Ozone_cci+

End to End ECV Uncertainty Budget (E3UB)

Date: 17/03/2025

Version: 6.4

WP Manager:

WP Manager Organization:

Other partners:

EOST: DLR-IMF, RAL, KNMI, IUP, FMI, ULB
VALT: BIRA-IASB, AUTH
CRG: DLR-PA

DOCUMENT PROPERTIES

Title	**End to End ECV Uncertainty Budget**
Reference	Ozone_cci+_D2.2_E3UB_v6.4
Issue	6
Revision	4
Status	Final
Date of issue	17/03/2025
Document type	Deliverable

	FUNCTION	NAME	DATE	SIGNATURE
LEAD AUTHOR	Project partner		29/11/2024	
CONTRIBUTING AUTHORS	Project partner	Viktoria Sofieva Melanie Coldewey-Egbers Klaus Peter Heue Katerina Garane Daan Hubert Arno Keppens Barry Latter Richard Siddans Ronald van der A Catherine Wespes	29/11/2024	
APPROVED BY	Technical Officer	Michael Eisinger	05/03/2025	
ISSUED BY	Science Leader	Daan Hubert	17/03/2025	

DOCUMENT CHANGE RECORD

Issue	Revision	Date	Modified items	Observations
0	0	20/02/2020	Initial template	Creation of document
1	1	08/05/2020	First complete version	
2	1	09/11/2020	Implementation of comments	
3	0	20/01/2021	Minor updates	
4	0	24/09/2021	Minor updates	
5	0	20/10/2021	Major updates	
6	0	27/12/2023	Major updates: only datasets in Phase 2 are considered	
6	1	5/5/2024	Minor updates	
6	2	7/10/2024	Minor updates in Sect. 2.1.2 and 3.3.2	
6	3	10/12/2024	Addressing the comments of Daan Hubert; updates in Sections 2 and 3.	
6	4	17/03/2025	Document header; updated reference documents.	Approved by ESA

Table of Contents

1 Purpose and scope

1.1 Purpose

An End-to-End ECV Uncertainty Budget (E3UB) describes all steps of uncertainty assessment from comprehensive uncertainty estimates of individual measurements to the full error budget of higher level data products. Error budget studies in the O3 CCI+ project (https://climate.esa.int/en/projects/ozone) are based on both error propagation and geophysical validation of ozone measurements and their uncertainties. Instrumental drift issues are investigated as well. The purpose of this document is to collect in one place the characterization and geophysical validation of uncertainty estimates of all individual Level 2 datasets participating in the project and provide characterizations of errors of the Level 3 and Level 4 Climate Data Records (CDR) generated within the project.

In this document the focus is on algorithm activities during Phase II of the O3 CCI+ project (2022-2024). For datasets developed in O3 CCI Phase II (2014-2017) and O3 CCI+ Phase I (2019-2022), we refer to the Comprehensive Error Characterization Report (CERC) document [RD3] and corresponding End to End ECV Uncertainty Budget (E3UB) [RD4].

1.2 Reference documents

[RD1] CCI Data Standards v2.3, online at
https://climate.esa.int/media/documents/CCI_DataStandards_v2-3.pdf

[RD2] CCI+ Phase 2 Product Validation and Intercomparison Report (PVIR) v5.1, online at
https://climate.esa.int/en/projects/ozone/key-documents

[RD3] CCI Comprehensive Error Characterization Report (CERC) v2, online at
https://climate.esa.int/en/projects/ozone/key-documents

[RD4] CCI+ Phase 1 End to End ECV Uncertainty Budget (E3UB) v5.0, online at
https://climate.esa.int/en/projects/ozone/key-documents

[RD5] CCI+ Phase 2 Algorithm Theoretical Basis Document (ATBD) v3.2, online at
https://climate.esa.int/en/projects/ozone/key-documents

1.3 Summary and terminology

The **"precision"** of an instrument/retrieval is its random (in the time domain) error. It is the debiased root mean square deviation of the measured values from the true values. The precision can also be seen as scatter of multiple measurements of the same quantity. The

difference between the measured and the true state can still be large, because there still can be a large systematic error component unaccounted by the precision.

The **"bias"** of an instrument/retrieval characterizes its systematic (in the time domain) error. It is the mean difference of the measured values from the true values.

The **"total error"** of an instrument/retrieval characterizes the estimated total difference between the measured and the true value. In parts of the literature the expected total error is called "accuracy" but we suggest not using this particular term because its use in the literature is ambiguous.

1.1. Acronyms

ACE-FTS	Atmospheric Chemistry Experiment – Fourier Transform Spectrometer
CCI	Climate Change Initiative
CDR	Climate Data Record
ECMWF	European Centre for Medium-range Weather Forecast
ECV	Essential Climate Variable
ENVISAT	Environmental Satellite (ESA)
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FMI	Finnish Meteorological Institute
FORLI	Fast Optimal Retrievals on Layers for IASI
GODFIT	GOME-type Direct-FITting
GOME	Global Ozone Monitoring Experiment
GOMOS	Global Ozone Monitoring by Occultation of Stars
IASI	Infrared Atmospheric Sounding Interferometer
ISS	International Space Station
KNMI	Royal Netherlands Meteorological Institute
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detection of Atmospheric Composition Change
OMI	Ozone Monitoring Instrument (aboard EOS-Aura)
OMPS-LP	Ozone Mapper and Profile Suite - Limb Profiler (aboard Suomi-NPP)
OSIRIS	Optical and Spectroscopic Remote Imaging System (aboard Odin)
POAM	Polar Ozone and Aerosol Measurement (aboard SPOT 4)
RAL	Rutherford Appleton Laboratory
SABER	Sounding of the Atmosphere using Broadband Emission Radiometry

SAGE	Stratospheric Aerosol and Gas Experiment
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric CHartography
UTLS	Upper Troposphere Lower Stratosphere

2 Uncertainty of Level 2 data

The Level 2 data (individual ozone profiles or column data) are the starting point for creating climate data variables. For remote sensing measurements, the uncertainty budget is estimated via propagation of measurement noise (random) and other uncertainties (random or systematic) through the inversion algorithm (e.g., Rodgers, 2000). Von Clarmann et al. (2020) uses the term **"ex-ante"** for the uncertainty estimates by an inversion algorithm, so do we in this document. Some CCI projects prefer the term "prognostic" uncertainty.

