphere		
	AD	RD	STDEV	AD	RD	STDEV	AD	RD	STDEV	<u>AD</u>	<u>RD</u>	<u>STDEV</u>
<i>Metop A</i>												
Northern Polar	1.3	9.7	8.0	<u>3.6</u>	<u>15.9</u>	<u>16.9</u>	-6.1	-3.9	3.7	-0.2	-7.8	0.2
Northern Mid-Lat.	2.3	14.9	12.3	<u>4.1</u>	<u>26.0</u>	<u>26.2</u>	-6.7	-4.5	4.0	-0.1	-2.8 <u>-2.6</u>	0.2
Tropics	3.6	27.3	25.5	<u>0.1</u>	<u>2.2</u>	<u>6.2</u>	-3.6	-3.5	3.1	-0.3 <u>-0.2</u>	-3.9 <u>-5.2</u>	0.1
Southern Mid-Lat.	0.7	6.7	13.9	<u>1.1</u>	<u>11.9</u>	<u>24.9</u>	-1.6	-1.1	3.3	-0.1	-1.1	0.1
Southern Polar	-0.6	-7.4	11.1	<u>-1.4</u>	<u>-8.9</u>	<u>14.2</u>	0.5	0.4	4.1	-	-	-
<i>Metop B</i>												
Northern Polar	1.2	8.6	9.5	<u>1.5</u>	<u>8.6</u>	<u>23.5</u>	-4.5	-3.0	4.9	-0.3	-10.5	0.2
Northern Mid-Lat.	2.3	14.7	10.7	<u>4.0</u>	<u>23.1</u>	<u>19.6</u>	-5.9	-3.9	3.2	-0.1	-3.6 <u>-3.2</u>	0.2
Tropics	2.8	21.9	18.7	<u>0.4</u>	<u>5.1</u>	<u>5.1</u>	-2.7	-2.6	2.2	-0.1 <u>-0.2</u>	-3.5 <u>-5.1</u>	0.1
Southern Mid-Lat.	0.5	5.3	10.0	<u>0.1</u>	<u>5.5</u>	<u>18.6</u>	-0.1	-0.2	2.3	-0.1	-1.6	0.1
Southern Polar	-0.8	-9.5	8.2	<u>-2.5</u>	<u>-15.0</u>	<u>10.9</u>	1.1	0.9	3.8	-	-	-
<i>Metop C</i>												
Northern Polar	0.5	3.5	9.0	<u>-0.1</u>	<u>2.4</u>	<u>26.9</u>	-3.3	-2.5	6.1	-0.4	-13.5	0.2
Northern Mid-Lat.	2.4	15.0	10.4	<u>4.9</u>	<u>30.8</u>	<u>20.6</u>	-6.8	-4.7	2.9	-0.2	-4.2 <u>-2.8</u>	0.2
Tropics	7.5	57.6	28.0	<u>0.4</u>	<u>5.3</u>	<u>6.8</u>	-7.0	-6.8	2.4	-0.2 <u>-0.3</u>	-4.2 <u>-7.3</u>	0.1
Southern Mid-Lat.	1.1	11.8	15.7	<u>1.7</u>	<u>19.0</u>	<u>25.1</u>	-2.1	1.7 <u>-1.7</u>	2.8	-0.1	-1.1	0.2
Southern Polar	-0.6	-7.5	8.8	<u>-1.6</u>	<u>-9.6</u>	<u>12.8</u>	-0.1	0.1 <u>-0.1</u>	3.1	-	-	-

In the 30–40 km altitude range, the mean relative difference between the satellite product and the smoothed lidar observations at mid-latitude sites is approximately –3.5%, with a station-to-station spread of about –3% to –5%. Comparisons with smoothed MWR observations yield a slightly smaller mean bias of about –2.8%, with a comparable spread (–1.2% to –4%).

350 In the 40–50 km layer, the satellite–smoothed lidar differences increase to around –4%, while the satellite–smoothed MWR differences remain closer to –2%. Despite these differences in magnitude, the results from lidar and MWR observations are consistent in sign and order of magnitude, and all mean biases remain within approximately ±5% for the mid-latitude stations.

355 Comparable bias magnitudes are found at the tropical sites (Figure 4) for the lidar, FTIR, and MWR station. However, the Réunion lidar data above 40 km exhibit increased variability and reduced consistency compared to the other stations. As a result, these data are treated with caution and are not used as primary reference in the uppermost altitude range.

A comparison with the original (unsmoothed) ground-based data reveals substantially larger station-to-station variability and, in some cases, even a change in the sign of the bias. Applying the averaging-kernel-based smoothing markedly reduces

these discrepancies, highlighting the importance of accounting for differences in vertical resolution in satellite-ground-based comparisons.

360 5.1 Validation results for the troposphere, UTLS and lower ~~stroposphere~~stratosphere

To assess the consistency of the GOME-2A/B/C level-3 ozone profile product in the troposphere, UTLS-zone and lower stratosphere, we compare the ~~integrated ozone values for the tropospheric and lower stratospheric part of GOME-2A/B/C~~ vertically integrated ozone partial profiles with the corresponding pressure levels derived from ozonesonde data, following (Table 4). The averaging kernels are applied to high-resolution ozone profiles and those values are used in the evaluation. Table 5 summarizes
365 the validation statistics as function of latitude and vertical regimes specified the three instruments respectively. In addition, Figure 5 ~~shows the outcome for~~ supplements the validation results from the mid-latitude stations in terms of time-series approach for the different latitude belts.

In the troposphere, all three GOME-2 instruments show predominantly positive biases in the Northern Hemisphere, with relative differences reaching 14–15% in the northern mid-latitudes for all sensors and up to 57.6% in the tropics for Metop-C.
370 Absolute differences in the northern mid-latitudes are consistently around 2–2.4 DU. Tropical biases are moderate for Metop-A and ~~Table 5 summarizes the results for the time periods under consideration for the different latitude belts. The comparison with~~ balloon sondes ~~Metop-B (22–27%) but substantially larger for Metop-C as earlier mentioned. In the southern mid-latitudes,~~ positive biases are smaller (5–12%), while systematic negative biases are observed in the southern polar region (-7% to -10%).
Overall, the tropospheric results indicate consistent Northern Hemisphere overestimation and enhanced variability compared
375 to the UTLS, with notable inter-sensor differences in the tropics. For the UTLS region, all three GOME-2 instruments show consistent hemispheric differences. Positive biases are observed in the Northern Hemisphere, particularly in the northern mid-latitudes, where relative differences reach +26.0% (Metop-A), +23.1% (Metop-B), and +30.8% (Metop-C), corresponding to absolute differences of 4–5 DU. For northern polar stations, the biases are smaller but remain positive for Metop-A and ~~Metop-B, while Metop-C shows near-neutral absolute bias with elevated variability. In the tropics, all sensors exhibit small~~ positive differences (2–5%) and low variability, indicating stable performance. In contrast, systematic negative biases are found
380 in the southern polar region (-9% to -15%). For the lower stratosphere, the intercomparison with ozonesondes shows that all three sensors ~~have comparable results and are all produce comparable results, with relative differences within 10% for the lower stratosphere. We can observe a seasonal dependency, present. A seasonal dependence is observed in the dataset; this was also the case in the level,~~ consistent with the Level 2 product. Delcloo et al. (2024) did also describe this seasonal dependency for
385 both, ~~also reported this seasonal dependence in both~~ the lower stratosphere and the troposphere. In the lower stratosphere, all three instruments exhibit systematic negative biases in the Northern Hemisphere and tropics, with relative differences typically between -2.5% and -6.8% and absolute differences of -3 to -7 DU. The strongest negative biases are observed in the northern mid-latitudes for all sensors. In contrast, the southern polar region shows small positive or near-neutral differences (0–1%), while southern mid-latitude biases remain weakly negative. Variability is comparatively low across all latitude bands and
390 markedly smaller than in the troposphere ~~. For the troposphere, we see among all sensors the same behavior, an overestimation of the tropospheric ozone part in the Tropical belt (Deleloo, 2024). Here we note that in this specific region, the retrieved ozone~~

concentrations shows very high differences for the GOME-2C sensor, also in the lower troposphere (see Table 5), for which we have no explanation yet.

395 and UTLS. Overall, the lower stratospheric results demonstrate consistent underestimation in most regions but relatively stable performance with limited dispersion. For the first two years of the GOME-2A mission, there was a particular instrumental integration time setting for the spectral bands ~~1A and 1B~~ 1a and 1b in the far- and near-UV regions, that made it difficult to distinguish the troposphere from the stratosphere on small ground pixels. This was changed by the end of 2008 (also listed in Table 1). This ~~inconvenient~~ instrument setting has affected the ozone profile across the full atmosphere for the first two years, but this effect should not cause the sustained bias for the first four years of the mission, as seen in Figure 5. Our hypothesis is
400 that the degradation correction, which spans multiple years in order to get the seasonal effect right, has a lingering effect.

5.2 Validation results for the upper stratosphere

To assess the consistency of the GOME-2 monthly mean level-3 ozone profile product in the upper stratosphere, monthly averaged satellite partial columns were compared with ground-based observations. In this altitude range, ground-based observations from lidar, (MWR), and FTIR provide complementary information with different vertical sensitivities (see Table 3). The analysis focuses on the 30–50 km altitude range, where nadir-viewing UV instruments provide good sensitivity to stratospheric ozone due to Rayleigh scattering. However, when comparing vertically resolved ozone profiles or partial columns in narrow altitude ranges, differences in vertical resolution and smoothing between satellite and ground-based observations need to be taken into account. Therefore, the application of the averaging kernel is ~~essential due to the limited vertical sensitivity of nadir-viewing instruments.~~ required to ensure a consistent comparison between satellite and ground-based profile measurements. This is
410 consistent with previous studies showing that averaging-kernel effects are less critical for total or broad stratospheric column ozone but become important for profile-based or layer-resolved comparisons (e.g., Liu et al., 2010; Keppens et al., 2024). The results of the comparison for mid-latitude and tropical stations are presented in Figure 6 and Figure 7, respectively. The mean ~~retrieved smoothed~~ ground-based profiles (left panel) closely follow the satellite retrievals, particularly between 40 and 50 km.

