Ozone_cci+

End to End ECV Uncertainty Budget (E3UB)

Date: 27/10/2021

Version: 5

WP Manager:

WP Manager Organization:

Other partners:

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

DOCUMENT PROPERTIES

Title	**End to End ECV Uncertainty Budget**
Reference	Ozone_cci+_UBR_E3UB_v5
Issue	5
Revision	1
Status	V5
Date of issue	27/10/2021
Document type	Deliverable

	FUNCTION	NAME	DATE	SIGNATURE
LEAD AUTHOR	Project partner			
CONTRIBUTING AUTHORS	Project partner	Viktoria Sofieva Carlo Arosio Melanie Coldewey-Egbers Klaus Peter Heue Katerina Garane Daan Hubert Arno Keppens J.-C. Lambert Barry Latter Christophe Lerot Richard Siddans Ronald van der A Tijl Verhoelst Catherine Wespes Michel van Roozendael		
REVIEWED BY	Project partner			
ISSUED BY	Science Leader			

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	27/10/2021	Major updates in all sections, Inclusion of the section with an overview of validation methods	

Table of Contents

1 Purpose and scope

1.1. Purpose

The 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 Level 3 data. Error budget studies in this project 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 all Level 3 ECV generated within the project.

In this document the focus is brought to activities within O3 CCI+ project, for datasets developed in CCI Phase II, we refer to the Comprehensive Error Characterization Report (CERC) document.

1.2. Reference documents

Data Standards Requirements for CCI Data Producers. Latest version at time of writing is v2.2:
https://climate.esa.int/media/documents/CCIDataStandards_v2-2_CCI-PRGM-EOPS-TN-13-0009.pdf

CCI Product Validation and Intercomparison Report (PVIR), online at:
https://climate.esa.int/sites/default/files/Ozone_cci_Phase-II_PVIR_2.0_2016.pdf

CCI Comprehensive Error Characterization Report (CERC), online at:
https://climate.esa.int/sites/default/files/filedepot/incoming/Ozone_cci_KIT_CECR_02_01_02.pdf

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.

Some teams use "smoothing error" concept: pros and cons of smoothing error are discussed in detail in von Clarmann (2014).

1.4. Acronyms

ACE-FTS	Atmospheric Chemistry Experiment – Fourier Transform Spectrometer
CCI	Climate Change Initiative
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 Cartography
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.

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.

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.

Table 1: Summary of error budget characterization and precision validation publications for Ozone_cci Level 2 datasets.

summarizes the status of publications on error budget evaluation and uncertainties validation of Level 2 ECV's generated within Ozone_cci.

Sensor	Product	Algorithm	Error budget publication	Uncertainty validation publication
GOME GOME-2 SCIAMACHY OMI TROPOMI	Total column	GODFIT 4	Lerot et al. 2014	Verhoelst et al. 2015 Sofieva et al. 2021b
GOME GOME-2 SCIAMACHY OMI	Nadir profile	RAL	Siddans et al. 1998, Siddans et al. 2003, Kerridge et al. 2002 Miles et al. 2015	Keppens et al. 2018
IASI		FORLI	Hurtmans et al. 2012 Wespes et al. 2016 Boynard et al. 2018	Keppens et al. 2018 Boynard et al. 2018
ACE-FTS	Limb profile	UToronto, v3.6/4.0	Dupuy et al., 2009	Ozone_cci CECR
GOMOS		IPF, V.6; FMI ALGOM2s v1	Tamminen et al. 2010 Sofieva et al. 2017a	Sofieva et al. 2014a Ozone_cci CECR

HALOE		v19	Brühl et al. 1996	
MIPAS		IMK Scientific	Steck et al. 2007	Laeng et al. 2015 Ozone_cci CECR
MLS		v4.2	Froidevaux et al. 2008, Livesey et al. 2008, Livesey et al. 2020	
OMPS-LP		Usask, v1.1.0	Zawada et al. 2018	Ozone_cci CECR
		UBr, v3.3	Arosio et al. 2018	In progress
OSIRIS		Usask, v5.10	Bourassa et al. 2012	Bourassa et al. 2012 Ozone_cci CECR
POAM III		NASA NRL, v4	Lumpe et al. 2002	
SCIAMACHY		IUP Scientific, v3.5	Rahpoe et al. 2013	Ozone_cci CECR
SABER		NASA, v2.0	Rong et al. 2009	
SAGE II		NASA, v7.0	Damadeo et al. 2013	
SAGE III M3M		NASA, v4	Rault et al. 2005	
SAGE III ISS		AO3, v5.1	McCormick et al. 2020	
SMR		Chalmers, v2.1	Urban et al. 2005	Ozone_cci CECR
GOME-2 OMI TROPOMI	Tropospheric column	CCD	Valks et al. 2014	
IASI		FORLI	Dufour et al. 2012, Boynard et al. 2018	Keppens et al. 2018 Boynard et al. 2018

Table 1: Summary of error budget characterization and precision validation publications for Ozone_cci Level 2 datasets.

2.1 Total ozone

Within the Ozone_cci project, the baseline algorithm for total ozone retrieval from backscatter UV sensors is the GOME-type direct-fitting (GODFIT) algorithm. During Ozone_cci project, several algorithmic improvements have been realized, with a reprocessing of the time series of GOME, OMI, SCIAMACHY and GOME-2A/B using version 4 of GODFIT. The most important update was the adaptation of the L1 soft-calibration scheme in order to restore the full independency of the satellite observations with respect to the ground-based measurements. Dominant error sources are described in Lerot et al., 2014:

- Ozone cross-sections uncertainties;
- Level-1 calibration limitations;
- Interferences with other species, including aerosols;
- Cloud contamination;
- A priori O3 profile shape, especially at large solar zenith angles.

Figure 1 shows that the mean total ozone error due to the profile shape is less than 0.5 % at low SZAs and is as large as 4% at extreme SZA for clear sky pixels. In the case of cloud contamination, the error increases, especially at low SZA.
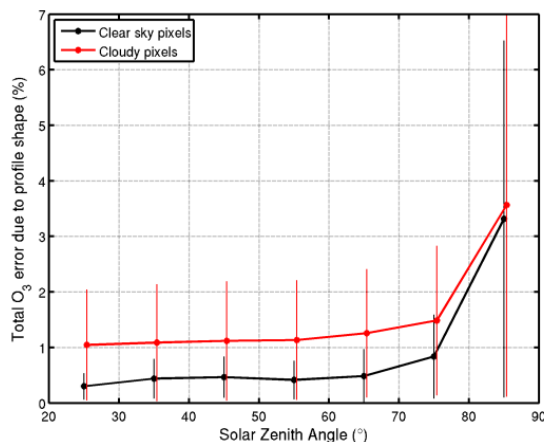

Figure 1: mean total ozone error due to a priori O_3 profile shape, as a function of the SZA for clear sky and cloudy pixels. Error bars represent the standard deviation of the errors.

Figure 2 illustrates the total ozone errors due to the neglect of aerosols in the forward model for different types of atmospheric aerosols content. These errors are generally within 1%. In particular, it is shown that the fit of the surface albedo leads to a minimization of the errors, which would be much larger without that procedure (up to 4%) in case of heavily polluted conditions. For a scenario with a strong injection of stratospheric aerosols due to a major volcanic eruption such as Pinatubo, the total errors may reach 10% (right panel).

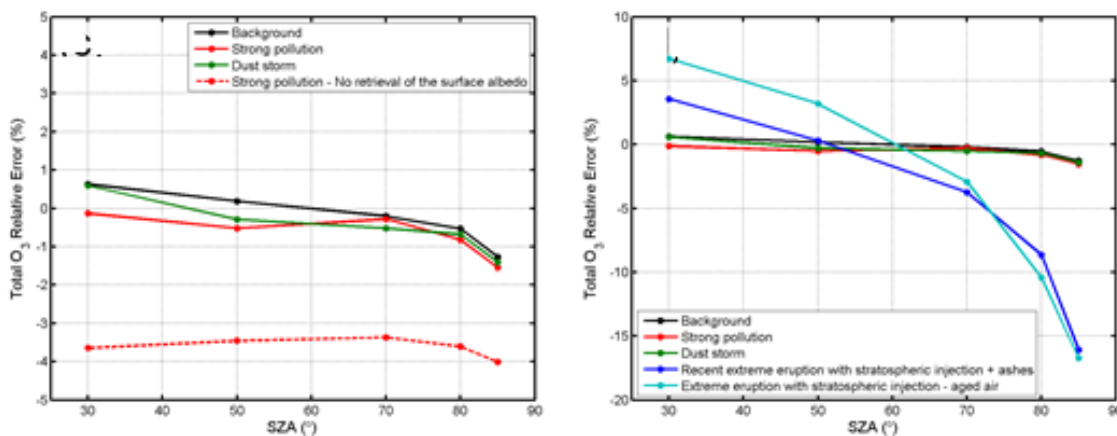*Figure 2: Left: Total ozone error (%) due to neglect of aerosols in the retrieval scheme, polluted and dust storm scenarios. Right: Same for strong volcanic eruption scenarios.*

Table 2 summarizes the assessment of the main contributions to the global error budget on total ozone retrieval by GODFIT (Lerot et al., 2014). Total errors are computed assuming all contributions are mutually uncorrelated.