Ex-ante uncertainty estimates might be incomplete: this might be due to incomplete/simplified models of the processes that describe the satellite measurements or/and unknown/unresolved atmospheric features. Other contributing factors might be the imperfect estimates of measurement uncertainties, as well as the uncertainties of external auxiliary data. Therefore, validation of theoretical (ex-ante) uncertainty estimates is desired for remote-sensing measurements. For atmospheric measurements specifically, a distinction shall be made between a baseline validation, which consists in checking biases and comparing dispersions with the requirements, and a proper validation of ex-ante uncertainty estimates, which is not a trivial task because the measurements are performed in a continuously changing atmosphere. The experimental estimates of uncertainty estimates are called **"ex-post"** estimates in von Clarmann et al. (2020), and we follow this terminology. Other CCI projects refer to such estimates as "diagnostic" uncertainty.

This section presents the characterization of Level 2 uncertainties (ex-ante) and the results of the uncertainty validation. The overview of the methods for uncertainty validation is collected in Section 4 of E3UB Ozone_cci Phase 1 [RD 4], together with a summary of publications on error budget evaluation and uncertainties validation of Level 2 ECV's generated within Ozone_cci.

2.1 Ozone profiles from nadir sensors

In the O3 CCI+ project, ozone profiles are processed with RAL and FORLI algorithms. The RAL algorithm is applied to GOME-type sensors (e.g., OMI and TROPOMI), while the FORLI algorithm is applied to IASI sensors.

2.1.1 RAL processor

This section is unchanged from the CCI Phase 1 document [RD4] and remains relevant for the current publicly released products. In CCI phase 2 a re-engineered and updated version of

the profile scheme has been developed, but products are not yet released. Uncertainties on that product are described in the ATBD covering the re-engineered scheme [RD5].

Analysis of error budget of RAL scheme, reported in (Siddans, 2003), is based on performing retrieval simulations for a set of basic geophysical scenarios, which had been defined for the GOME-2 Error Study (Kerridge et al., 2002). Figure 1 shows retrieval precision and base-line mapped errors for GOME-1 and the April 55°N scenario from Siddans (2003). Dashed and solid lines refer to the 80% and 5% surface albedo cases respectively. Colours distinguish results for the 3 across-track ground pixels in Band 1 (the legend shows the pixel mean off-nadir angle in degrees; positive angle are East of nadir). Dotted lines in each panel other than the top left show (for comparison) the precision where the scale permits. The black dash-dot curve is the a priori error input to the B1 retrieval. Retrieval precision and a priori are also plotted as negative values for comparison with negative mapped errors.

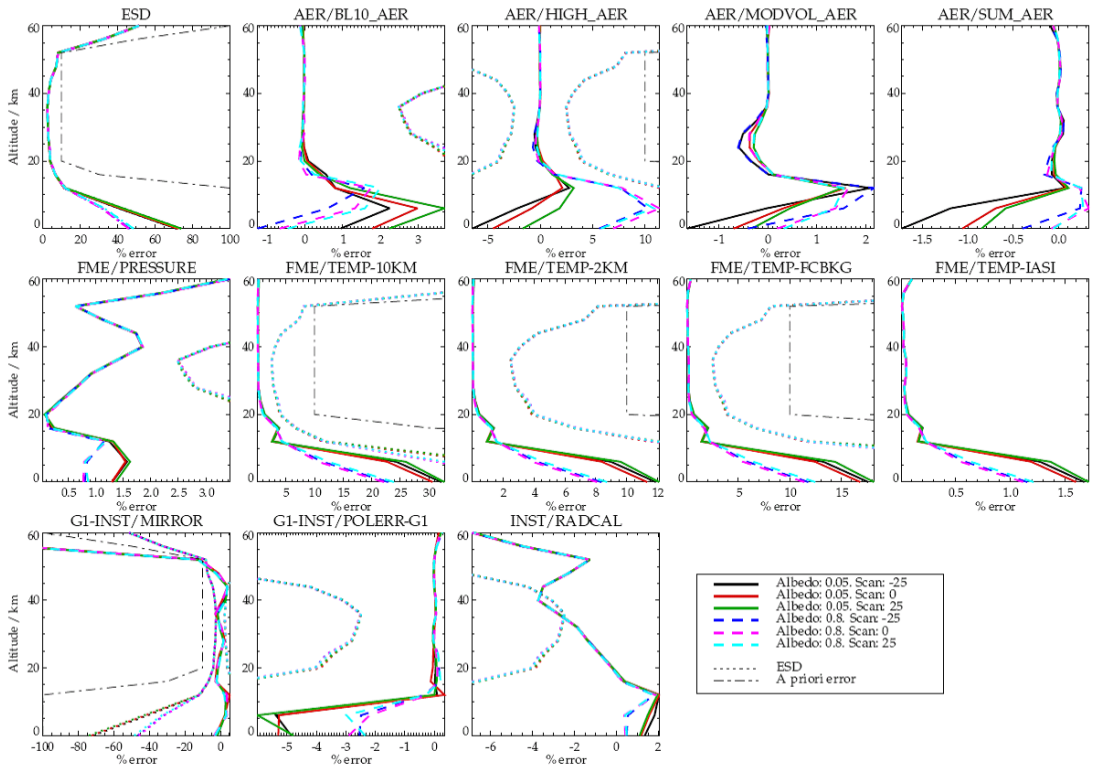

Figure 1: Retrieval precision and base-line mapped errors for GOME-1 and the April 55°N scenario.

(Miles et al., 2015) assessed the performance of the RAL ozone profile retrieval scheme for GOME-2 with a focus on tropospheric ozone. The retrieval precision, as given by the square roots of diagonals of the solution error covariance matrix is generally in the few percent range

in the stratosphere increasing to a few tens of percent in the lowest retrieval levels. An example is presented in Figure 2 for a mid-latitude profile in Northern Hemisphere summer. In this case, the retrieval precision on retrieval levels is typically much smaller than a priori error throughout the profile. The retrieval noise error is around a factor of 2 smaller than the retrieval precision.

Figure 3 shows an example of how the retrieval precision varies for a typical orbit cross-section; the uncertainty values are higher at lower altitudes in tropical and polar conditions.