The time series of monthly mean ozone partial columns between 30–40 km and 40–50 km (middle panels) show good agree-
415 ment in both long-term variability and seasonal cycles, with maxima in summer and minima in winter. In the 40–50 km layer, the overall seasonal behavior is also reproduced, although a systematic phase shift between the satellite and ground-based time series is apparent. Although a bias in absolute values is apparent between the satellite and ground-based datasets, the satellite product reliably reproduces the temporal evolution of ozone in the upper stratosphere. The correlation coefficients for each altitude range are indicated in the middle panels. The higher correlation between the ~~retrieved smoothed~~ ground-based and level-3 time series, compared to that between the original ground-based and level-3 data, highlights the effectiveness of the ~~a priori and averaging kernel corrections and confirms the vertical representativeness of the satellite product. The correlation coefficients for each altitude range are indicated in the middle panels.~~ apriori and averaging kernel-based smoothing in taking into account differences in vertical resolution. Here r denotes to the correlation coefficient calculated from the monthly mean time series. In the 40–50 km altitude range, the combination of narrow layer integration, occasional crossings
425 of satellite and ground-based profiles, and the increased influence of the apriori profile can lead to reduced or even negative

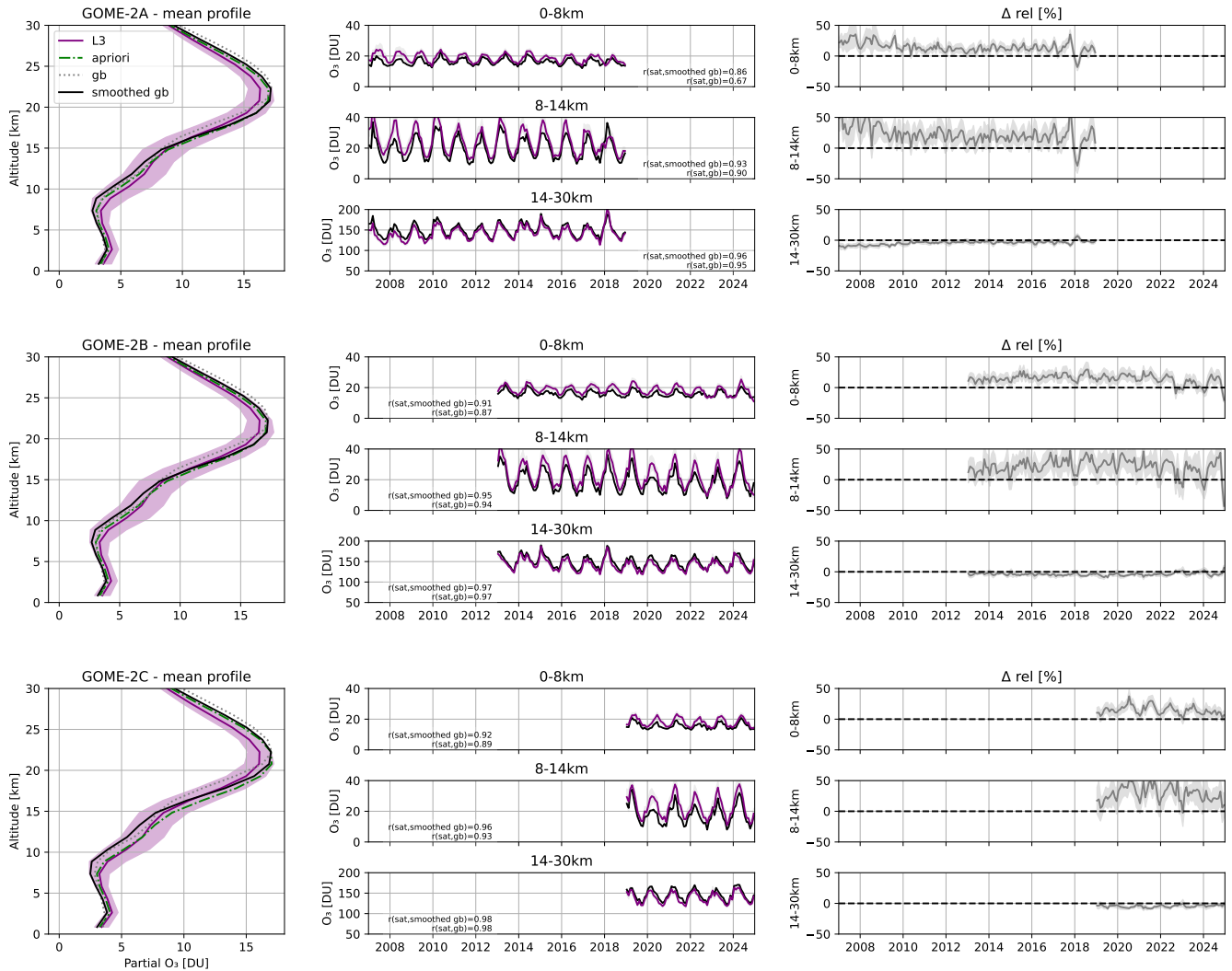


Figure 5. Comparison of ozone partial columns derived from satellite and ground-based observations for GOME-2A/B/C. For each sensor (from top to bottom), the three panels show: (left) the mean vertical profile of ozone partial columns in Dobson Units (DU) averaged over the entire observation period, including the [L3-level-3](#) satellite retrieval (purple), the ozonesonde measurement (grey), the ozonesonde profile [adjusted-smoothed](#) by the satellite averaging kernel and a priori profile (black), the [L3-level-3](#) a priori profile (green), and the $\pm 1\sigma$ uncertainty of the satellite profile (shaded); (middle) the time series of monthly mean ozone partial columns in the 0–8 km altitude range (troposphere) and in the 14–30 km altitude range (lower stratosphere) for the same data sources; and (right) the relative difference between the [L3-level-3](#) and [adjusted-smoothed](#) ozonesonde data in this altitude range. This comprehensive comparison illustrates the consistency between satellite and ground-based ozone observations, as well as systematic differences over time and altitude.

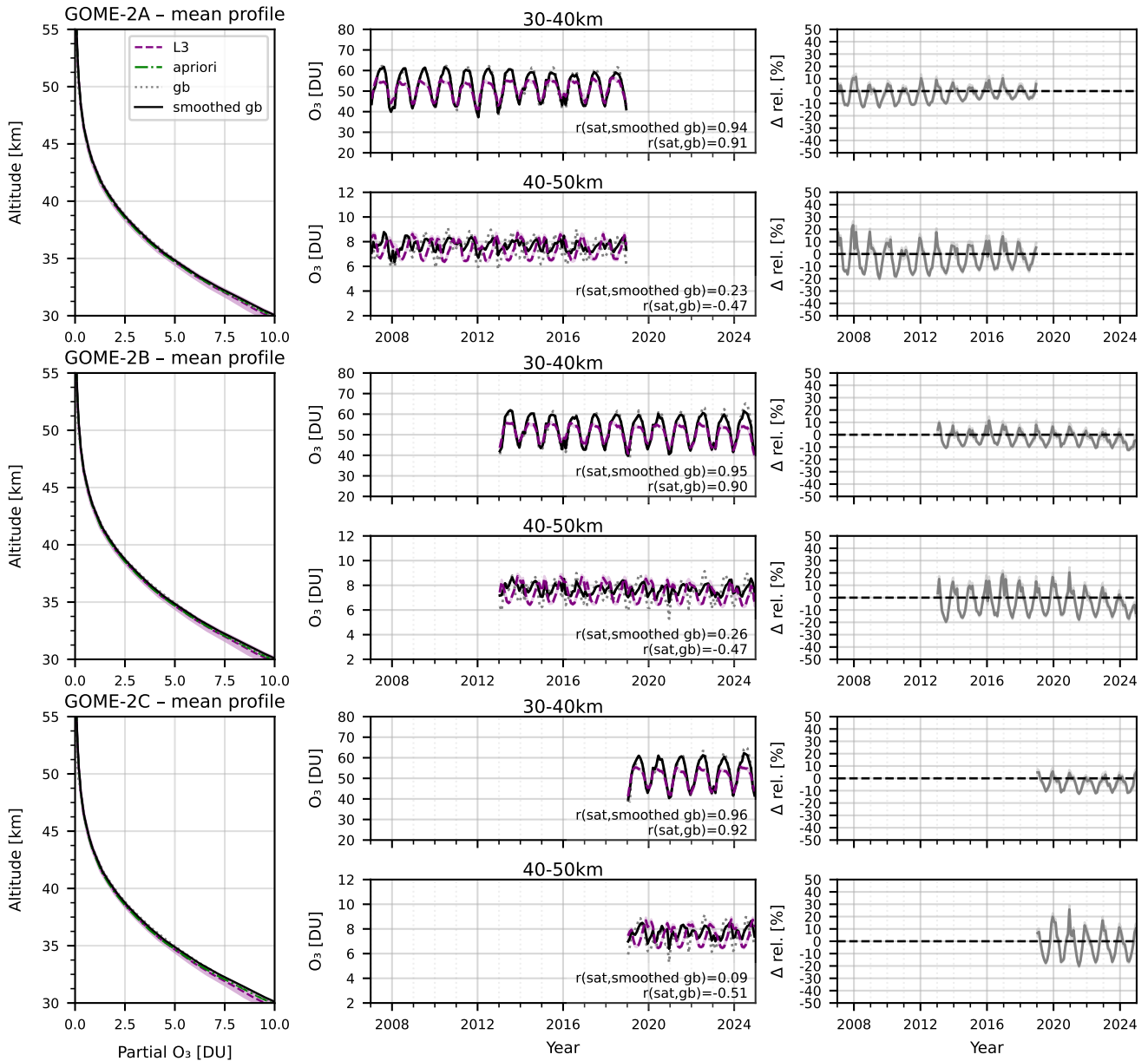


Figure 6. Comparison of ozone partial columns derived from satellite and all mid-latitude mid-latitude ground-based observations as shown in Figure 4, for GOME-2A/B/C. For each sensor (from top to bottom), the three panels show: (left) the mean vertical profile of ozone partial columns in Dobson Units (DU) averaged over the entire observation period, including the L3-level-3 satellite retrieval (purple), the lidar ground-based measurement (grey), the retrieved lidar-smoothed ground-based profile (black), the L3-level-3 a priori profile (green), and the $\pm 1\sigma$ uncertainty of the satellite profile (shaded); (middle) the time series of monthly mean ozone partial columns in the 30–40 km and 40–50 km altitude range for the same data sources with the Pearson correlation coefficient (r) between the level-3 and ground-based as well as smoothed ground-based time series; and (right) the relative difference between the L3-level-3 and retrieved lidar-smoothed data in this altitude range. This comprehensive comparison illustrates the consistency between satellite and ground-based ozone observations, as well as systematic differences over time and altitude.