Error source	Per cent error	
	SZA < 80°	SZA > 80°
Instrument signal-to-noise	< 0.5	< 2
Soft calibration: Absolute recalibration + structures removal	< 1.5	< 1.5
O_3 absorption cross-sections and its atmospheric temperature	< 2.5	< 2.5
Interferences with other species (except in case of volcanic eruption)	< 1.5	< 1
Aerosols (except in case of volcanic eruption)	< 1	< 1.5
Instrument spectral stability (wavelength registration)	< 0.5	< 0.5
Solar I_0-effect	< 0.2	< 0.2
Ring effect (Rotational Raman Scattering)	< 0.1	< 0.5
O_3 profile shape	< 1	< 4
Cloud fraction	< 0.5	< 0.5
Cloud top height	< 1.5	< 1.5
Total random error (including cloud fields)	**< 1.7**	**< 2.6**
Total systematic error	**< 3.6**	**< 5.3**

Table 2: Estimation of the error sources of the direct-fitting total ozone retrieval (single pixel retrieval). Blue fields indicate random errors (precision) associated with instrument signal-to-noise and which can be derived easily by the propagation of radiance and irradiance statistical errors provided in the level-1 products through the inversion algorithm, and red fields systematic errors. The errors due to the cloud parameters (orange) are random or systematic depending on the time scale.

Validation using independent ground-based reference measurements (reported on in the CCI Product Validation and Intercomparison Report, i.e. the PVIR) shows that the total ozone column products more than meet the official User Requirements, i.e. that the stability of the satellite TOC measurements has to be between 1 and 3% per decade (w.r.t. the truth/reference) and that the short-term variability (due to measurement uncertainty) has to be less than 3 %. In detail, it was found that:

- the individual Level-2 data sets show excellent inter-sensor consistency with mean differences within 1.5% at moderate latitudes (+/-50°);
- the mean bias between GODFIT v4 satellite and Brewer and Dobson reported TOCs is well within 1.5 ± 1.0 % for all sensors;
- the drift per decade spans between -1.5 to +0.5 %, depending on the sensor.

The peak-to-peak seasonality ranges between ~0.7 % for GOME, GOME-2C and OMPS, to ~2% for SCIAMACHY.

Sofieva et al., (2021b) applied the structure function method for validation of TROPOMI random uncertainty estimates. It was found that the random uncertainties reported by the TROPOMI inversion algorithm, which are in the range 1-2 DU, agree well with the experimental uncertainty estimates by the structure function, for clear-sky conditions. This is also confirmed by the consistency test presented in 4.4. In cloudy conditions, the experimental uncertainties (ex-post) are higher than the reported by the inversion algorithm (ex-post), because pseudo-random errors due to presence of clouds are not characterized by the inversion algorithm at the moment.

2.2 Ozone profiles from nadir sensors

The status of publications on error budget evaluation and uncertainties validation of Level 2 ozone profiles from nadir sensors generated within Ozone_cci+ is presented in Table 1.

In the project, the ozone profiles are processed with RAL and FORLI algorithms

2.2.1 **RAL processor**

Analysis of error budget of RAL scheme, reported in (Siddans, 1998, 2003), is based on performing retrieval simulations for a set of basic geo-physical scenarios, which had been defined for the GOME-2 Error Study (Kerridge et al., 2002). Figure 3 shows retrieval precision and base-line mapped errors for GOME-1 and the April 55ºN 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 B1 (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.
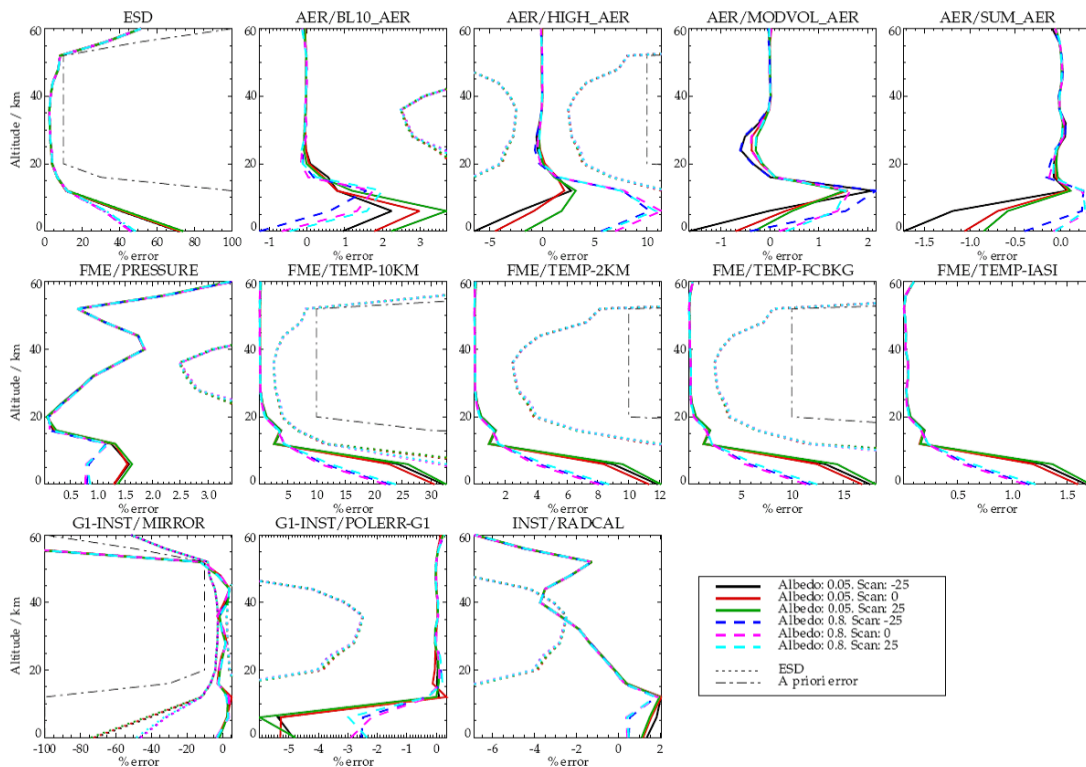
Figure 3: retrieval precision and base-line mapped errors for GOME-1 and the April 55ºN scenario.