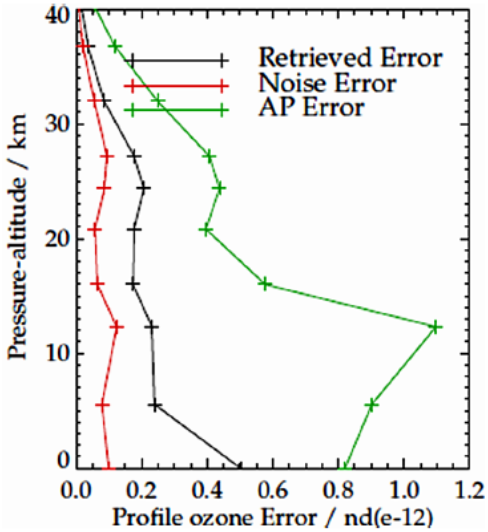

Figure 2: Error assessment for ozone profile for a GOME-2 nadir pixel at 45°N on 25 August 2008.

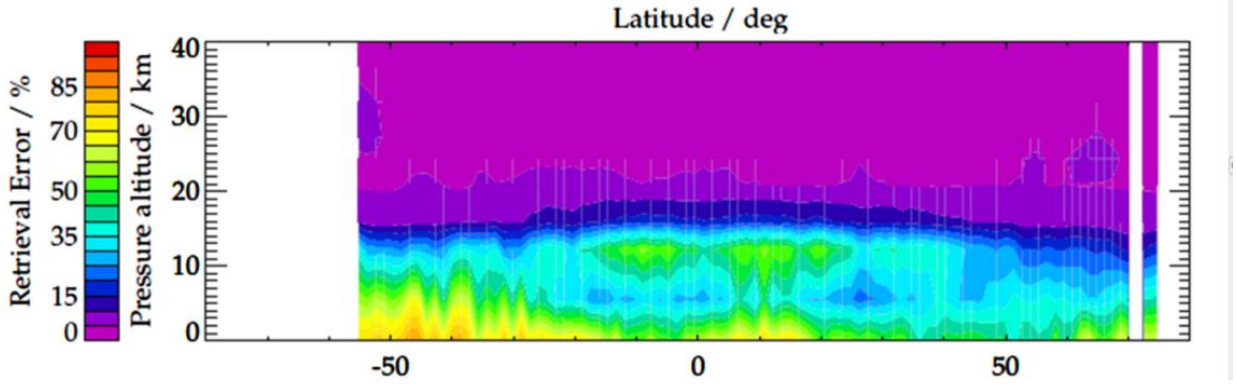

Figure 3: Relative retrieval error of ozone product from GOME-2 retrieved with RAL scheme.

Keppens et al. (2018) reported a relative random error for RAL v2.14 of about:

- 5 % at the altitude of the ozone maximum;
- up to about 10 % at higher altitudes;
- up to 40 % in the lower troposphere.

By comparing nadir ozone profiles and ground-based ozonesonde and lidar measurements, the authors conclude that "the total satellite measurement and retrieval uncertainty is typically underestimated in the RAL v2.14 nadir ozone profile products, because the ex-ante uncertainty under consideration only includes random noise errors."

Regarding satellite drifts, Keppens et al. (2018) reported:

- negative and insignificant decadal drift on the order of 5 % per decade for GOME;
- insignificant (except for the tropics) drift on the order of −15 and 10 % per decade for OMI's L2 stratospheric and tropospheric observations, respectively;
- a significant positive drift of ~40 % per decade for SCIAMACHY and GOME-2A below the tropopause.
- a significant 30 % per decade negative drift in the UTLS at all latitudes for GOME-2A.

2.1.2 IASI FORLI processor

The estimated statistical uncertainties on the ozone vertical profiles retrieved from FORLI (v20191122) are calculated for three latitude bands and different altitudes for the year 2014 in Figure 4 (adapted from Wespes et al. (2016) using FORLI v20151001). The retrieval total errors depend on the latitude and the season, reflecting, amongst other things, the influence of signal intensity, of interfering water lines and of thermal contrast under certain conditions (e.g. temperature inversion, high thermal contrast at the surface). It usually ranges between 10 and 30 % in the troposphere and in the UTLS, except in the equatorial belt due to the low ozone amounts which leads to larger relative errors (Wespes et al., 2016). The retrieval errors are usually less than 10 % in the stratosphere. The error is larger above cold surfaces, possibly due to a misrepresentation of the emissivity in the polar regions (Hurtmans et al., 2012;.

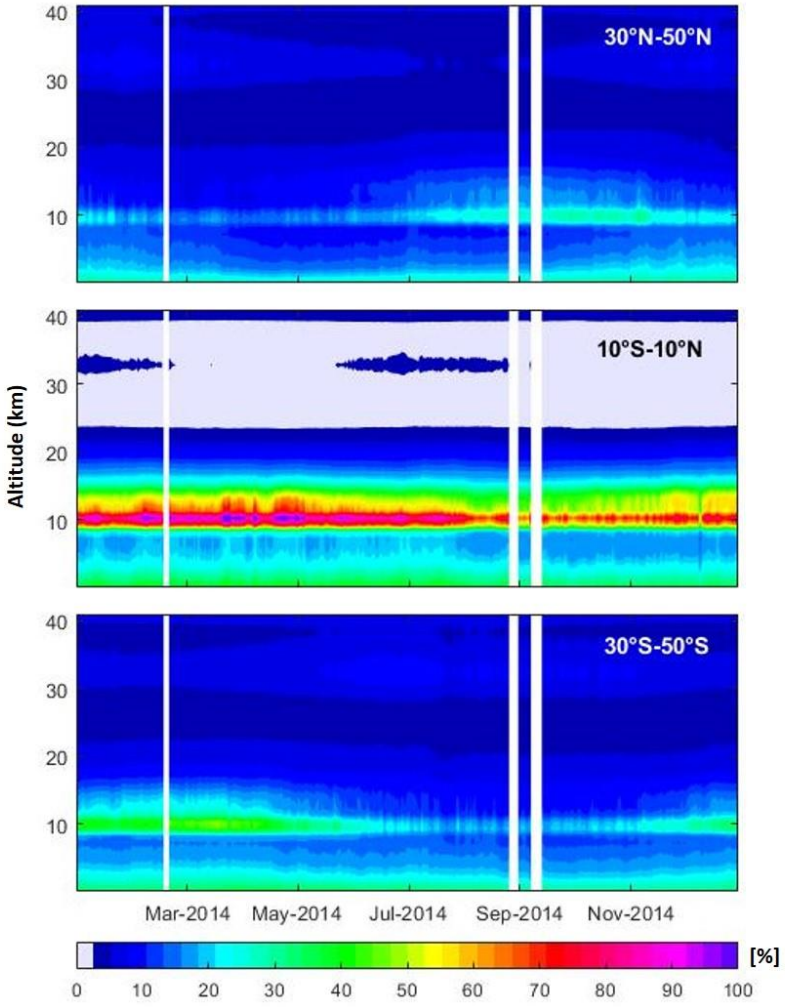

Figure 4: Daily estimated FORLI-O3 (v20191122) total retrieval errors (%) for the year 2014 as a function of time and altitude, for three latitude bands: 30-50°N, 10°s-10°N and 30-50°S.