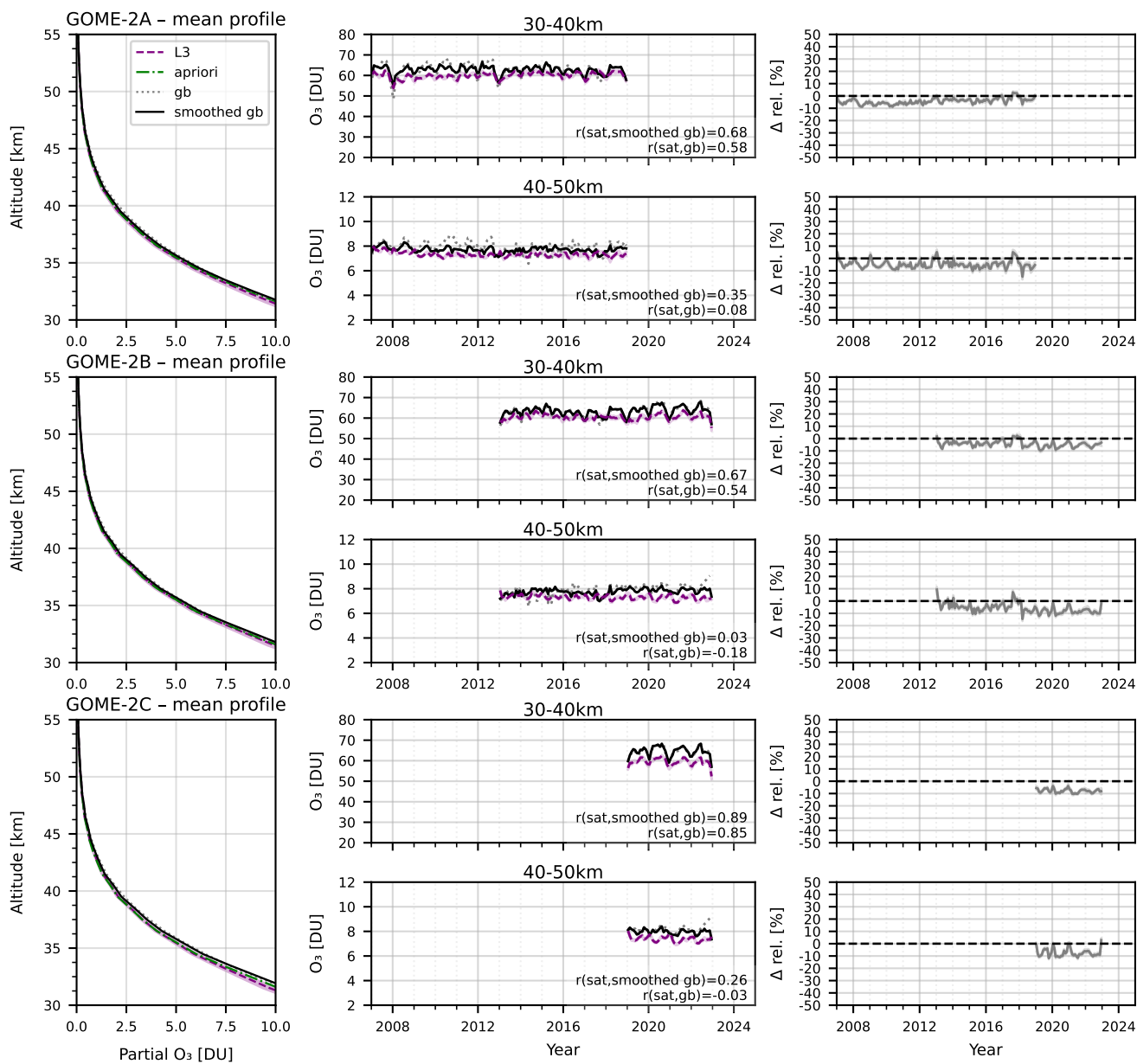


Figure 7. Same as Figure 6 but for all tropical ground-based stations as shown in Figure 4.

correlation coefficients, despite the overall consistency of the mean profiles. Such behavior is a known feature of nadir profile retrievals, where the vertical sensitivity and averaging kernels redistribute information across the profile and relative to the a priori, affecting the apparent temporal phasing and trend structure when compared to high-vertical-resolution references (Keppens et al., 20215; Rivoire et al., 2025).

430 The relative bias (left panels in Figure 6 and Figure 7) between the three satellite sensors and the adjusted-lidar-smoothed
ground-based data (right panels) remains within $\pm 10\%$ in the 30–40 km altitude range. In the 40–50 km region, the relative differences are larger, around 20%, particularly during periods of ozone maxima and minima. Nevertheless, the mean relative difference, as reported in Table 5, is below 5%. Overall, the lidar-based-ground-based validation confirms the robustness of the GOME-2 level-3 ozone profile product and its suitability for cross-validation with other satellite datasets and the construction
435 of long-term ozone time series.

5.3 Inter sensor differences

Figure 8 illustrates the inter-sensor differences between the three GOME-2 instruments respectively in the upper stratosphere for the altitude ranges 30–40 km (gray line) and 40–50 km (black line). Figure 9 is showing this for the troposphere within the altitude range 0-8 km (gray line) and for the lower stratosphere within the altitude range 14-30 km, averaged across all mid-
440 latitude stations. The comparison reveals a high degree of consistency among the sensors, with differences generally within 2. The smallest discrepancies are observed in the 40–50 km layer, indicating robust long-term stability in this altitude range. It is also shown in Figure 9 that for GOME-2A, the data after 2018 is not consistent anymore and should not be used after that date.

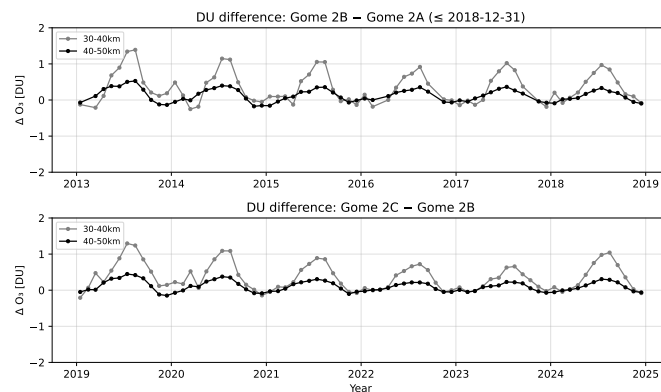


Figure 8. Time series of differences in partial ozone columns (DU) between the GOME-2 satellite pairs for two altitude ranges. **Top:** GOME-2B minus GOME-2A, limited to the period up to 31 December 2018 due to the reduced data quality of GOME-2A. **Bottom:** GOME-2C minus GOME-2B over the respective overlapping time period. Both panels show the altitude ranges 30–40 km (gray line) and 40–50 km (black line).

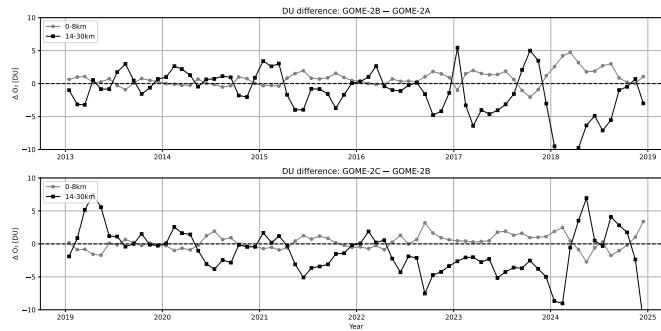


Figure 9. Time series of differences in partial ozone columns (DU) between the GOME-2 satellite pairs for two altitude ranges. **Top:** GOME-2B minus GOME-2A, limited to the period up to 31 December 2018 due to the reduced data quality of GOME-2A. **Bottom:** GOME-2C minus GOME-2B over the respective overlapping time period. Both panels show the altitude ranges 0–8 km (gray line) and 14–30 km (black line).