Miles et al., (2015) assessed the performance of the RAL ozone profile retrieval scheme for the 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 4 for a mid-latitude profile in Northern Hemisphere summer. In this case, the retrieval precision on retrieval levels is typically much smaller than the a priori error throughout the profile. The retrieval noise error is around a factor of 2 smaller than the retrieval precision. Figure 5 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.
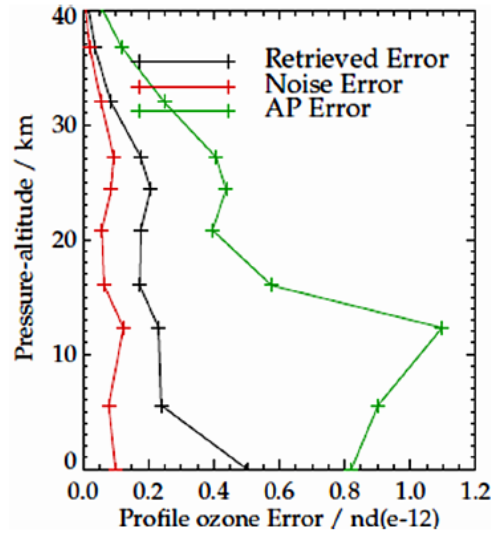
Figure 4: error assessment for ozone profile for a GOME-2 nadir pixel at 45ºN on 25 August 2008.
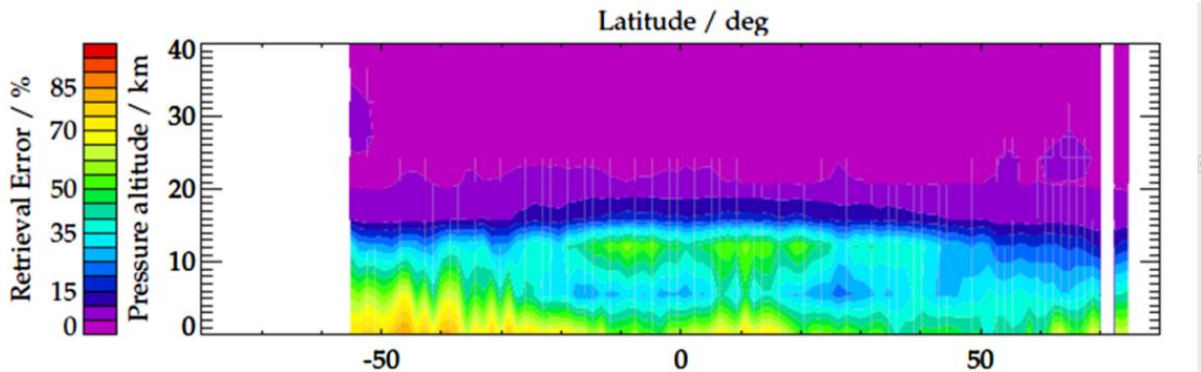

Figure 5: relative retrieval error of ozone product from GOME-2 retrieved with RAL scheme.

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 % for GOME

- insignificant (except for the tropics) drift on the order of −15 and 10 % for OMI's L2 stratospheric and tropospheric observations, respectively;
- a significant positive drift of ~40 % decade^{-1} for SCIAMACHY and GOME-2A below the tropopause.
- a significant 30 % decade^{-1} negative drift in the UTLS at all latitudes for GOME-2A.

2.2.2 IASI FORLI processor

The estimated statistical uncertainties on the ozone vertical profiles retrieved from FORLI (v20140922) are calculated for three latitude bands in Wespes et al., (2016) and are shown in Figure 6. This total retrieval error depends 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 (upper troposphere–lower stratosphere), except in the equatorial belt (above 30%) due to the stronger influence of interfering water lines and the low O$_3$ amounts which leads to larger relative errors (Wespes et al., 2016). The retrieval errors are usually less than 5 % in the stratosphere. The error is larger above cold polar surfaces, characterized by low sensitivity and a likely misrepresentation of the surface emissivity (Hurtmans et al., 2012, Wespes et al., 2016).

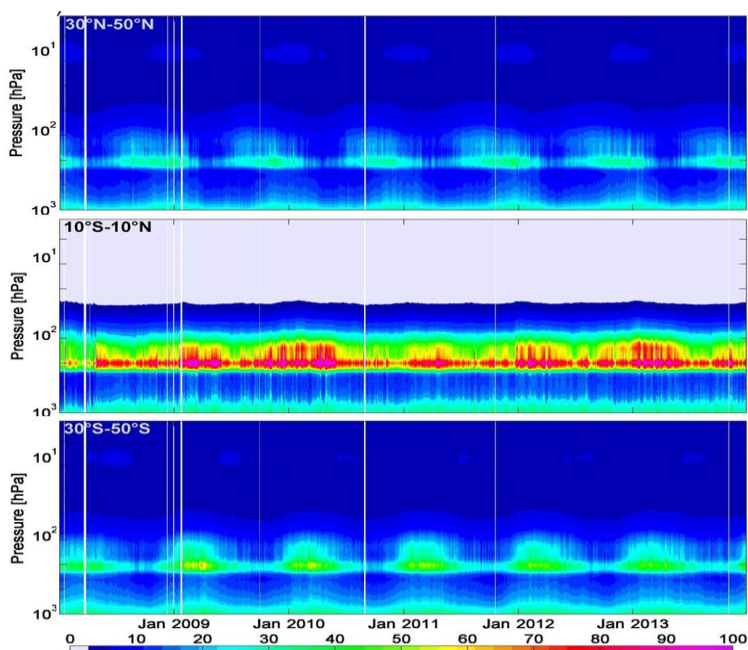*Figure 6: Daily estimated total retrieval errors (%) for the period 2008-2013 as a function of time and altitude, for three latitude bands: 30-50°N, 10°s-10°N and 30-50°S.*

As shown in Figure 7, the main contributions to the total error are the limited vertical sensitivity (smoothing error), the measurement noise and uncertainties on the fixed parameters, such as

surface emissivity and temperature profiles, but in the routine processing of the error matrix, the latter are not taken into account.
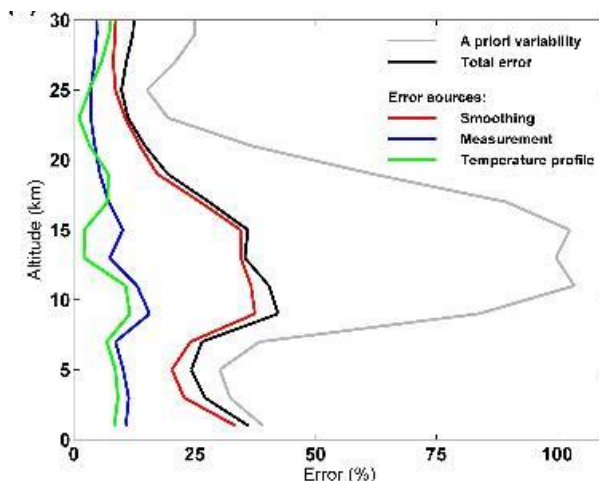

Figure 7: Error budget analysis for IASI ozone retrieval (Boynard et al., 2009). The a priori variability and total errors are given by the square root of the diagonal elements of the a priori covariance matrix and the error covariance matrix, respectively.

There is no significant bias due to instrument aging: when comparing IASI/MetOp-A and IASI/MetOp-B total ozone columns for the year 2014 (FORLI-v20151001), the bias is generally within 0.5% as shown by Figure 8.
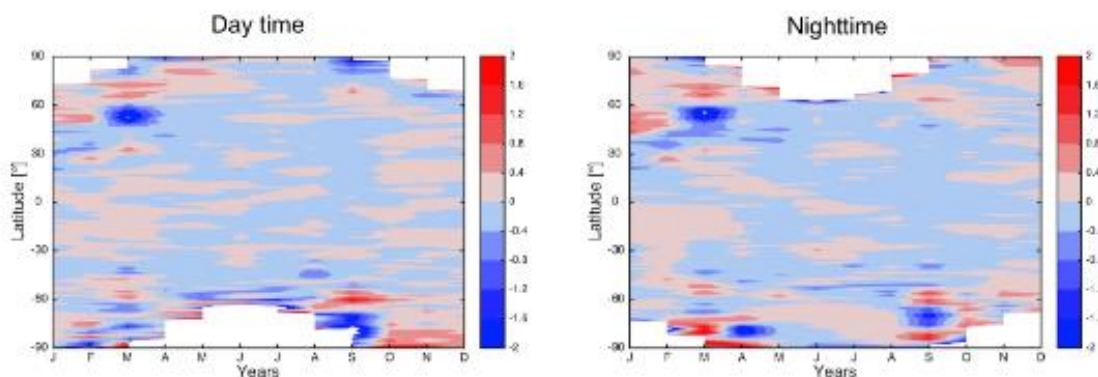
Figure 8: Contour representation of the relative difference (in percent) between IASI-A and IASI-B total ozone column retrieved using FORLI as a function of latitude and time for the year 2014 for day time data (left) and night-time data (right). The relative differences are calculated as 100 x (IASI-A – IASI-B) / IASI-A (Boynard et al., 2018).

The IASI FORLI-O_3 profiles dataset (v20151001) has been extensively validated in Boynard et al., (2018) and Keppens et al., (2018). They demonstrated a good degree of accuracy, precision and vertical sensitivity, and no instrumental drift, The retrieval data products showed an insignificant stratospheric bias usually smaller than 10 %, a large positive bias (10 to 40 %) in the UTLS, likely due to strong ozone variability in that poor-O3 region inducing inadequate a priori

information and larger total retrieval errors (Wespes et al., 2016), and and a ~4 to 19 % bias in the troposphere, depending on latitudes. Keppens et al., (2018) also reports that the ex-ante IASI uncertainties provided in L2 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 (v20151001) nadir ozone profile products.