As shown in Figure 5, the main contributions to the total error are the limited vertical sensitivity (smoothing error), the measurement noise and the uncertainty of the fitted water vapour column. Another contribution comes from the uncertainty on fixed parameters, such as surface emissivity and temperature profile, but in the routine processing of the error matrix, these are not taken into account.

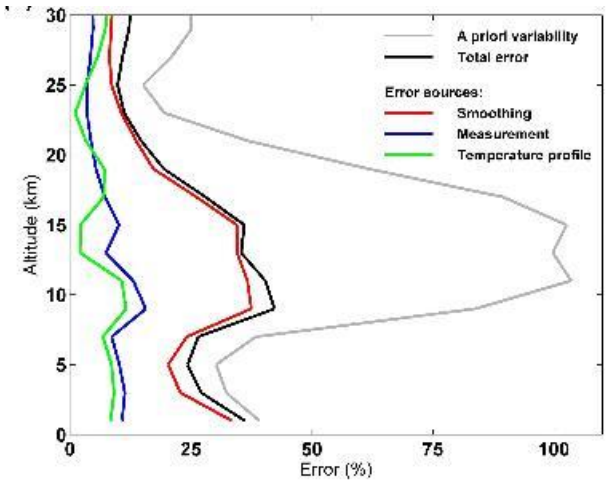

Figure 5: Error budget analysis for IASI ozone retrieval (Boynard et al., 2009). The a priori variability and total errors are given by the square root of the diagonal elements of the a priori covariance matrix and the error covariance matrix, respectively.

For ozone, the error is larger in the tropics (above 30%) due to the increase in humidity and also above cold surfaces, possibly due to a misrepresentation of the emissivity in the polar regions (Hurtmans et al., 2012, Wespes et al., 2016).

The IASI FORLI-O_3 profile dataset (v20151001) has been extensively validated by Boynard et al. (2018) and Keppens et al. (2018). They demonstrated a good degree of accuracy, precision and vertical sensitivity with no instrumental drift. The retrieval data products showed an insignificant stratospheric bias smaller than 10 %, a 10 to 40 % positive bias in the UTLS, and a ~4 to 20 % bias in the troposphere, depending on latitudes. Keppens et al. (2018) also report that the ex-ante IASI uncertainties provided in Level 2 files are typically of the order of the bias above the UTLS. The ex-post random uncertainty, as estimated by the spread, is about twice as large, except for the lower tropics. This means that, overall, the total satellite measurement and retrieval uncertainty is underestimated in the IASI FORLI nadir ozone profile products.

Typical uncertainty values and retrieval characteristics (for both FORLI-v20151001 and -v201991122) are reported in Table 1.

Altitude range	0-40 km: Retrieval performed on a uniform 1 km vertical grid on 40 layers from surface up to 40 km with an extra layer from 40 km to 60 km.
Vertical resolution	~7 km troposphere, ~15 km stratosphere
Random errors: Measurement error & Smoothing error	< 10% in the total O3 columns ; 10-30% in the troposphere and in the UTLS 5-20% in the stratosphere
Systematic errors: - Uncertainty in cross-sections - Temperature uncertainty	 ~4% <10% over all the profile

Table 1: IASI ozone profiles characteristics and error budget

El Aabaribaoune et al. (2021) estimated inter-channel error covariances and evaluated their impact on ozone assimilation with a chemistry transport model.

There is no significant reported bias due to instrument aging. The excellent inter-platform consistency is in line with the good stability of the IASI instruments. When comparing IASI/MetOpA vs IASI/MetOpB vs IASI/MetOpC, the radiance signals are similar, with differences below the IASI radiometric noise (Chinaud et al. 2019). The ozone profile data retrieved by FORLI-O3 from IASI radiances are highly consistent as well, with differences generally below 1% (within 0.5% as shown by Figure 6 for the year 2014). However, discontinuity was found in the IASI FORLI (v20151001 and v20191122) datasets. It was suspected to result from coincident updates in the Eumetsat Level-2 algorithm that produces humidity and temperature profiles used as inputs in FORLI. In order to solve this issue, the IASI FORLI (v20151001) dataset has been reprocessed using a climate data record of Level-2 inputs (EUMETSAT L2 CDR) to produce the IASI-O3 CDR product [RD5].

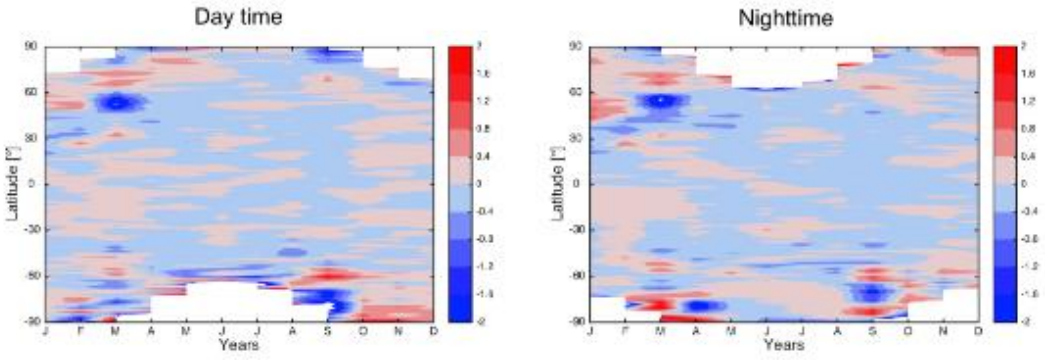

Figure 6: Contour representation of the relative difference (in percent) between IASI-A and IASI-B total ozone column retrieved using FORLI as a function of latitude and time for the year 2014 for day time data (left) and night-time data (right). The relative differences are calculated as 100 x (IASI-A – IASI-B) / IASI-A (Boynard et al., 2018).