6 Conclusions

445 A comparison study of ozone retrieval statistics from the GOME-2 instruments onboard of Metop-A, -B, and -C shows a high level of consistency across latitude bands and atmospheric layers, confirming the maturity and stability of the retrieval algorithm. In the **troposphere**, all three sensors tend to slightly overestimate ozone, with absolute differences (AD) of roughly +1 to +3 DU in mid-latitudes and up to +7.5 DU in the tropics for Metop-C. Relative differences increase toward lower latitudes, reflecting higher retrieval uncertainty under variable cloud and humidity conditions. **The standard deviations (STDEV) of 8–15**

450 **DU across all sensors indicate robust internal consistency.**

In the UTLS, GOME-2 instruments show positive biases in the Northern Hemisphere, largest in mid-latitudes, and negative biases in the southern polar region. Tropical biases are small (2 – 5%) with low variability, while mid- and high-latitude variability is higher (up to 27%). In the **lower stratosphere**, all sensors show a small negative bias, typically between -3 and -7 DU (RD \approx -3% to -7%), corresponding to a modest underestimation of ozone concentrations. The lowest scatter is observed

455 **for Metop-B (STDEV \approx 2–3), suggesting particularly stable performance. Regional patterns are coherent across sensors, with slightly weaker negative biases in the tropics and southern mid-latitudes compared to the northern hemisphere.**

In the upper stratosphere, biases are minimal across all sensors, with absolute difference values, close to zero (-0.1 to -0.4 DU) and extremely low variability (STDEV \approx 0.1–0.2-DU). These results confirm excellent consistency in the upper-stratospheric ozone retrievals, where radiometric calibration and spectral fitting are least affected by tropospheric interference.

460 **Overall, the three Metop GOME-2 instruments provide mutually consistent ozone profiles with small and systematic inter-satellite differences. The main distinctions are a stronger positive tropospheric bias in Metop-C—particularly in the tropics—and slightly smaller negative deviations in the lower stratosphere for Metop-B. These findings underscore the reliability of the GOME-2 constellation for long-term ozone trend analyses and the potential for merged multi-sensor time series without significant inter-calibration artifacts.**

465 The observed inter-sensor consistency across Metop-A, -B, and -C supports the creation of a coherent, multi-decadal ozone
climate data record. Minor biases in the troposphere and lower stratosphere are small relative to natural ozone variability
and can be corrected through standard inter-calibration procedures, ~~ensuring that combined datasets remain robust for trend
470 detection and model evaluation~~. The minimal deviations in the upper stratosphere further reinforce the stability of GOME-2
instruments over time, making the Metop constellation a reliable backbone for long-term monitoring of ozone distribution,
stratosphere-troposphere interactions, [model evaluations](#) and assessment of policy-driven ozone recovery. In conclusion, these
results confirm the use of merged Metop GOME-2 datasets in constructing consistent, high-quality ozone time series, suitable
for climate and atmospheric research.

Data availability. The level-3 gridded ozone profile data record used for this study can be obtained via the AC SAF helpdesk or via links
on the AC SAF website (<https://www.acsaf.org/>), or via the EUMETSAT Data Center (EUMDC) (<https://user.eumetsat.int/data-access/>)
475 data-centre). Its reserved DOI is: 10.15770/EUM_SAF_AC_0052

1 [Appendix Averaged Averaging Kernels by latitude band](#)

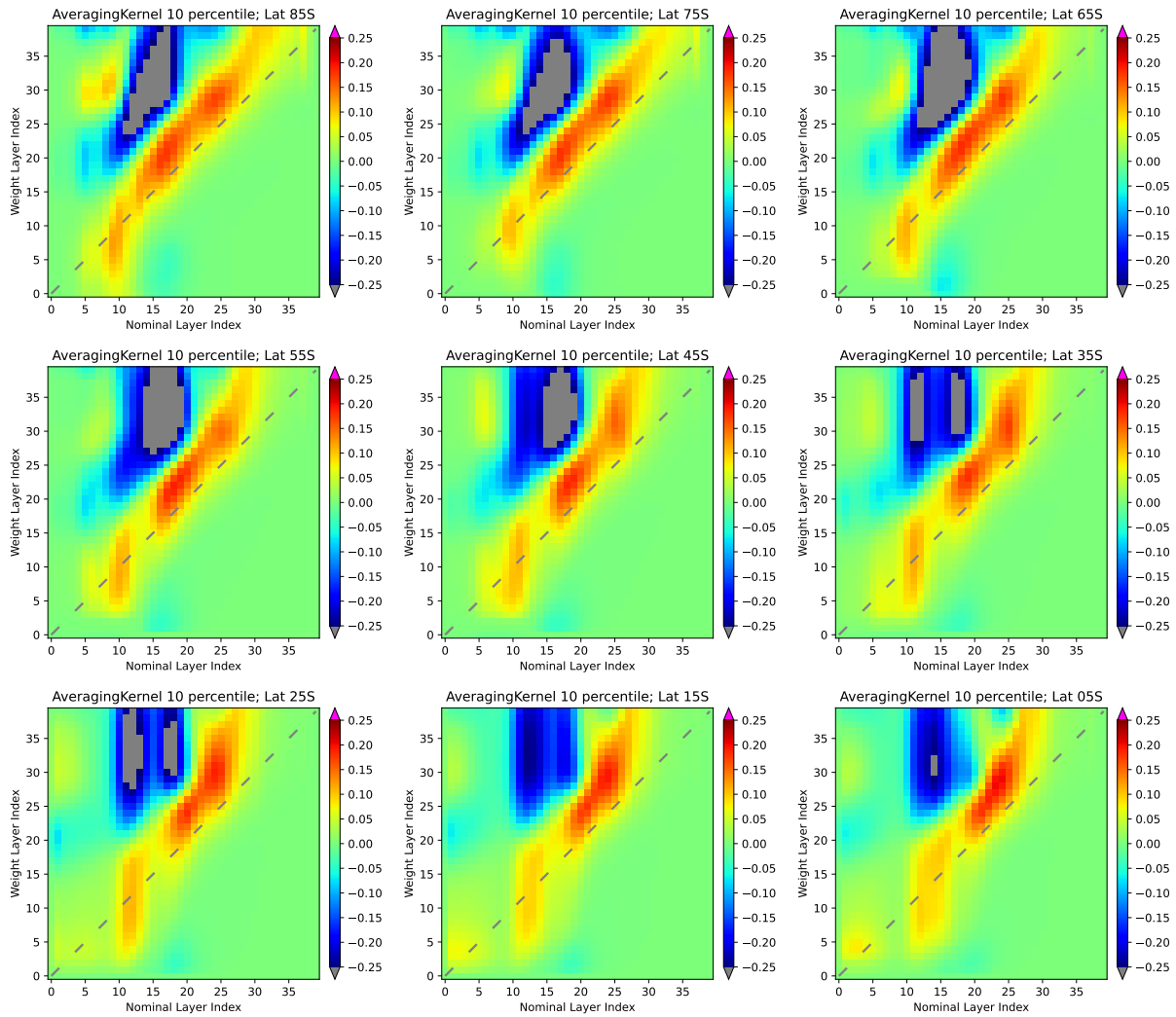


Figure A1. 10th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Northern Hemisphere).

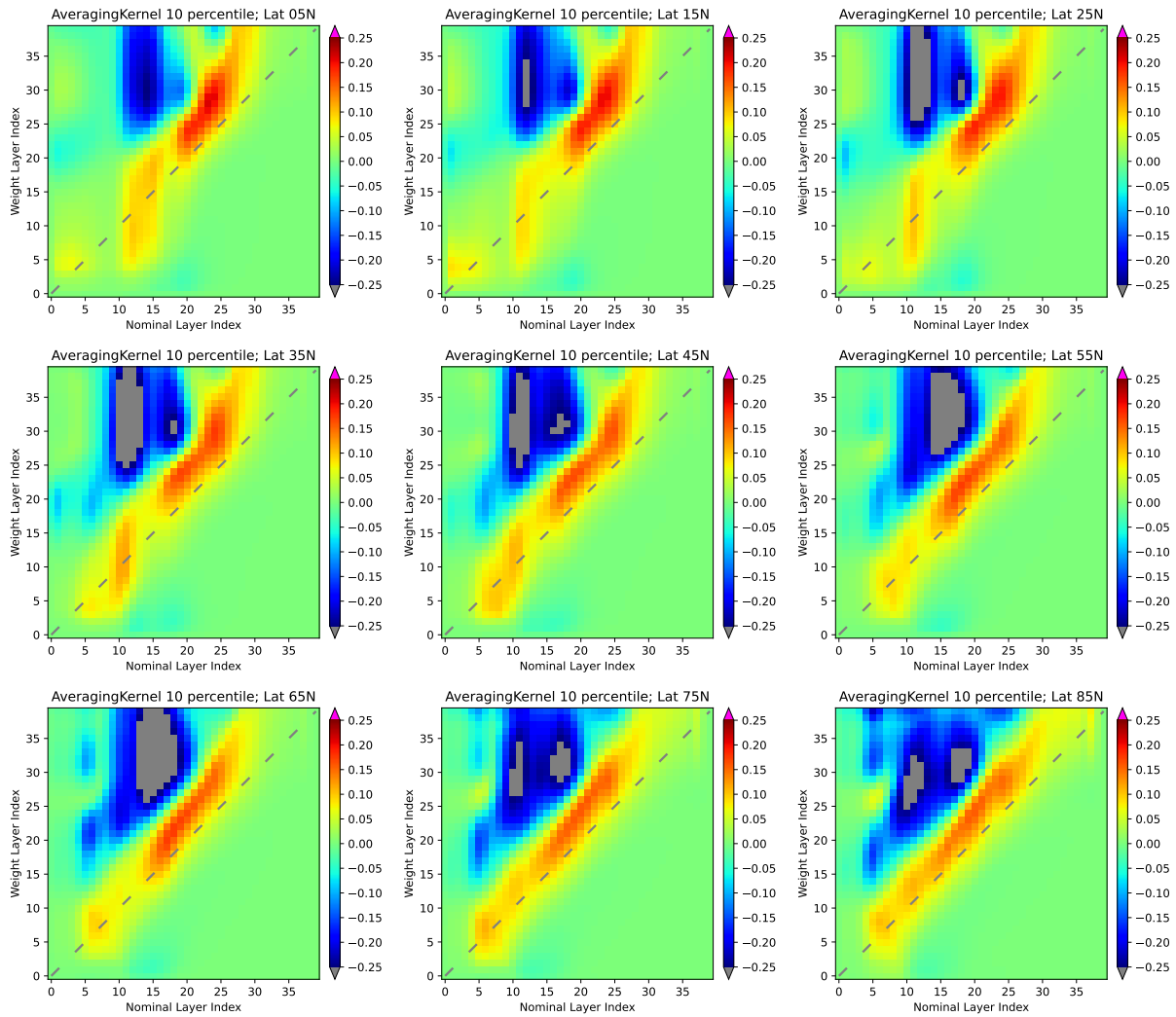


Figure A2. 10th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Southern Hemisphere).