Typical uncertainty values and retrieval characteristics are reported in the table below:

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 & in the UTLS; <5% in the lower & the middle-upper stratosphere.
Systematic errors: Uncertainty in cross-sections Temperature uncertainty	~4% <10% over all the profile

Table 3: IASI ozone profiles characteristics and error budget

2.3 Ozone profiles from limb sensors

In the following table the links to the limb data sets from Table 1 are listed:

Sensor L2 data	Link to L2 data

ACE-FTS	https://database.scisat.ca/level2/
GOMOS ALGOM2s	https://earth.esa.int/web/sppa/activities/instrument-characterization-studies/algom/data-resources
HALOE	https://disc.gsfc.nasa.gov/datasets/UARHA2FN_019/summary
MIPAS IMK	https://www.imk-asf.kit.edu/english/308.php#org0f1a3a1
MLS	https://mls.jpl.nasa.gov/eos-aura-mls/data-access
OMPS-LP Usask	https://arg.usask.ca/projects/omps-lp/
OSIRIS	https://research-groups.usask.ca/osiris/data-products.php
POAM III	https://search.earthdata.nasa.gov/search?q=POAM%20III
SCIAMACHY IUP	https://www.iup.uni-bremen.de/scia-arc/
SABER	http://saber.gats-inc.com/data.php
SAGE II	https://zenodo.org/record/3710518#.YW7oahwxnb0
SAGE III M3M	https://asdc.larc.nasa.gov/project/SAGE%20III-M3M/g3assp_4
SAGE III ISS	https://asdc.larc.nasa.gov/project/SAGE%20III-ISS/g3bssp_52
SMR	http://odin.rss.chalmers.se/level2_download/

Table 4:Links to the limb L2 data sets listed in Tab 5.

2.3.1 **ACE-FTS**

The uncertainties reported in the data files are the statistical fitting errors from the least-squares process and do not include systematic components or parameter correlations (Boone et al., 2005). The mean relative fitting errors are lower than 3% between 12 and 62 km and typically less than 2% around 30–35 km, as shown in Figure 9.
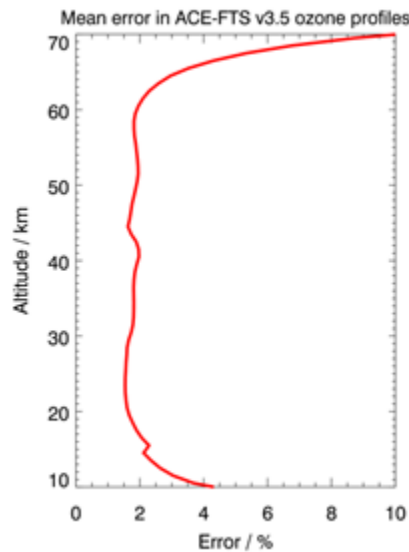
Figure 9: mean uncertainties reported in ACE-FTS v3.5 ozone vertical profiles

There is no simple function to use to calculate the vertical resolution, so the data provider estimates a value of 3 km as an average for all measurements.
The global distribution of absolute uncertainties reported with ACE-FTS ozone vertical profiles is shown in Figure 10. The absolute uncertainty values are higher in the tropical region, while relative uncertainties are higher in polar regions.
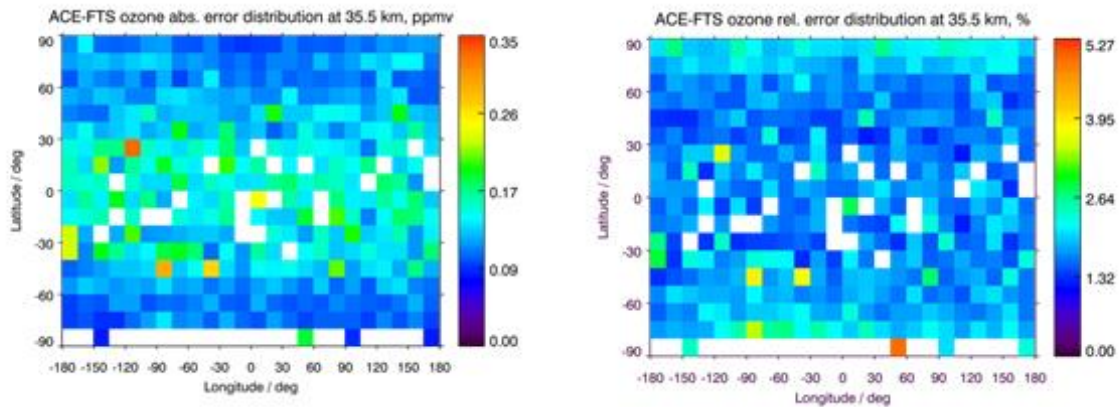
Figure 10: distribution of uncertainties for ozone products from ACE-FTS on 35.5 km height.

Evolution of ACE-FTS ozone uncertainties with time are shown on Figure 11. One can see that the uncertainties are slightly growing with time, going from 1.7% in the beginning of the mission to 2.0% in the recent period. The significance of this estimation is in progress.
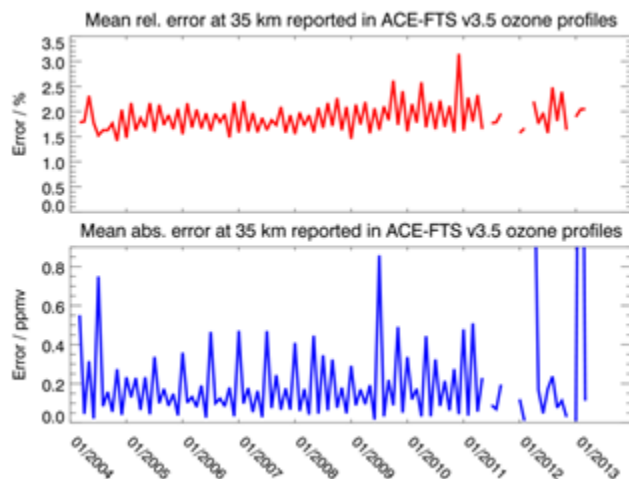
Figure 11: evolution with time of uncertainties of ACE-FTS ozone vertical profiles at 35 km height. Top panel: relative uncertainty, bottom panel: absolute uncertainty.

Sheese et al., (2021) presented an extensive validation of v3.6 and v4.0 ACE-FTS data sets. They report that a significant in v3.6 O3 at all altitudes above 20 km was detected, whereas no drift in v4 at any analysed altitude was found.

2.3.2 GOMOS ALGOM2s

In the CCI project, the new GOMOS data processed with ALGOM2s v.1 Scientific Processor are used (Sofieva et al., 2017a). The error propagation scheme is similar to that used in GOMOS IPF v.6 processor, as the ALGOM 2S ozone profiles are identical to those of IPF v.6 in the stratosphere, and differ in the UTLS. The error estimates (square roots of the diagonal elements of the covariance matrix) are provided in the Level 2 data. The covariance matrix of retrieved profile uncertainties is obtained via Gaussian error propagation through the GOMOS inversion, see Tamminen et al., (2010) for details. Both noise and the dominating random modelling error (due to scintillations) are taken into account in GOMOS inversion. Thus, error estimates provided in Level 2 files represent the total precision estimates.

The precision of GOMOS ozone profiles depends on stellar brightness, spectral class and obliquity of occultation. Typical values of ozone precision values based on real GOMOS data are presented in Figure 12 which shows GOMOS precision estimates of ozone for representative cases: bright star (first column), typical star (middle column) and dim star (last column). The dashed lines correspond to oblique occultations (O) and the solid lines to vertical (in orbital plane, V) occultations. Stellar temperature is indicated with the line colour: hot stars (red), medium stars (green) and cool stars (blue). The uncertainty values are from GOMOS data processed with IPF version 6.
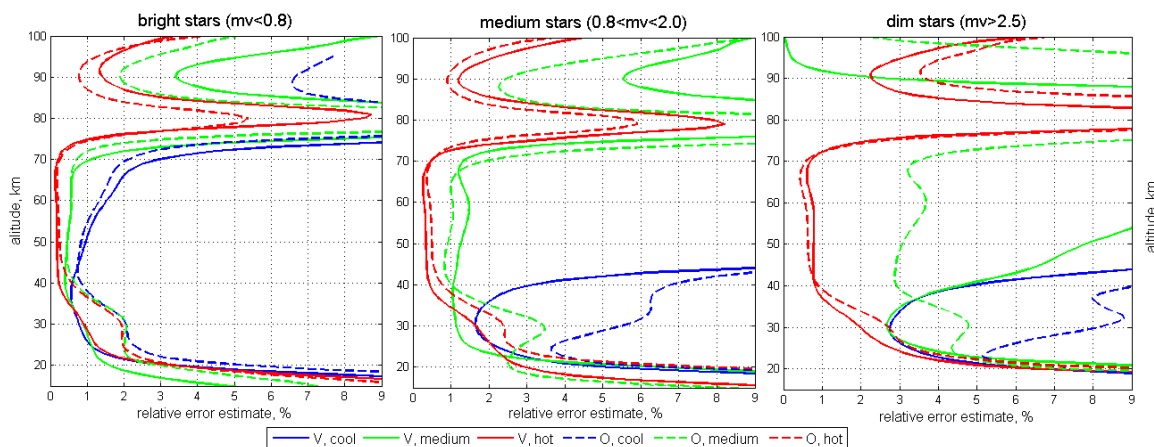
Figure 12: Typical values of ozone precision values based on real GOMOS data.