2.2 Ozone profiles from limb sensors

2.2.1 OMPS-LP on board NOAA-21

The preliminary version of ozone profiles retrieved from the OMPS-LP instrument on board NOAA-21 (N21) is obtained using the same retrieval approach based on spectral fitting as for observations from OMPS-LP on board Suomi-NPP (SNPP). Retrieval settings are also similar, as described in (Arosio et al., 2018).

A study was performed to estimate the error budget for the retrieved profiles from SNPP (Arosio et al., 2022). Results from that study are here briefly summarized and are expected to be valid for the ozone retrievals from N21 observations.

Synthetic simulations with SCIATRAN were used to estimate the error budget. Main source of parameter uncertainties is the pointing accuracy of the instrument, especially in the upper stratosphere. The precision of the retrieved aerosol extinction profiles and surface albedo plays a relevant role below 25 km with parameter uncertainty up to 5 %. The cloud filtering is also relevant below 20 km. This contribution is rather systematic and was added to the retrieval bias. Another relevant source of uncertainty was found to be related to the used ozone cross-section. The cross-section error propagated into Level 2 profiles is estimated to be on the order of 2 % in the lower stratosphere and decreasing with altitude. In addition, the analysis of the error from model approximations revealed that the radiative transfer solver and polarisation effect may add to the total ozone error a contribution up to 1-2 %. Errors were classified in random and systematic components, providing a total estimation of the two as a function of latitude.

In addition to this assessment, a further term shall be added for N21 retrievals, related to the usage of the aerosol profile. Since there was no time to develop an adequate aerosol retrieval product for N21 observations and, at the same time, OMPS-LP SNPP is still operational, we decided to use the aerosol information retrieved from SNPP in the N21 ozone retrieval, after averaging it zonally and daily. A sensitivity study was performed on SNPP observations to estimate the expected difference between using the corresponding aerosol profiles for each observation or a zonally and daily averaged extinction profile. Results have shown that the impact of this approximation affects mostly altitudes below 18 km, with an estimated random uncertainty of about 2 % at 17 km, increasing to 5 % at 12 km.

3 Uncertainty of level 3 data

3.1 Tropospheric ozone column

3.1.1 Convective Cloud Differential (CCD) Method

The convective cloud differential (CCD) method to retrieve tropospheric ozone columns is applied to level 2 GOME-type data, i.e., ozone total columns and cloud data. This study was performed in O3 CCI Phase 2 (CERC, RD3). The algorithm is applied to level 2 data of the following instruments: GOME, SCIAMACHY, OMI, GOME-2 (A, B, C). CCD tropospheric ozone columns are operationally retrieved by the Sentinel 5p (S5P) processors. However, the S5P data are gridded with a higher spatial and temporal resolution (Heue et al., 2021) and are averaged to 1°x 1° x 1 month.

The data are gridded on 1° x 1° grid with a monthly resolution, therefore within one grid cell several individual observations are averaged. Depending on the instrument the maximum number of averaged observations per grid cell varies between 20 for GOME and roughly 4000 for S5P. For SCIAMACHY, GOME-2 and OMI the numbers vary around 50-100. Because of that the influence off the individual measurement and the propagated error is limited and the variation between the individual observations dominates the uncertainty. In the following we will estimate the influence of both effects. The propagation of the individual measurement error to final product is called error and the standard deviation within a grid cell is called uncertainty. This naming scheme is also used in the data files.

We estimated the error caused by the total column error σ_{toc} to the stratospheric columns and propagating to the tropospheric columns. The effect of the cloud data (cloud fraction and cloud height mainly) on the error of the retrieved tropospheric column was assumed to be negligible. The squared error of the stratospheric column error σ_{soc} is given as the mean of the squared errors within the reference region (70° E to 170° W)

$$\sigma_{soc}^2 = \frac{\sum_{reference} \sigma_{toc}^2}{n_{ref}} \quad (3.1.1.1)$$

This error propagated into the error of each tropospheric columns (troc) as

$$\sigma_{(i,troc)}^2 = \sigma_{(i,toc)}^2 + \sigma_{(i,soc)}^2, \quad (3.1.1.2)$$

where *i* indicates the cloud free total column observation or the respective latitude band for the stratospheric column. The final tropospheric error is calculated as square root of the mean tropospheric squared errors

$$\sigma_{troc} = \sqrt{\frac{\sum_i^n \sigma_{(i,troc)}^2}{n}} \quad (3.1.1.3)$$

As long as enough measurements are averaged the standard deviation of the measurements per grid cell is at least 5 times higher as the propagated error. The error derived by the error propagation was about 0.3 mmol/m² (0.6 DU) for GOME in August 1996. For the same period the standard deviation of tropospheric columns within a grid cell was about 2-3 mmol/m² (~4.5 DU). Similar results were found for the other sensors. One can easily calculate that in the above-mentioned example the overall uncertainty hardly changes when the squared error and standard deviations are added.

$$\sigma_{(i,troc)} = \sqrt{\left(\sigma_{(std,troc)}^2 + \sigma_{(err,troc)}^2\right)} = \sqrt{(3^2 + 0.3^2)} \approx 3.015 \approx 3 \quad (3.1.1.4)$$

Both the standard deviation and the propagated error are given for each grid cell and for each instrument in the respective files. However, as the standard deviation clearly dominates the uncertainty, the merged final results contain the deviation only.

In case only very few data were averaged the concept of the standard deviation as error is no longer valid. In future releases we therefore plan to provide a rough estimate of the propagated errors as well.

For S5P, here the quality flag is reduced for cases with low numbers of observations in the operational dataset. Because S5P data are averaged from 3 days to monthly data the uncertainty cannot be directly retrieved from all observations within a month. The 3 days uncertainty σ_{daily} is hence propagated to the monthly $\sigma_{monthly}$:

$$\sigma_{month}^2 = \frac{\sum_{month} \sigma_{daily}^2}{month} \quad (3.1.1.5)$$

3.1.2 OMI-LIMB and GTO-LIMB tropospheric ozone column

The random uncertainty of the resulting tropospheric ozone column is estimated via error propagation through all steps of the retrieval algorithm.