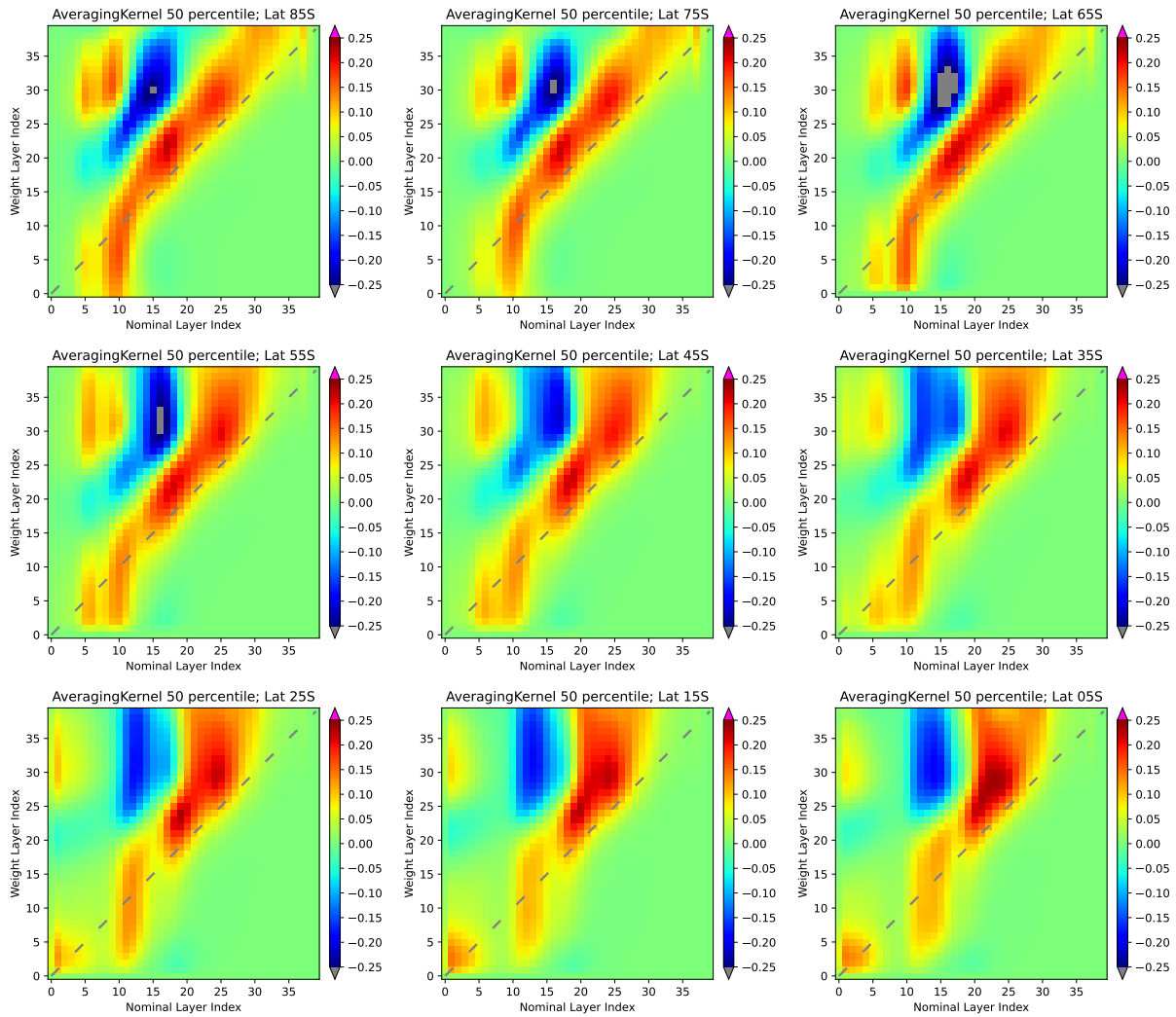


Figure A3. 50th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Northern Hemisphere).

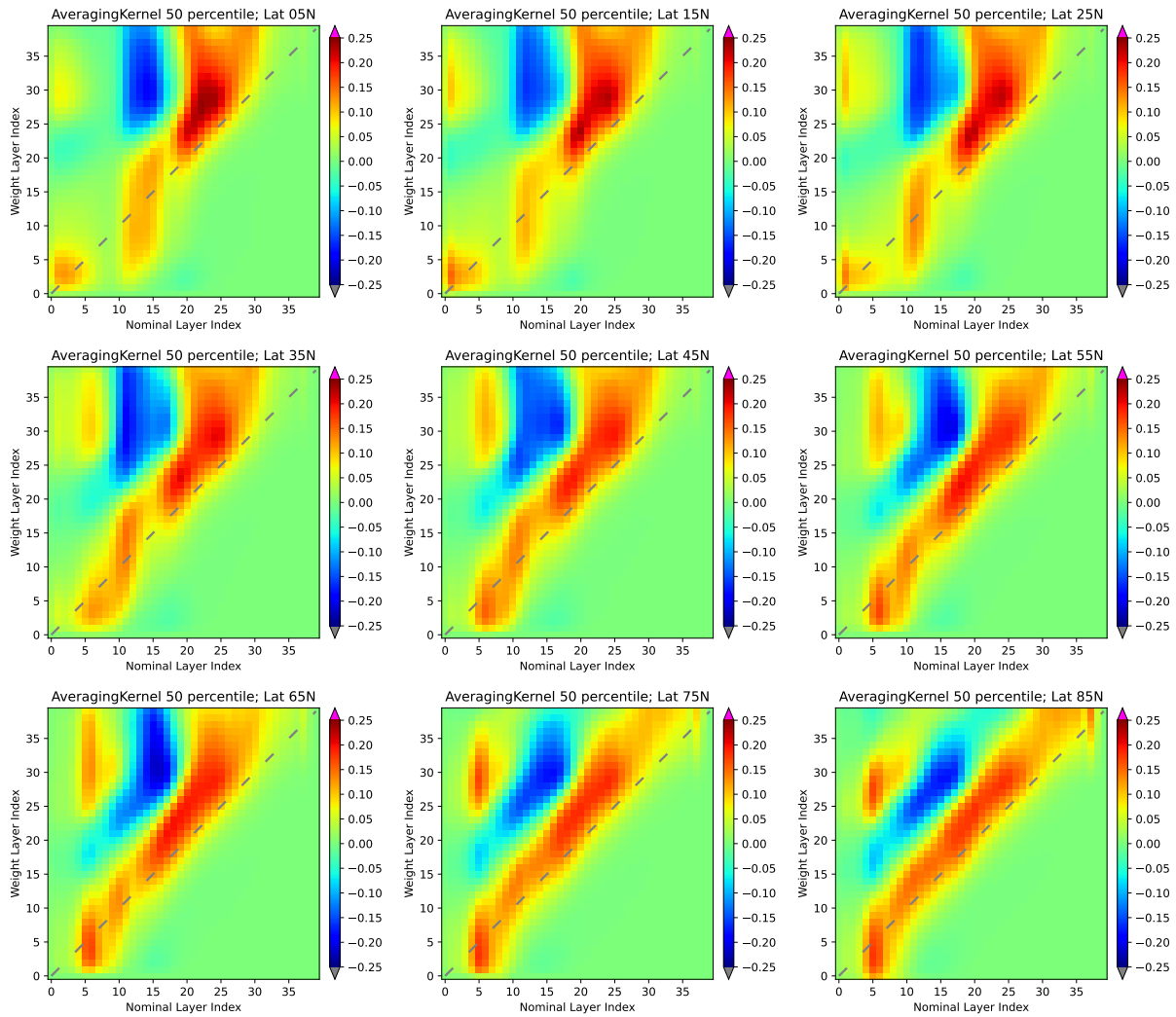


Figure A4. 50th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Southern Hemisphere).