Other sources of systematic errors are imperfect modelling of the aerosol extinction, uncertainties in the absorption cross sections and temperature. Uncertainties of air density profile, ray tracing and potentially missing constituents have a negligible impact on ozone retrievals. The characteristics of GOMOS ozone profiles together with the random and the systematic errors are summarized in Table 6.

Altitude range	15-100 km
Vertical resolution	2 km below 30 km, 3 km above 40 km
Random errors: measurement noise and scintillations	0.4-4% stratosphere, 2-10 % MLT, ~10% at 15 km
Systematic errors: Uncertainty in cross-sections Aerosol model selection Temperature uncertainty Air density uncertainty	~1 % ~20% below 20 km, 1-5% at 20-25 km, <1% above 25 km <0.5% at 30-60 km, negligible elsewhere <1% below 20 km, negligible elsewhere

Table 5:GOMOS ozone profiles characteristics and error budget.

The validation of GOMOS random uncertainty estimates is performed using the differential method and reported in Sofieva et al., (2014a). It was shown that GOMOS random uncertainty estimates are realistic for not-dim stars.

2.3.3 HALOE

An extensive assessment of uncertainties on L2 profiles is presented in Brühl et al., (1996).
The total estimated error was calculated by taking the root-sum-square of all the individual error sources. The error estimates were obtained by simulating a profile with an accurate forward model, by using realistic temperature and pressure information, H_2O, O_3, and aerosol extinction profiles. The reference profile was then perturbed to simulate the various types of error sources and the error estimate could then be derived by comparing the resulting retrievals to the unperturbed profile

retrieval. In the upper stratosphere the most important contribution comes from retrieval noise and the systematic uncertainty on pointing knowledge. In the lower stratosphere instead, model errors, pointing accuracy and aerosol extinction play relevant roles.

2.3.4 **MIPAS IMK Scientific**

Figure 13 shows the MIPAS ozone error budget for full spectral resolution (FR) period of MIPAS instrument (2002-2004) averaged over one orbit. Below 15 km the percentage errors are rapidly increasing to values in the order of 25% for polar and midlatitude conditions or more than 50% for tropical conditions, where the vmr is small. The error in vmr remains below 0.5 ppmv. The estimated random error is dominated by the instrumental noise above 14 km. Below 14 km, the error due to uncertain water vapor concentration becomes dominant because water vapor increases exponentially with decreasing altitude, these strong water vapor lines are slightly interfering with ozone lines leading to a dependence of the retrieved ozone on the pre-retrieved water vapor amount.

Figure 13: Estimated ozone error budget of MIPAS for a typical FR mid-latitude retrieval (night and day). Left: absolute, right: percentage errors.

The total error is dominated by uncertainties in spectroscopic data (dark blue). The altitude-dependence of errors due to spectroscopic data is due to the fact that the microwindows used in the retrieval are varying with altitude. Errors caused by uncertainties in the ILS (instrumental line shape) are in the order of 1 to 4% and thus nearly negligible compared to spectroscopic uncertainties.

Figure 14 reports a similar error budget analysis of MIPAS IMK ozone retrieval for reduced spectral resolution (RR) period of MIPAS instrument (2005-2012). Spectroscopy is still the leading error source. The error in vmr at heights below 15 km remains below 0.1 ppmv.
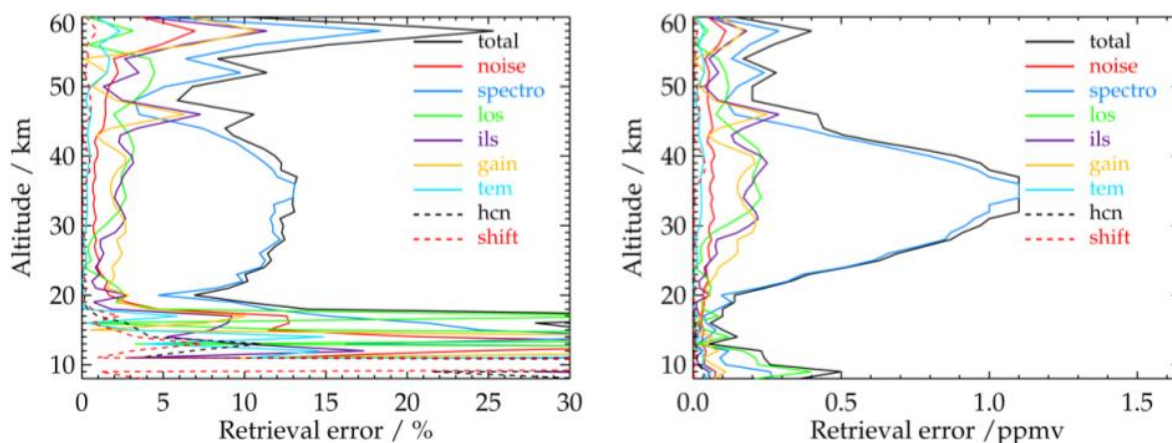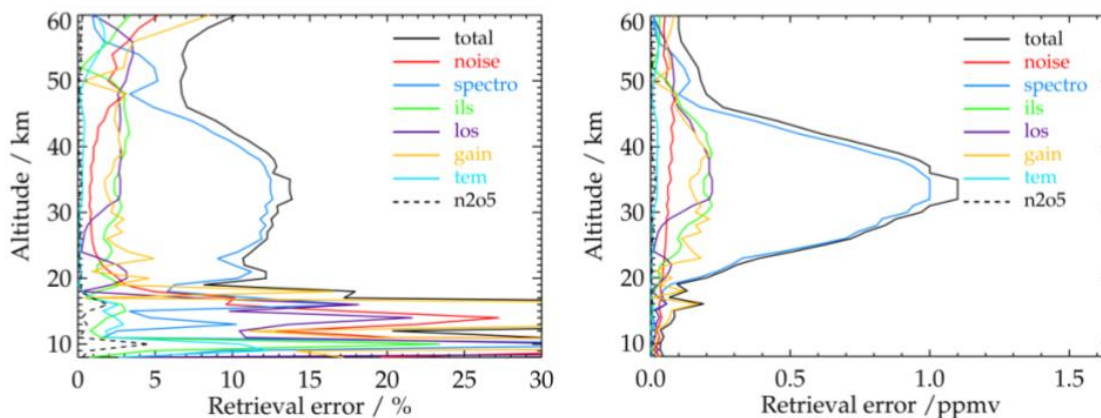
Figure 14: Estimated ozone error budget of MIPAS for a typical RR mid-latitude retrieval (night and day). Left: absolute, right: percentage errors.

2.3.5 **MLS v4.2**

The typical vertical resolution, precision and accuracy values of the v4.2 MLS ozone profiles are reported in table below (from Livesey et al., 2020). In order to assess the accuracy component, estimated systematic errors are propagated, e.g., from calibration and spectroscopy. Details in Froidevaux et al., (2008), Livesey et al., (2008).

Several quantities are provided in the Level 2 data, such as convergence, status and quality: detailed description of their use is provided in the user guide Livesey et al., (2020).

A known artefact of retrieved profiles is oscillations in the tropical UTLS. The very good long-term stability of the data set has been assessed by Hubert et al., (2016).