For OMI daily total ozone column, the uncertainty is computed as:

$$\sigma^2 = \frac{1}{N}\sum_i \sigma_i^2 + \frac{1}{N} var(\rho_i) \qquad (3.1.2.1)$$

where σ_i are uncertainties reported by the retrieval algorithm and $\mathrm{var}(\rho_i)$ is the variance of N individual ozone values in the bin. The typical daily gridded clear-sky total ozone column and the corresponding random uncertainties are in the range of 0.5 - 5 DU (typically 1-2 DU).

The uncertainty estimation of daily gridded and interpolated ozone profiles is described in Sect. 3.3.3. Uncertainties of the stratospheric ozone column are then estimated using error propagation, it is mostly 5-8 DU (< 2%).

Uncertainties of daily tropospheric ozone values are estimated as:

$$\sigma_{TrOC}^2 = \sigma_{TOC}^2 + \sigma_{SOC}^2 , \qquad (3.1.2.2)$$

where σ_{TOC}^2 is the uncertainty of the total ozone column and σ_{SOC}^2 is the uncertainty of the stratospheric ozone column.

The uncertainties of monthly average data are estimated similarly to uncertainties of the gridded data, i.e., by Eq. (3.1.2.1). In this case, all parameters are related to tropospheric ozone column. In the majority of tropical locations, the estimated uncertainty of the tropospheric ozone column is 5-7 DU. Over Indonesia, where the tropospheric ozone column has the smallest values, the relative uncertainty increases to 100%. In the mid-latitudes, the estimated uncertainties are mostly within the range of 15-40%. The largest uncertainty is close to the polar vortex boundary, as expected. The illustrations of the uncertainties are shown in Figure 7 and in Sofieva et al. (2022).

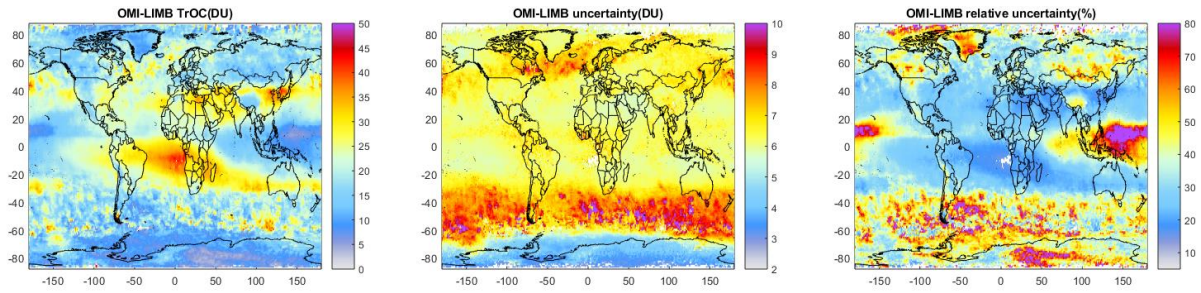

Figure 7. Adapted from Sofieva et al. (2022): OMI-LIMB tropospheric ozone distributions (DU, color) for September 2018 (left). The stratospheric ozone column is estimated from 3 km below the tropopause. The corresponding estimated uncertainties are shown in absolute values (DU) on the central panel and in relative values (%) on the right panel.

3.2 Merged data sets

3.2.1 GOP-ECV

The merged ozone profile data record GOP-ECV (GOME-type Ozone Profile Essential Climate Variable) from the nadir ultraviolet-visible satellite sensors consists of monthly mean ozone partial columns for 19 atmospheric layers covering 0-80 km and 5°x5° latitude-longitude grid cells. Data from five sensors (GOME, SCIAMACHY, OMI, GOME-2A, and GOME-2B) are merged into a homogeneous record spanning the period 1995-2021. The merged profiles are finally adjusted with respect to the GTO-ECV total column data record (Coldewey-Egbers et al., 2015, 2022).

The uncertainties are approximated analogously as described in detail in Sofieva et al. (2021). At first, monthly mean profiles $\rho_{k,i}$ from each individual sensor k are computed for each spatio-temporal bin i, and the uncertainties of the averages are evaluated by the standard error $\sigma_{k,i}$ of the mean

$$\sigma_{k,i} = \frac{s_{k,i}}{\sqrt{n}} \qquad (3.2.1.1)$$

$s_{k,i}$ is the standard deviation and n is the number of measurements in that spatio-temporal bin. Since the merging is based on de-seasonalized anomalies, the climatological means $\rho_{k,m}$ and their corresponding uncertainties $\sigma_{k,m}$ for each sensor k and each month m from January to December are computed as well. The uncertainty of the seasonal cycle is estimated via

$$\sigma_{k,m}^2 = \frac{1}{N_m}\sum_{j=1}^{N_m} {\sigma_{k,j}}^2 \qquad (3.2.1.2)$$

where N_m is the number of monthly mean values for month m for all available years. The uncertainty $\sigma_{k,i}$ of the monthly mean profile and the uncertainty $\sigma_{k,m}$ of the climatological mean are used to obtain an estimate of the uncertainty of the de-seasonalized anomaly $\sigma_{\Delta_{k,i}}$:

$$\sigma_{\Delta_{k,i}} = \Delta_{k,i}\sqrt{\frac{\sigma_{k,i}^2}{\rho_{k,i}^2} + \frac{\sigma_{k,m}^2}{\rho_{k,m}^2}} \quad (3.2.1.3)$$

$\Delta_{k,i}$ is the de-seasonalized relative anomaly for sensor k and spatio-temporal bin i and computed as

$$\Delta_{k,i} = \frac{\rho_{k,i} - \rho_{k,m}}{\rho_{k,m}} \quad (3.2.1.4)$$

The merged de-seasonalized anomaly $\Delta_{\text{merged},i}$ then corresponds to the median of the individual sensors

$$\Delta_{\text{merged},i} = \text{median}\left(\Delta_{k,i}\right) \quad (3.2.1.5)$$

Following Sofieva et al. (2021), the uncertainty $\sigma_{\Delta_{\text{merged},i}}$ of the merged de-seasonalized anomaly $\Delta_{\text{merged,i}}$ is finally estimated as

$$\sigma_{\Delta_{\text{merged},i}} = \min\left(\sigma_{\Delta_{k_{\text{median}}}}, \sqrt{\frac{1}{N}\sum_{k=1}^{N}\sigma_{\Delta_{k,i}}^2 + \frac{1}{N^2}\sum_{k=1}^{N}\left(\Delta_{k,i} - \Delta_{\text{merged},i}\right)^2}\right) \quad (3.2.1.6)$$

In that equation the first term $\sigma_{\Delta_{k_{\text{median}}}}$ is the uncertainty of the anomaly of the sensor corresponding to the median value. For more details we refer to Sofieva et al. (2021).