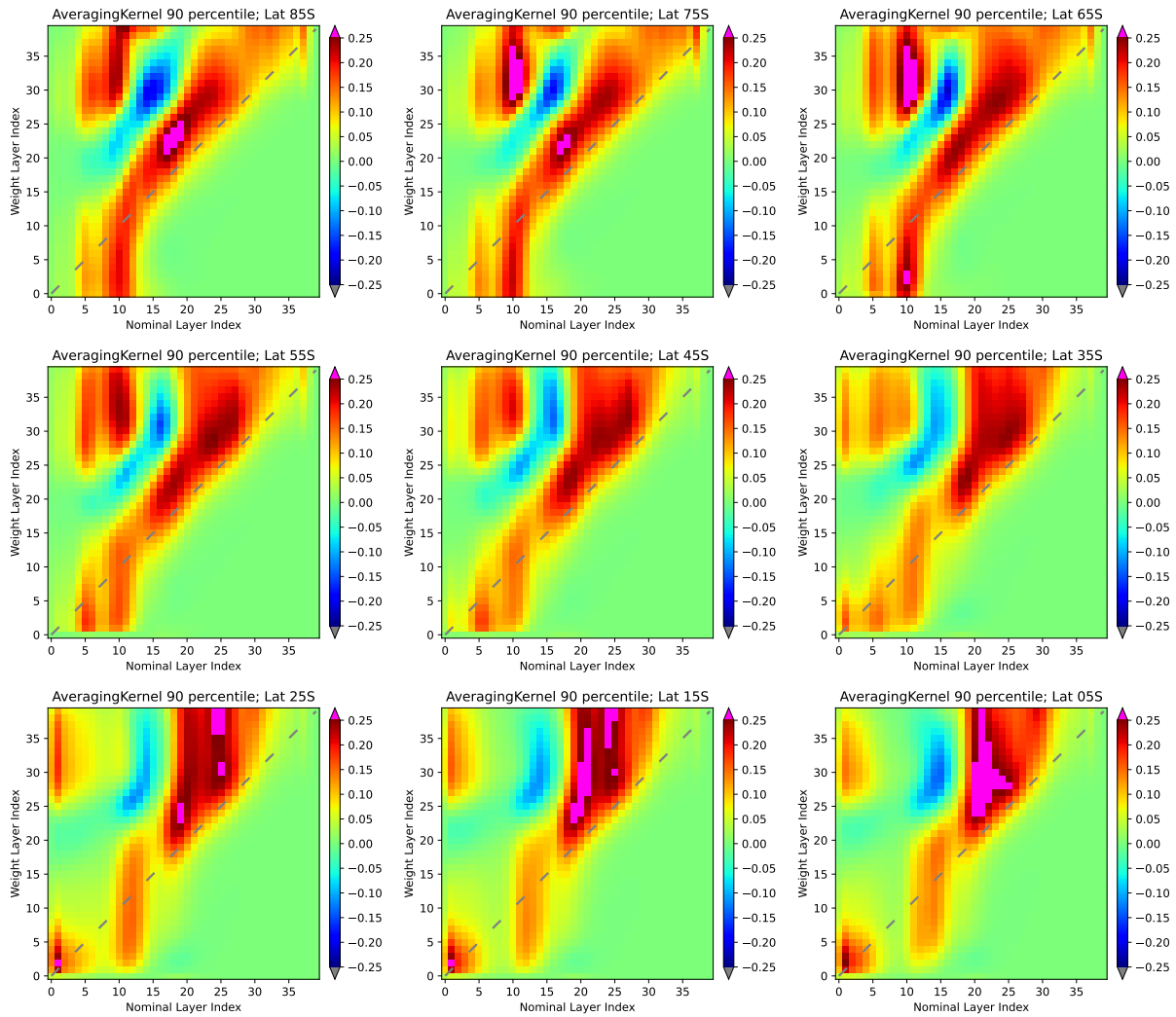


Figure A5. 90th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Northern Hemisphere).

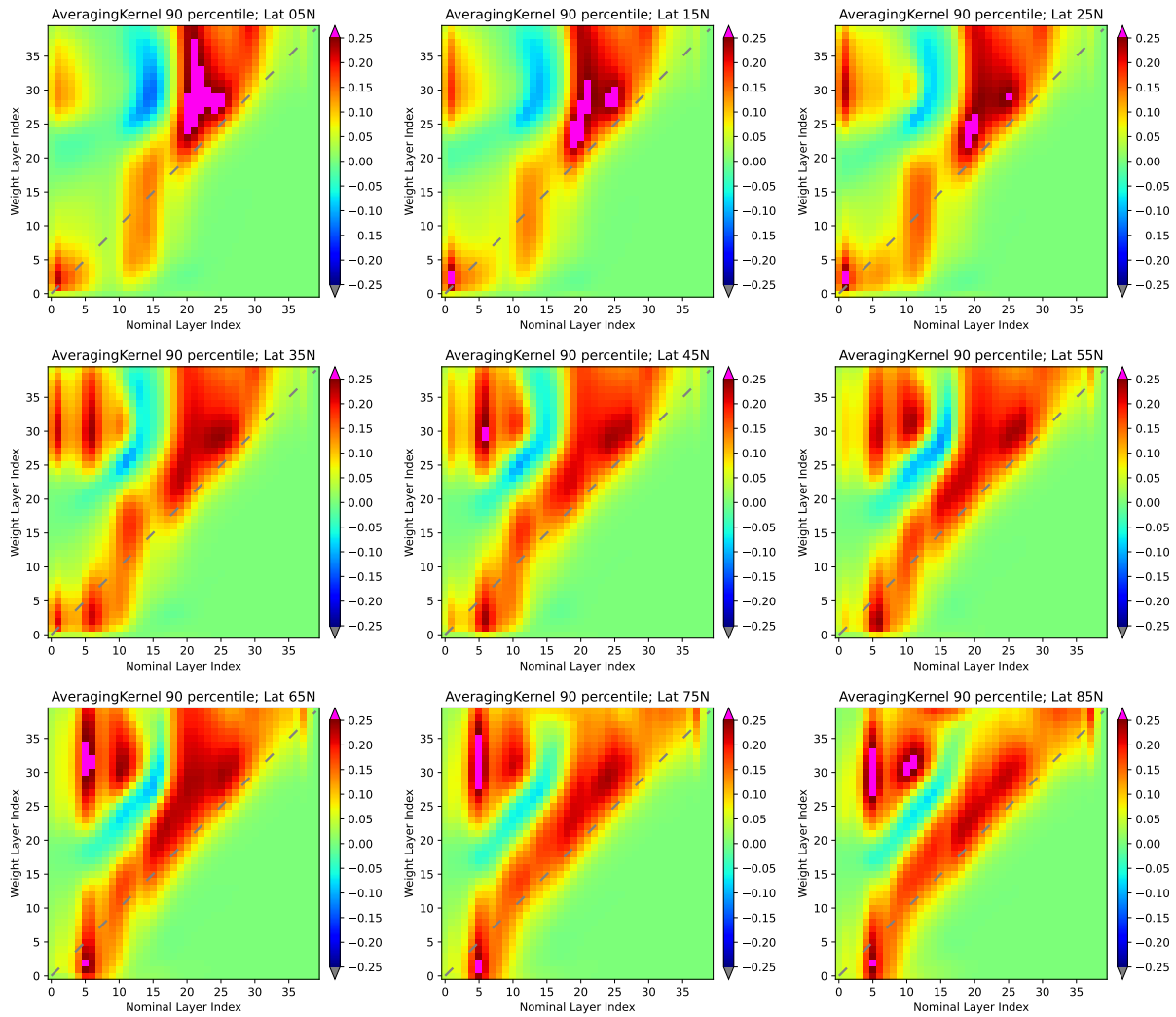


Figure A6. 90th percentile of the averaged averaging kernels of GOME-2A by 10-degree latitude band in 2009 (Southern Hemisphere).

Author contributions. OT developed the algorithm, produced the level-2 and gridded the level-3 data. AD and PTA did the analyses of the gridded data and compared them with balloon sondes and LIDAR & microwave respectively. All authors contributed to the interpretation of the results. ground-based ozone reference data (lidar, MWR, FTIR, and ozonesonde observations) were collected and processed by the
480 respective NDACC station Principal Investigators.

Competing interests. There are no competing interests.

Acknowledgements. The authors would like to thank the AC SAF / EUMETSAT for funding the study via the CDOP-4 project, and their on-going support though GOME-2 L1b data provision. Author AD has received additional funding from PRODEX, contract number 4000140643 B-ACSAF. ~~The~~ Ground-based ozone reference data used in this publication study, including ozone lidar, microwave radiometer (MWR),
485 FTIR, and ozonesonde observations, were obtained from the Network for the Detection of Atmospheric Composition Change (NDACC) and are available through the NDACC website (NDACC, <https://www.ndacc.org>). We thank the station Principal Investigators for providing and maintaining the long-term ozonesonde records used in this study.

References

- Arosio, C., Sofieva, V., Orfanoz-Cheuquelaf, A., Rozanov, A., Heue, K.-P., Loyola, D., Malina, E., Stauffer, R. M., Tarasick, D., Van
490 Malderen, R., Ziemke, J. R., & Weber, M. (2025). *Intercomparison of tropospheric ozone column datasets from combined nadir and limb satellite observations*. *Atmospheric Measurement Techniques*, 18, 3247–3265. doi:10.5194/amt-18-3247-2025
- Bognar, K., Tegtmeier, S., Bourassa, A., Roth, C., Warnock, T., Zawada, D. and Degenstein, D. (2022). Stratospheric ozone trends for 1984–2021 in the SAGE II–OSIRIS–SAGE III/ISS composite dataset. *Atmospheric Chemistry and Physics*, 22(14), 9553–9569.
- Boyd, I. S., A. D. Parrish, L. Froidevaux, T. von Clarmann, E. Kyrölä, J. M. Russell III, and J. M. Zawodny (2007), Ground-based microwave
495 ozone radiometer measurements compared with Aura-MLS v2.2 and other instruments at two Network for Detection of Atmospheric Composition Change sites, *J. Geophys. Res.*, 112, D24S33, doi:10.1029/2007JD008720.
- [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.](#)
- [Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weissenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M. and Perner, D. \(1999\). The global ozone monitoring experiment \(GOME\): Mission concept and first scientific results. *Journal of the Atmospheric Sciences*, 56\(2\), 151–175.](#)
- [Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., Kiessling, W., O'Connor, M., Pandolfi, J.M., Parmesan, C., Schwing, F.B., Sydeman, W.J., and Richardson, A. J. \(2011\). The pace of shifting climate in marine and terrestrial ecosystems. *Science*, 334\(6056\), 652–655.](#)
- [Calisesi, Y., V. T. Soebijanta, and R. van Oss \(2005\), Regridding of remote soundings: Formulation and application to ozone profile comparison, *J. Geophys. Res.*, 110, D23306, doi:10.1029/2005JD006122.](#)
- [Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A. \(2000\). GOME-2-Metop's second-generation sensor for operational ozone monitoring. *ESA bulletin*, 102, 28–36.](#)
- [Chipperfield, M. P. and Bekki, S.: Opinion: Stratospheric ozone – depletion, recovery and new challenges, *Atmos. Chem. Phys.*, 24, 2783–2802, <https://doi.org/10.5194/acp-24-2783-2024>.](#)
- Coldewey-Egbers, M., Loyola-R., D. G., Latter, B., Siddans, R., Kerridge, B., Hubert, D., van Roozendael, M., & Eisinger, M. (2025). *The novel GOME-type Ozone Profile Essential Climate Variable (GOP-ECV) data record covering the past 26 years*. *Atmospheric Measurement Techniques*, 18, 5485–5505. doi:10.5194/amt-18-5485-2025
- Delcloo A. (2020): AC SAF Validation report on GOME-2C near real-time and offline global tropospheric ozone
https://acsaf.org/docs/vr/Validation_Report_TrO3-C_Jun_2020.pdf
- Delcloo A. (2024): AC SAF Validation report on GOME-2 reprocessed global tropospheric ozone.
https://acsaf.org/docs/vr/Validation_Report_TrO3_DR_Jan_2024.pdf
- 520 Delcloo A., K., Garane, P., Achtert (2020), AC SAF Validation report on GOME-2 near real-time ozone profiles for GOME-2C.
https://acsaf.org/docs/vr/Validation_Report_NHP-C_OHP-C_Jun_2020.pdf.
- Delcloo A., K., Garane, P., Achtert (2024), AC SAF Validation report on GOME-2 near real-time and reprocessed ozone profiles.
https://acsaf.org/docs/vr/Validation_Report_O3_prof_DR_Jan_2024.pdf