Pressure	Vertical resolution	Precision [%]	Accuracy [%]
0.02	5.5	300	50
0.05	5.5	150	30
0.1	4	60	20
0.2	3	30	10
0.5	3.5	20	10
1	3	7	10
2	3.5	3	7
5	3	2	7
10	3	2	6
22	2.5	2	5
46	2.5	3	8
68	2.5	4	7
100	3	15-25	[+0.005 +7%]
150	3	5-70	[+0.005 +7%]
215	3.5	5-100	[+0.01 +10%]
261	3.5	5-100	[+0.02 +10%]

Table 6: characterization of MLS v4.2 ozone profiles

2.3.6 **OSIRIS Usask v 5.10**

Typical relative precision of the OSIRIS ozone retrievals is within 3–5% over the 20 - 50 km altitude range, showing an increase at lower altitudes, due to decreasing sensitivity, and at higher altitudes, due to decreasing limb scattered signal (Bourassa et al, 2012). This estimate was found to be consistent with the standard deviation of pairs of OSIRIS retrievals that are close in time and space and located in the tropics where natural variability is weak. In addition, OSIRIS precision were validated using a set of coincident SAGE II measurements. The results showed that the relative precision is typically a few percent higher than that determined through the numerical algorithm.

To estimate the OSIRIS ozone error budget, a random sampling of scans was chosen and the ozone was repeatedly retrieved with randomly perturbed inputs. The inputs were adjusted by a random factor chosen from a normal distribution of values with a 3σ of 10%. This was performed in turn for the aerosol profile, albedo, neutral density profile, and NO_2 profile. For the altitude registration a 3σ of 300m was used. The total error is calculated using a sum in quadrature of the error components and was found to be dominated by the precision.

Figure 15 illustrates the dominance of the precision over the total error budget, which peaks around 7% at approximately 15 km. This is followed by contributions from potential errors in altitude registration, which provides about 2% uncertainty above 35km and below 20km. Errors in the neutral density potentially contribute up to 2% uncertainty at the lowest bounds of the retrieval and are negligible above 30km. Errors from other sources are much less than 1% at all altitudes.
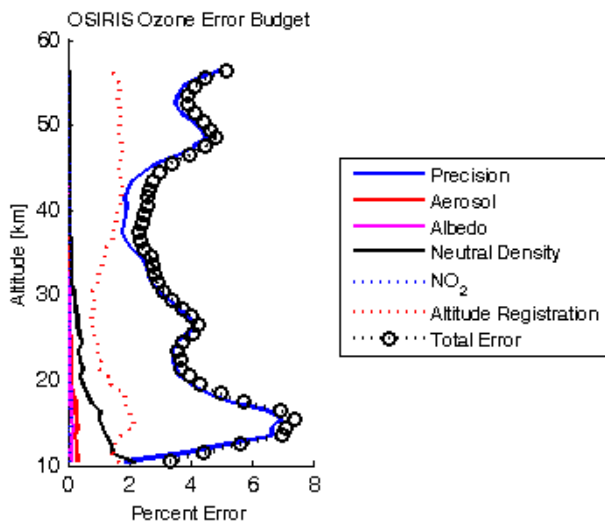
Figure 15: Dominance of the precision over the total error budget of OSIRIS.

A drift was found in the time series starting from 2012 and corrected at L2 by applying a RSAS pointing correction to the limb radiances before the retrieval (Bourassa et al., 2018, v5.10 ozone profiles).

2.3.7 **OMPS-LP Usask 2D**

The OMPS-LP USask 2D retrieval process uses Gaussian error propagation to estimate the covariance of the retrieved solution. Currently only the random error component of the radiance measurements is accounted for (Zawada et al., 2018). The reported precision is the square root of the diagonal elements of the converged solution covariance matrix, which is plotted in Figure 16. Smoothing error is not included in the reported error estimate, however representative averaging kernels are available as diagnostic quantities.
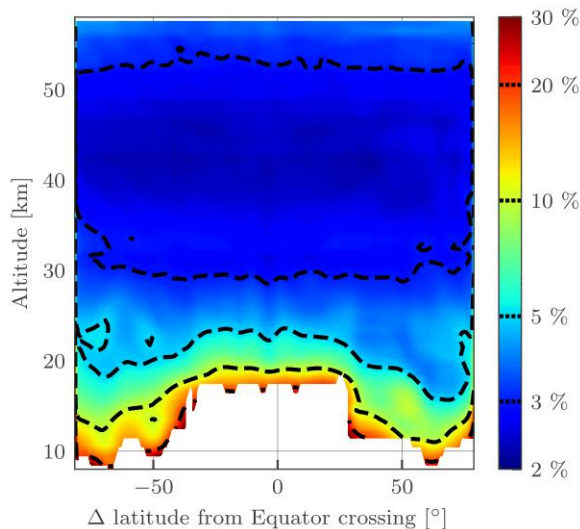
Figure 16: Measurement noise estimate for one orbit of OMPS-LP, Zawada et al., (2018).

2.3.8 **OMPS-LP UBr**

Uncertainty estimation for OMPS-LP retrieval at the UBr is ongoing. Arosio et al., (2018) presented the retrieval algorithm and its characterization: the measurement noise is typically within 2 and 5% in the range 20-50 km, the measurement response is close to 1 and the vertical resolution about 2.5 km above the tropopause with values up to 4 km at 30-35 km, as shown in Figure 17.
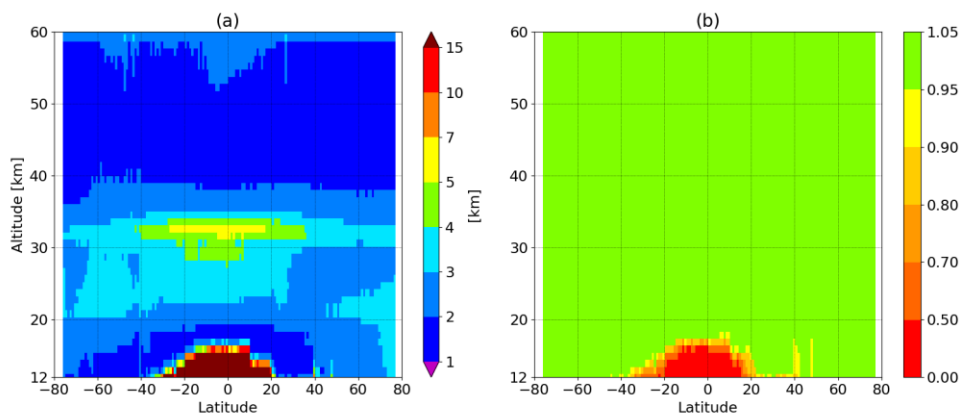
Figure 17: Vertical resolution and measurement response for OMPS-LP retrieval at UBr.

A study is ongoing to estimate an error budget for the retrieved profiles: preliminary results to assess a total random uncertainty and systematic bias for mid-latitudes are shown in Figure 18. 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 uncertainties was found to be related to the used ozone cross section. In particular, changing the used ozone cross section in the UV leads to variations up to 2 % above 30 km. The cross section error propagated into L2 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 %.

Uncertainties have been assessed also using self-collocated OMPS-LP measurements at high latitudes, showing that the standard deviation of the collocated profiles approaches the typical retrieval noise values, whereas the total uncertainty estimated in a root mean square fashion is an upper value of the spread of L2 profiles.
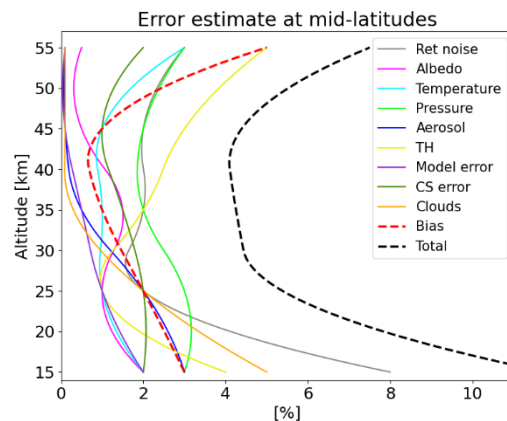
Figure 18: Error budget estimate for OMPS-LP ozone profiles at UBr.

2.3.9 **POAM III**

Detailed description of retrieval, uncertainties and vertical resolution can be found in Lumpe et al., (2002). The total error provided in v4 of Level 2 data is the rms of 3 error sources: a total random error obtained from uncertainty propagation, a component related to sunspots and an aerosol feedback loading error. Uncertainties on ozone profiles are found to be around 5% or less between 13 and 60 km, increasing to 10% by 10 km. The timing uncertainty is the largest error component above 20 km, but below 15 km the measurement noise rapidly becomes the dominant error source. The typical values of the vertical resolution are reported in the user guide and in Lumpe et al., (2002), and is ~1 km between 15 and 50 km.