3.2.2 Merged IASI dataset

The results from previous phases of the CCI project showed an excellent consistency between the IASI-A, -B and -C ozone profile retrievals by the standard FORLI-O3 algorithm (v20191122). The merged IASI-A, -B and -C dataset of ozone profiles consists of daily gridded clear-sky weighted mean ozone profiles in the pressure range from the surface to 40 km with one extra layer from 40 to 60 km, and with a horizontal resolution of 1°x 1°.

The weighted average ozone value and its associated error estimate in a specific level and in a level-3 grid cell are calculated (the details are provided in Sect. 3.1 of [RD4]). The weighted

mean is computed over all IASI-A, -B and -C values assigned to that bin with the weight of each measurement is the reciprocal square of the total retrieval random errors at that level (i.e. measurement and smoothing errors) .

3.2.3 High-resolution gap-free dataset of limb ozone profiles

The high-resolution gap-free dataset of ozone profiles consists of daily gridded profiles in the pressure range from 900 to 0.02 hPa with a horizontal resolution of 1°x 1°.

Ozone profiles from satellite limb instruments are debiased to Aura MLS and their random uncertainties are a-posteriori estimated by the method described in Sofieva et al. (2022). These uncertainties are used in the kriging-type interpolation.

The estimation of uncertainties associated with the interpolated dataset of ozone profiles are evaluated as follows. Firstly, we used the error propagation to evaluate the uncertainty after the kriging step. In addition, we estimated the interpolation uncertainty using the SILAM data: we run the same interpolation but on the SILAM fields sub-sampled at the measurements locations, and evaluated the error as the absolute difference of true and interpolated data. The final uncertainty is the root-mean-square of error propagation and model-assessed interpolation errors. The illustration of error estimation can be found in the Supplement of Sofieva et al. (2022). It is presented also in Figure 8.

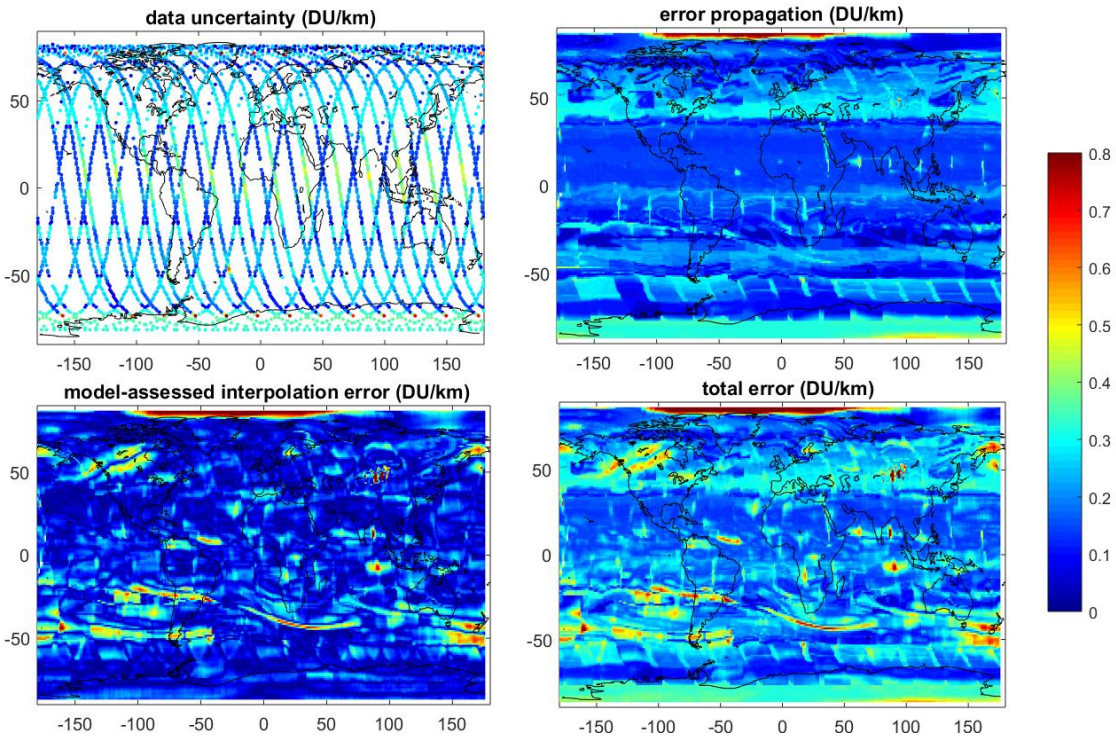

Figure 8. From Sofieva et al. (2022): Illustration of uncertainty estimation of the interpolated ozone data. Data are for 1 Sep 2018, at 10 hPa. Top left: uncertainty of satellite data, top right: error propagation after kriging-type interpolation, bottom left: interpolation error from SILAM data, bottom right: total uncertainty.

4 Uncertainty of level 4 data

4.1 Multi-sensor reanalysis total ozone

Error analysis

We have compared the observed OmF (by comparing the forecast with individual observations) with the estimate of the OmF. The latter is calculated from the combination of the model forecast error as computed in TMDAM and the given individual measurement error bars on the observations. This approach can be seen as an extension of the much-used $\chi 2$ test, which essentially checks if the mean of both quantities are consistent. The OmF is much smaller for the more modern satellite instruments in recent years, mainly because of the higher number of observations per model grid cell and the daily revisit cycle.

To evaluate the performance of the MSR2 we routinely produce and store OmF and observation-minus-analysis (OmA) statistics. These statistics provide an internal consistency

check on the error estimates for the total ozone retrieval as well as for the model performance.

Results

To evaluate the quality of the MSR2 data, the OmF and OmA statistics have been analysed. The OmA of this data set is less than 1 %, which is better than for the assimilation of observations of a single sensor. The model bias as estimated by the difference between OmF and OmA is in general small: for periods of a couple of days with no data, the bias remains within 1 %. This holds also for the period with only sparse BUV observations, although model biases of several percent as a function of latitude can become visible. The RMS errors are around 2–3 % after 1979, which is small given that the RMS errors contain contributions from the representativity errors, forecast errors and instrumental noise.