- 525 [Donzelli, G.; Suarez-Varela, M.M. Tropospheric Ozone: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. *Atmosphere* 2024, 15, 779. <https://doi.org/10.3390/atmos15070779>](#)
- [Eskes, H.J., Velthoven, P.F.J.V., Valks, P.J.M. and Kelder, H.M. \(2003\). Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model. *Q.J.R. Meteorol. Soc.*, 129: 1663-1681. <https://doi.org/10.1256/qj.02.14>](#)
- EUMETSAT (2022), GOME-2 Metop-A and -B FDR Product Validation Report Reprocessing R3, (EUM/OPS/DOC/21/1237264), EUMETSAT, 4e, 2022, <https://user.eumetsat.int/>
- 530 [García, O. E., Sanromá, E., Schneider, M., Hase, F., León-Luis, S. F., Blumenstock, T., Sepúlveda, E., Redondas, A., Carreño, V., Torres, C., and Prats, N.: Improved ozone monitoring by ground-based FTIR spectrometry, *Atmos. Meas. Tech.*, 15, 2557–2577, <https://doi.org/10.5194/amt-15-2557-2022>, 2022](#)
- Gaudel, A., Bourgeois, I., Li, M., Chang, K.-L., Ziemke, J., Sauvage, B., Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Smith, N., Hubert, D., Keppens, A., Cuesta, J., Heue, K.-P., Veeffkind, P., Aikin, K., Peischl, J., Thompson, C. R., Ryerson, T. B., Frost, G. J.,
- 535 McDonald, B. C., & Cooper, O. R. (2024). *Tropical tropospheric ozone distribution and trends from in situ and satellite data. Atmospheric Chemistry and Physics*, 24, 9975–10000. doi:10.5194/acp-24-9975-2024
- Godin S., G. Megie, J. Pelon: Systematic lidar measurements of the stratospheric ozone vertical distribution, *GRL*, 16, 6, doi/10.1029/GL016i006p00547, 1998
- Godin-Beekmann, S., Azouz, N., Sofieva, V. F., Hubert, D., Petropavlovskikh, I., Effertz, P., Ancellet, G., Degenstein, D. A., Zawada, D.,
- 540 Froidevaux, L., Frith, S., Wild, J., Davis, S., Steinbrecht, W., Leblanc, T., Querel, R., Tourpali, K., Damadeo, R., Maillard Barras, E., Stübi, R., Vigouroux, C., Arosio, C., Nedoluha, G., Boyd, I., Van Malderen, R., Mahieu, E., Smale, D., and Sussmann, R.: Updated trends of the stratospheric ozone vertical distribution in the 60° S–60° N latitude range based on the LOTUS regression model, *Atmos. Chem. Phys.*, 22, 11657–11673, <https://doi.org/10.5194/acp-22-11657-2022>, 2022.
- ~~García, O. E., Sanromá, E., Schneider, M., Hase, F., León-Luis, S. F., Blumenstock, T., Sepúlveda, E., Redondas, A., Carreño, V., Torres, C., and Prats, N.: Improved ozone monitoring by ground-based FTIR spectrometry, *Atmos. Meas. Tech.*, 15, 2557–2577, <https://doi.org/10.5194/amt-15-2557-2022>, 2022~~
- 545 Hocke, K., Kämpfer, N., Ruffieux, D., Froidevaux, L., Parrish, A., Boyd, I., von Clarmann, T., Steck, T., Timofeyev, Y. M., Polyakov, A. V., and Kyrölä, E.: Comparison and synergy of stratospheric ozone measurements by satellite limb sounders and the ground-based microwave radiometer SOMORA, *Atmos. Chem. Phys.*, 7, 4117–4131, <https://doi.org/10.5194/acp-7-4117-2007>, 2007.
- 550 [IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change\[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou \(eds.\)\]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:10.1017/9781009157896.](#)
- 555 Johnson, M. S., Liu, X., Zoogman, P., Sullivan, J., Newchurch, M. J., Kuang, S., Leblanc, T. and McGee, T., Evaluation of potential sources of a priori ozone profiles for TEMPO tropospheric ozone retrievals, *Atmospheric Measurement Techniques*, 11, 6, 3457–3477, <https://amt.copernicus.org/articles/11/3457/2018/>, 10.5194/amt-11-3457-2018, 2018
- Keckhut, Philippe, et al. "Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change." *Journal of Environmental Monitoring* 6.9 (2004): 721-733, <https://doi.org/10.1029/GL016i006p00547>.

- 560 [Keppens, A., Lambert, J.-C., Granville, J., Miles, G., Siddans, R., van Peet, J. C. A., van der A, R. J., Hubert, D., Verhoelst, T., Delcloo, A., Godin-Beekmann, S., Kivi, R., Stübi, R., and Zehner, C.: Round-robin evaluation of nadir ozone profile retrievals: methodology and application to MetOp-A GOME-2, *Atmos. Meas. Tech.*, **8**, 2093–2120, <https://doi.org/10.5194/amt-8-2093-2015>, 2015.](#)
- [Keppens, A., Di Pede, S., Hubert, D., Lambert, J.-C., Veefkind, P., Sneep, M., De Haan, J., ter Linden, M., Leblanc, T., Compernelle, S., Verhoelst, T., Granville, J., Nath, O., Fjæraa, A. M., Boyd, I., Niemeijer, S., Van Malderen, R., Smit, H. G. J., Dufлот, V., Godin-Beekmann, S., Johnson, B. J., Steinbrecht, W., Tarasick, D. W., Kollonige, D. E., Stauffer, R. M., Thompson, A. M., Dehn, A., and Zehner, C.: 5 years of Sentinel-5P TROPOMI operational ozone profiling and geophysical validation using ozonesonde and lidar ground-based networks, *Atmos. Meas. Tech.*, **17**, 3969–3993, <https://doi.org/10.5194/amt-17-3969-2024>, 2024.](#)
- 565 [Keppens, A., Hubert, D., Granville, J., Nath, O., Lambert, J.-C., Wespes, C., Coheur, P.-F., Clerbaux, C., Boynard, A., Siddans, R., Latter, B., Kerridge, B., Di Pede, S., Veefkind, P., Cuesta, J., Dufour, G., Heue, K.-P., Coldewey-Egbers, M., Loyola, D., Orfanoz-Cheuquelaf, A., Maratt Satheesan, S., Eichmann, K.-U., Rozanov, A., Sofieva, V. F., Ziemke, J. R., Inness, A., Van Malderen, R., & Hoffmann, L. \(2025\). *Harmonisation of sixteen tropospheric ozone satellite data records*. *EGUsphere* \[preprint\]. doi:10.5194/egusphere-2024-3746](#)
- [Leblanc, T., and I. S. McDermid, Stratospheric Ozone Climatology From Lidar Measurements at Table Mountain \(34.4°N, 117.7°W\) and Mauna Loa \(19.5°N, 155.6°W\), *J. Geophysical Research*, **105**, 14,613-14,623, 2000.](#)
- 575 [Liu, X., Bhartia, P. K., Chance, K., Froidevaux, L., Spurr, R. J. D., and Kurosu, T. P.: Validation of Ozone Monitoring Instrument \(OMI\) ozone profiles and stratospheric ozone columns with Microwave Limb Sounder \(MLS\) measurements, *Atmos. Chem. Phys.*, **10**, 2539–2549, <https://doi.org/10.5194/acp-10-2539-2010>, 2010.](#)
- [Lobsiger E., K.F. Künzi and H.U. Dütsch, Comparison of stratospheric ozone profiles retrieved from microwave-radiometer and Dobson-spectrometer data, *J. Atm. and Terr. Phys.*, **46**, 799-806, 1984.](#)
- 580 [McPeters, R. D. and Labow, G. J., Climatology 2011: An MLS and sonde derived ozone climatology for satellite retrieval algorithms, *Journal of Geophysical Research \(Atmospheres\)*, **117**, D10, 2156–2202, doi:10.1029/2011JD017006".](#)
- [Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, *Atmos. Chem. Phys.*, **15**, 8889–8973, <https://doi.org/10.5194/acp-15-8889-2015>, 2015.](#)
- 585 [Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, *Atmos. Meas. Tech.*, **9**, 1279–1301, <https://doi.org/10.5194/amt-9-1279-2016>, 2016.](#)
- [Okamoto, S., Cuesta, J., Dufour, G., Eremenko, M., Miyazaki, K., Boone, C., Tanimoto, H., Peischl, J. and Thompson, C. \(2024\). *Natural and anthropogenic influence on tropospheric ozone variability over the Tropical Atlantic unveiled by satellite and in situ observations*. *EGUsphere* \[preprint\]. doi:10.5194/egusphere-2024-3758](#)
- [Ozone CCI, 2010, <https://climate.esa.int/en/projects/ozone/>, last access: 19 January 2025.](#)
- [Parrish, A., R.L. de Zafra, P.M. Solomon, and J.W. Barrett, A ground-based technique for millimeter wave spectroscopic observations of stratospheric trace constituents, *Radio Sci.*, **23**, 106-118, 1988.](#)
- 595 [Pope, R. J., Kerridge, B. J., Siddans, R., Latter, B. G., Chipperfield, M. P., Feng, W., Pimlott, M. A., Dhomse, S. S., Retscher, C., and Rigby, R.: Investigation of spatial and temporal variability in lower tropospheric ozone from RAL Space UV-Vis satellite products, *Atmos. Chem. Phys.*, **23**, 14933–14947, <https://doi.org/10.5194/acp-23-14933-2023>, 2023.](#)