2.3.10 **SABER**

Rong et al., (2009) present a validation of SABER observations and report an error budget estimation by performing synthetic retrievals. The total precision in the stratosphere varies from 8% to 1%, with the middle to upper stratosphere (<10 hPa) showing the best precision.

The authors performed an empirical estimate of the precision by using up- and down-scan pairs that should have high statistical repeatability since they are measured in very close local times and space. The two estimates were found to agree well.

The SABER vertical resolution is ~2 km for all channels and all altitudes (Rong et al., 2009).

2.3.11 **SCIAMACHY IUP Sciatran**

Total systematic ($\pm\sigma_{sys}$) and random ($\pm\sigma_{rnd}$) errors for retrievals of ozone profiles with SCIATRAN processor are calculated, for the three latitude bands and different altitudes in (Rahpoe et al., 2013) and shown as profiles in Figure 19. The contribution to total systematic error is coming from the aerosol (up to 15 %), albedo (up to 8 %), tangent height (up to 8 %), temperature (up to 1 %), and pressure (up to 2 %). The maximum random error is in the order of 43 % in the tropics at 10 km.
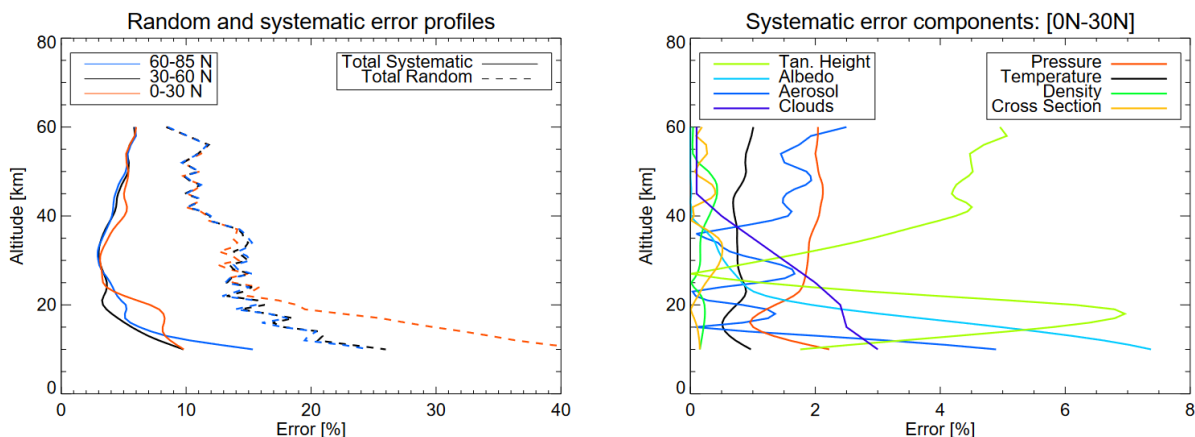
Figure 19: total systematic and random error profiles for three latitude bands for SCIATRAN ozone retrievals.

2.3.12 **SAGE II**

Damadeo et al., (2013) discussed v7.0 retrieval of ozone profiles from SAGE II, relevant uncertainties and validated the L2 data. Typical retrieval noise values are within 1 % between 18 and 52 km. No extensive error budget study is present nor uncertainty validation.

2.3.13 **SAGE III M3M (Solar occultation)**

SAGE III measurements are provided with uncertainty estimates for random components. Systematic uncertainties are normally secondary and can be assessed through sensitivity analysis.

Three are the primary sources of the random component of the uncertainty: the line-of-sight optical depth measurement errors, the Rayleigh optical depth estimate, and the removal of contributions by interfering species. Retrieval errors are estimated in Rault (2005), based on the inversion algorithm covariance matrices. The largest sources of uncertainty are the altitude registration, the stray light removal process and the dark current evaluation.

A constant vertical resolution of 0.5 km is assumed for all profiles and altitudes.

2.3.14 SAGE III ISS (Solar occultation, Least Square Ozone)

Error analysis for SAGE III ISS ozone profiles is ongoing, first validation of the results can be found in McCormick et al., (2020).

A constant vertical resolution of 0.5 km is currently assumed as a preliminary average value for all profiles and altitudes.

2.3.15 SMR/Odin

An advanced analysis of systematic uncertainties of ozone retrieval from ODIN/SMR 501.8 GHz band was performed in (Urban et al., 2005). Its description of uncertainty estimates can be found in CECR.

3 Uncertainty of level 3 data

Level 3 data are created using different spatio-temporal averaging. The associated uncertainties are usually estimated as a standard error of the mean:

$$\sigma_{mean}^2 = \frac{s^2}{N},$$

Eq. 1

where $s^2 = \left\langle (x_k - \overline{x})^2 \right\rangle$ is the sample variance and N is the number of measurements in a spatio-temporal bin. In addition, parameters characterizing spatio-temporal inhomogeneity are provided with the datasets. Below the details of Level 3 uncertainties are presented.

3.1 Monthly mean single instrument measurements

3.1.1 Ozone total column Level 3 data

The ESA-CCI total ozone level 3 total column ECV product contains monthly averages on a fixed global grid of 1° x 1° in latitude and longitude for six nadir-viewing satellite sensors: GOME/ERS-2, SCIAMACHY/ENVISAT, OMI/NASA-Aura, GOME-2/MetOp-A, GOME-2/MetOp-B, and TROPOMI/S5-P. A detailed description of the data records can be found in (Coldewey-Egbers 2015) and in Garane et al., (2018). Level-2 measurements processed with the GODFIT v4 retrieval algorithm (Lerot 2014) are mapped onto a regular global grid of 1°×1° in latitude and longitude to construct daily and monthly averages for each sensor. Besides the monthly mean total ozone column, the corresponding standard deviation, the standard error, and the number of measurements per month are provided. The sample standard deviation characterizes the scatter of the measured data encompassing the natural variability, the measurement error as well as the sampling uncertainty. Figure 20 shows standard deviations in April 1997 for GOME, in April 2005 for SCIAMACHY and in April 2008 for GOME-2.

Figure 20: Standard deviation of Level 3 ozone total column product in April 1997 for GOME (top left panel), in April 2004 for SCIAMACHY (top row, middle panel), in April 2006 for OMI (top right panel), in April 2008 for GOME-2A (bottom left panel), in April 2012014 for GOME-2B (bottom row, middle panel), and in April 2019 for TROPOMI (bottom right panel).