The fitted offset, trend and seasonality in the comparison between the MSR2 level 4 ozone fields and the average of the ground-based observations were negligible. The maximum fitted offset is 0.2 DU. All systematic effects found in the satellite data are removed by the simple corrections (using a few basic parameters) applied to the satellite observations.

More detailed information can be found in Van Der A et al. (2015).

5 References

Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., Clarmann, T. von, and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb scattering observations, Atmospheric Measurement Techniques, 11, 2135–2149, https://doi.org/10.5194/amt-11-2135-2018, 2018.

Arosio, C., Rozanov, A., Gorshelev, V., Laeng, A., and Burrows, J. P.: Assessment of the error budget for stratospheric ozone profiles retrieved from OMPS limb-scatter measurements, Atmospheric Measurement Techniques Discussions, 2022, 1–27, https://doi.org/10.5194/amt-2022-116, 2022.

Boynard, A., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Turquety, S., George, M., Hadji-Lazaro, J., Keim, C., and Meyer-Arnek, J.: Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations, Atmos. Chem. Phys., 9, 6255–6271, https://doi.org/10.5194/acp-9-6255-2009, 2009.

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.

Clarmann, T. von, Degenstein, D. A., Livesey, N. J., Bender, S., Braverman, A., Butz, A., Compernolle, S., Damadeo, R., Dueck, S., Eriksson, P., Funke, B., Johnson, M. C., Kasai, Y., Keppens, A., Kleinert, A., Kramarova, N. A., Laeng, A., Langerock, B., Payne, V. H., Rozanov, A., Sato, T. O., Schneider, M., Sheese, P., Sofieva, V., Stiller, G. P., Savigny, C. von, and Zawada, D.: Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature, Atmospheric Measurement Techniques, 13, 4393–4436, https://doi.org/10.5194/amt-13-4393-2020, 2020.

Coldewey-Egbers, M., Loyola, D. G., Koukouli, M., Balis, D., Lambert, J.-C., Verhoelst, T., Granville, J., Roozendael, M. van, Lerot, C., Spurr, R., Frith, S. M., and Zehner, C.: The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative, Atmospheric Measurement Techniques, 8, 3923–3940, https://doi.org/10.5194/amt-8-3923-2015, 2015.

Coldewey-Egbers, M., Loyola, D. G., Lerot, C., and Van Roozendael, M.: Global, regional and seasonal analysis of total ozone trends derived from the 1995–2020 GTO-ECV climate data record, Atmos. Chem. Phys., 22, 6861–6878, https://doi.org/10.5194/acp-22-6861-2022, 2022.

El Aabaribaoune, M., Emili, E., and Guidard, V.: Estimation of the error covariance matrix for IASI radiances and its impact on the assimilation of ozone in a chemistry transport model, Atmos. Meas. Tech., 14, 2841–2856, https://doi.org/10.5194/amt-14-2841-2021, 2021.

Heue, K.-P., Eichmann, K.-U., and Valks, P.: TROPOMI/S5P ATBD of tropospheric ozone data products Issue 2.3, 2021.

Hurtmans, D., Coheur, P. F., Wespes, C., Clarisse, L., Scharf, O., Clerbaux, C., Hadji-Lazaro, J., George, M., and Turquety, S.: FORLI radiative transfer and retrieval code for IASI, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 1391–1408, https://doi.org/10.1016/j.jqsrt.2012.02.036, 2012.

Keppens, A., Lambert, J.-C., Granville, J., Hubert, D., Verhoelst, T., Compernolle, S., Latter, B., Kerridge, B., Siddans, R., Boynard, A., Hadji-Lazaro, J., Clerbaux, C., Wespes, C., Hurtmans, D. R., Coheur, P.-F., Van Peet, J. C. A., Van Der A, R. J., Garane, K., Koukouli, M. E., Balis, D. S., Delcloo, A., Kivi, R., Stübi, R., Godin-Beekmann, S., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 2: Ground-based validation of nadir ozone profile data products, Atmos. Meas. Tech., 11, 3769–3800, https://doi.org/10.5194/amt-11-3769-2018, 2018.

Kerridge, B. J. K., R. Siddans, Latter, B. L., Burrows, J. P., Weber, M., Beek, R. De, Aben, I., and Hartman, W.: GOME-2 Error Assessment Study. Final Report EUMETSAT Contract No EUM/CO/01/901/DK., 2002.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, Atmospheric Measurement Techniques, 8, 385–398, https://doi.org/10.5194/amt-8-385-2015, 2015.

Rodgers, C. D.: Inverse Methods for Atmospheric sounding: Theory and Practice, World Scientific, 2000.

Siddans, R.: Height Resolved Ozone Retrievals from Global Ozone Monitoring Experiment., University of Reading, 2003, 2003.

Sofieva, V. F., Szeląg, M., Tamminen, J., Kyrölä, E., Degenstein, D., Roth, C., Zawada, D., Rozanov, A., Arosio, C., Burrows, J. P., Weber, M., Laeng, A., Stiller, G. P., Clarmann, T. von, Froidevaux, L., Livesey, N., Roozendael, M. van, and Retscher, C.: Measurement report: regional trends of stratospheric ozone evaluated using the MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP), Atmospheric Chemistry and Physics, 21, 6707–6720, https://doi.org/10.5194/acp-21-6707-2021, 2021.

Sofieva, V. F., Hänninen, R., Sofiev, M., Szeląg, M., Lee, H. S., Tamminen, J., and Retscher, C.: Synergy of Using Nadir and Limb Instruments for Tropospheric Ozone Monitoring (SUNLIT), Atmospheric Measurement Techniques, 15, 3193–3212, https://doi.org/10.5194/amt-15-3193-2022, 2022.

Van Der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012, Atmos. Meas. Tech., 8, 3021–3035, https://doi.org/10.5194/amt-8-3021-2015, 2015.

Wespes, C., Hurtmans, D., Emmons, L. K., Safieddine, S., Clerbaux, C., Edwards, D. P., and Coheur, P.-F.: Ozone variability in the troposphere and the stratosphere from the first 6 years of IASI observations (2008–2013), Atmospheric Chemistry and Physics, 16, 5721–5743, https://doi.org/10.5194/acp-16-5721-2016, 2016.