- Rivoire, L., Linz, M., Neu, J. L., Lin, P., and Santee, M. L.: Satellite nadir-viewing geometry affects the magnitude and detectability of long-term trends in stratospheric ozone, *Atmos. Chem. Phys.*, 25, 2269–2289, <https://doi.org/10.5194/acp-25-2269-2025>, 2025.
- 600 Rodgers C.D., Characterization and Error Analysis of Profiles Retrieved from Remote Sounding Measurements, *J. Geophys. Res.*, 95, 5587–5595, 1990.
- [Rodgers C.D., Inverse methods for atmospheric sounding, 2000, World Scientific Publishing.](#)
- [Rodgers, C. D., and Connor, B. J. \(2003\). Intercomparison of remote sounding instruments, *J. Geophys. Res.*, 108\(D3\), doi:10.1029/2002JD002299.](#)
- 605 Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stübi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., Posny, F. (2007). Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Jülich Ozone Sonde Intercomparison Experiment (JOSIE). *Journal of Geophysical Research*, 112, D19306. doi:10.1029/2006JD007308.
- [Smit, H. G. J., Poyraz, D., Van Malderen, R., Thompson, A. M., Tarasick, D. W., Stauffer, R. M., Johnson, B. J., and Kollonige, D. E.: New insights from the Jülich Ozone Sonde Intercomparison Experiment: calibration functions traceable to one ozone reference instrument, *Atmospheric Measurement Techniques*, 17, 73–112, 2024, <https://amt.copernicus.org/articles/17/73/2024/>, doi: 10.5194/amt-17-73-2024.](#)
- [Sofieva, V. F., Szélag, M., Tamminen, J., Arosio, C., Rozanov, A., Weber, M., Degenstein, D., Bourassa, A., Zawada, D., Kiefer, M., Laeng, A., Walker, K. A., Sheese, P., Hubert, D., van Roozendaal, M., Retscher, C., Damadeo, R., and Lumpe, J. D.: Updated merged SAGE-CCI-OMPS+ dataset for the evaluation of ozone trends in the stratosphere, *Atmos. Meas. Tech.*, 16, 1881–1899, <https://doi.org/10.5194/amt-16-1881-2023>, 2023.](#)
- 615 Spurr, R., de Haan, J., and van Oss, R., and Vasilkov, A., Discrete-ordinate radiative transfer in a stratified medium with first-order rotational Raman scattering, *Journal of Quantitative Spectroscopy & Radiative Transfer*, 2008, 109, 404–425, doi:10.1016/j.jqsrt.2007.08.011.
- Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Tarasick, D. W., Van Malderen, R., Smit, H. G. J., et al. (2022). An examination of the recent stability of ozonesonde global network data. *Earth and Space Science*, 9, e2022EA002459. <https://doi.org/10.1029/2022EA002459>
- 620 Steinbrecht W., et al. (2006), Long-term evolution of upper stratospheric ozone at selected stations of the Network for the Detection of Stratospheric Change (NDSC), *J. Geophys. Res.*, 111, D10308, doi:10.1029/2005J
- Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A., Degenstein, D., Damadeo, R., Zawodny, J., Frith, S., McPeters, R., Bhartia, P., Wild, J., Long, C., Davis, S., Rosenlof, K., Sofieva, V., Walker, K., Rapp, N., Rozanov, A., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Kramarova, N., Godin-Beekmann, S., Leblanc, T., Querel, R., Swart, D., Boyd, I., Hocke, K., Kämpfer, N., Maillard Barras, E., Moreira, L., Nedoluha, G., Vigouroux, C., Blumenstock, T., Schneider, M., García, O., Jones, N., Mahieu, E., Smale, D., Kotkamp, M., Robinson, J., Petropavlovskikh, I., Harris, N., Hassler, B., Hubert, D., and Tummou, F.: An update on ozone profile trends for the period 2000 to 2016, *Atmos. Chem. Phys.*, 17, 10675–10690, <https://doi.org/10.5194/acp-17-10675-2017>, 2017.
- 625 Steinbrecht, W., Velasco, V. A., Dirksen, R., Doppler, L., Oelsner, P., Van Malderen, R., et al. (2025). Ground-based monitoring of stratospheric ozone and temperature over Germany since the 1960s. *Earth and Space Science*, 12, e2024EA003821. <https://doi.org/10.1029/2024EA003821>
- [Stübi, R., Levrat, G., Hoegger, B., Viatte, P., Staehelin, J., Schmidlin, F. J. \(2008\). In-flight comparison of Brewer–Mast and electrochemical concentration cell ozonesondes. *Journal of Geophysical Research*, 113, D13302. doi:10.1029/2007JD009091.](#)
- [Solomon, S., Ivy, D. J., Kinnison, D., Mills, M. J., Neely III, R. R., and Schmidt, A. \(2016\). Emergence of healing in the Antarctic ozone layer. *Science*, 353\(6296\), 269–274.](#)
- 635

- L.G. Tilstra, O.N.E. Tuinder, P. Wang, and P. Stammes, Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY, *J. Geophys. Res. Atmos.* 122, 4084-4111, doi:10.1002/2016JD025940, 2017.
- L.G. Tilstra, O.N.E. Tuinder, P. Wang, and P. Stammes, Directionally dependent Lambertian-equivalent reflectivity (DLER) of the Earth's surface measured by the GOME-2 satellite instruments, *Atmos. Meas. Tech.* 14, 4219-4238, doi:10.5194/amt-14-4219-2021, 2021.
- 640 Tuinder, O. N. E., and van Oss, R. F., de Haan, J., and Delcloo, A., ACSAF Algorithm Theoretical Basis Document for NRT and Offline Vertical Ozone Profile and Tropospheric Ozone Column Products, ATBD, 2.1.3, 2022; <https://acsaf.org/atbds.php>.
- ~~Smit, H. G. J., Poyraz, D., Van Malderen, R., Thompson, A. M., Tarasiek, D.,~~
- ~~Vigouroux, C., Blumenstock, T., Coffey, M., Errera, Q., García, O., Jones, N. B., Hannigan, J. W., Stauffer, R. M., JohnsonHase, F., Liley, B. J., and Kollonige, D. E.: New insights from the Jülich Ozone Sonde Interecomparison Experiment: calibration functions traceable to one ozone reference instrument, *Atmospheric Measurement Techniques*, 17, 73–112, 2024, , doi: , Mahieu, E., Mellqvist, J., Notholt, J., Palm, M., Persson, G., Schneider, M., Servais, C., Smale, D., Thölix, L., and De Mazière, M.: Trends of ozone total columns and vertical distribution from FTIR observations at eight NDACC stations around the globe, *Atmos. Chem. Phys.*, 15, 2915–2933, <https://doi.org/10.5194/amt-17-73-2024>.[acp-15-2915-2015](https://doi.org/10.5194/amt-17-73-2024), 2015.~~
- 645 ~~Stübi, R., Levrat, G., Hoegger, B., Viatte, P., Stachelin, J., Schmidlin, F. J. (2008). In-flight comparison of Brewer–Mast and electrochemical concentration cell ozonesondes. *Journal of Geophysical Research*, 113, D13302. doi:10.1029/2007JD009091.~~
- ~~WMO: by Smit, H. G. J. and the ASOPOS Panel (2014). *Quality assurance and quality control for ozonesonde measurements in GAW*. WMO/GAW Report No. 201. Geneva: World Meteorological Organization. Available at: <https://www.en-sci.com/wp-content/uploads/2018/05/GAW-Report-201-Quality-Assurance-and-Quality-Control-For-Ozonesonde-Measurements-In-GAW.pdf>~~
- 650 WMO: Scientific assessment of ozone depletion: 2018. Global Ozone Research and Monitoring Project Report no. 58, Geneva, Switzerland:
- 655 World Meteorological Organization, ISBN: 978-1-7329317-1-8, 2018.
- WMO: Scientific assessment of ozone depletion: Global Ozone Research and Monitoring Project – GAW Report no. 278, Geneva, Switzerland: World Meteorological Organization, ISBN: 978-9914-733-97-6, 2022.
- ~~Xu, J., Zhang, Z., Rao, L., Wang, Y., Letu, H., Shi, C., Tana, G., Wang, W., Zhu, S., Liu, S. and Shi, J. (2024). Remote sensing of tropospheric ozone from space: Progress and challenges. *Journal of Remote Sensing*, 4, 0178, DOI:10.34133/remotesensing.~~