The standard error quantifies the spatial-temporal sampling errors inherent to the satellite measurements. These errors have been estimated using an Observing System Simulation Experiment (OSSE). High-resolution ECMWF (European Centre for Medium-Range Weather Forecasts, www.ecmwf.eu) data were taken as the reference data set. Then, daily observations were simulated from the reference using the sampling patterns appropriate to the individual sensors. Finally, the average monthly simulations are compared with the corresponding monthly reference in order to estimate the sampling errors corresponding to the total ozone monthly averages. Figure 21 shows standard error in April 1997 for GOME, in April 2005 for SCIAMACHY and in April 2008 for GOME-2.
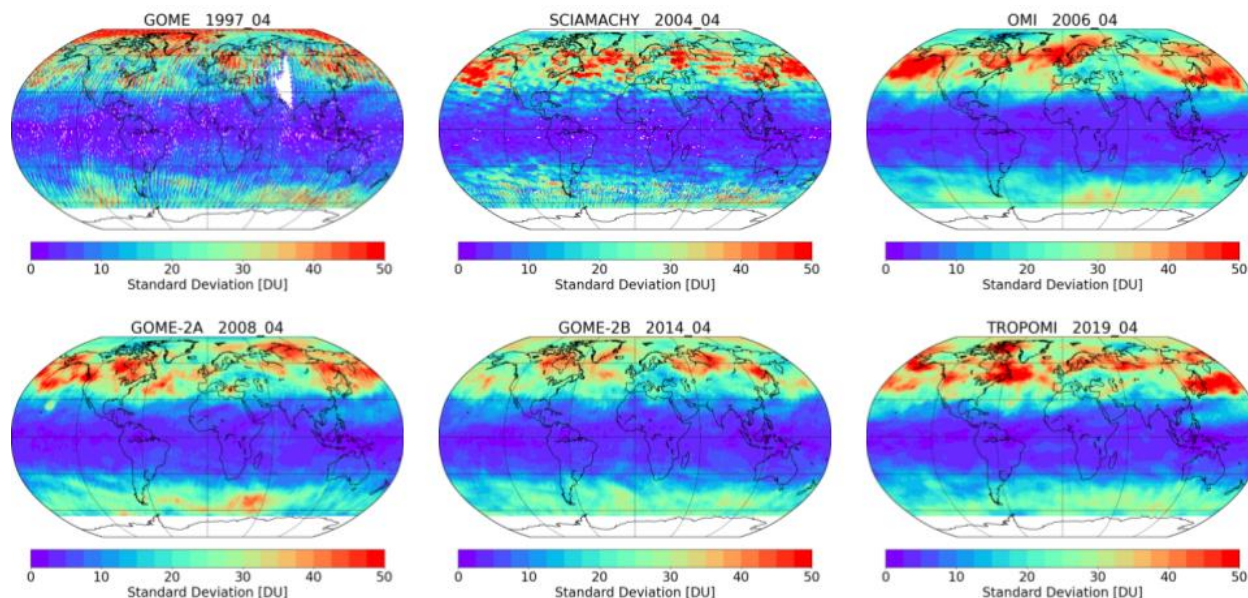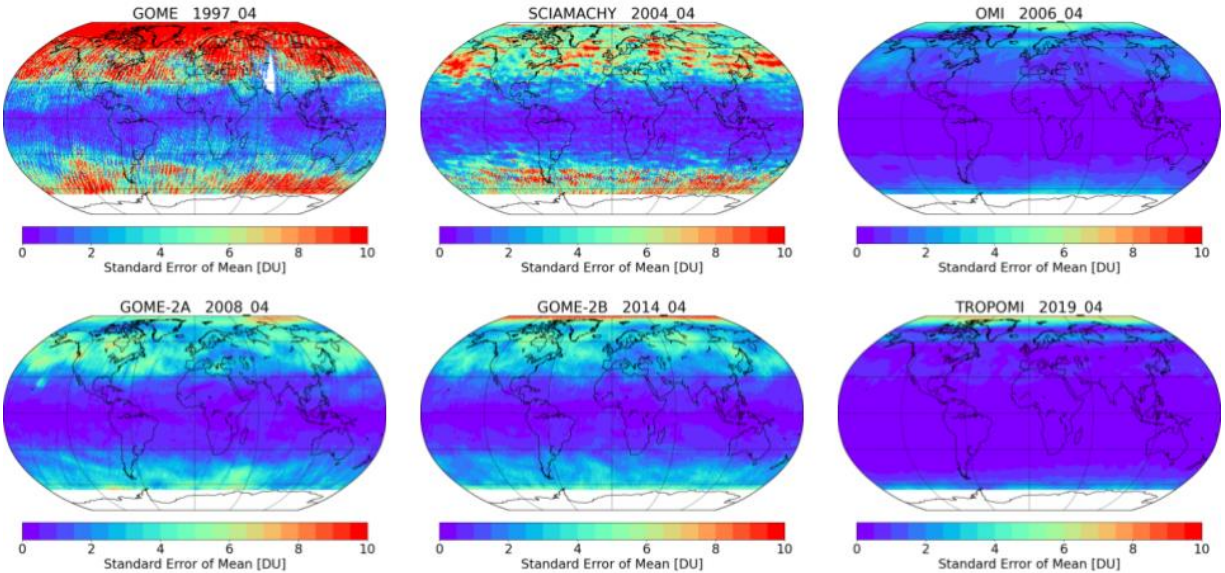
Figure 21: Standard error of Level 3 ozone total column product in April 1997 for GOME (top left panel), in April 2004 for SCIAMACHY (top row, middle panel), in April 2006 for OMI (top right panel), in April 2008 for GOME-2A (bottom left panel), in April 2014 for GOME-2B (bottom row, middle panel), and in April 2019 for TROPOMI (bottom right panel).

3.1.2 Ozone nadir profiles Level 3 data

The average value in a level-3 grid cell is a weighted average of all values assigned to that grid cell (and for that layer). The weights used for the averaging are equal to 1/variance, i.e. 1/(error^2) on the individual parameter. Nadir and off-nadir pixels are treated in the same way, which means that though the errors on individual profiles may have systematic differences across-track, the mathematical treatment is the same. If the data are uncorrelated, this estimate is optimal in the sense that it gives the smallest possible error. In mathematical notation the mean is calculated as:

$$mean = \frac{\sum \frac{x_i}{error_i{}^2}}{\sum \frac{1}{error_i{}^2}}$$

The error on the averaged values is the standard error of the weighted mean. With variance as weights, this error is calculated as:

$$Std_{Err} = \sqrt{\left(\frac{1}{\sum \frac{1}{error_i{}^2}}\right)}$$

The ozone values x_i and the associated $error_i$ come from the Level 2 profile data and are interpolated in the vertical to the standard Level 3 vertical grid.

The pixels in the satellite data (L2) are assumed to be ordered as indicated in Figure 22.
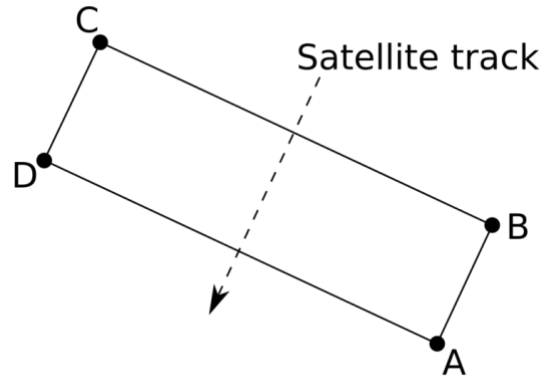
Figure 22: Pixel layout assumed in the nadir L3 algorithm.

If this is not the case, the reading routine should provide the appropriate transformation. **A** is the first corner in the longitude and latitude arrays, **B** the second etc. The across track direction is given by the lines the lines **A**-**D** and **B**-**C**, while the along track direction is given by the lines **A**-**B** and **D**-**C**. Note that corners **C** and **D** are reversed with respect to the GOME/GOME-2 convention.

The along track pixel edges **AB** and **DC** and cross track pixel edges **AD** and **BC** (see Figure 22) are divided into a number of points. The first point on **AB** and the first on **DC** form a line which is divided into the same number of points as **AD**. Each of these points is assigned to a grid cell, see for example Figure 23.
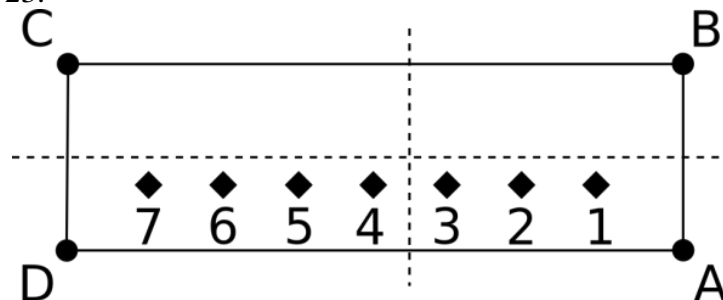
Figure 23: A L2 pixel is divided into subpixels (diamonds 1-7). Each subpixel is assigned to a TM5 grid cell (dashed) and the average and standard deviation are calculated.

Suppose that ABCD in Figure 23 is the pixel of interest and that the horizontal line marked with the diamonds are the subpixels (numbered 1 to 7). Furthermore, the two dashed lines denote the grid cell boundaries which are numbered the same way as the pixel corners (i.e. grid cell A is the lower right cell). In this case, subpixels $1 \sim 3$ are added to grid cell A, and the counter for grid cell A is increased by 3. Subpixels $4 \sim 7$ are added to grid cell D and the counter for grid cell D is increased by 4. The pixel values are weighted by $1/\sigma^2$ before adding, so the weighted mean

gridcell value and the corresponding standard deviation are given by $mean = \frac{\Sigma_i \frac{x_i}{\sigma_i^2}}{\Sigma_i \frac{1}{\sigma_i^2}}$ and $sdev =$

$\sqrt{\frac{1}{\Sigma_i \frac{1}{\sigma_i^2}}}$. These values are provided for partial columns in the L3 files on a layer-by-layer basis and

for the total column. An example is shown in Figure 24 for January 2008, based on the L2 dataset provided in phase 1 of the ozone CCI project.
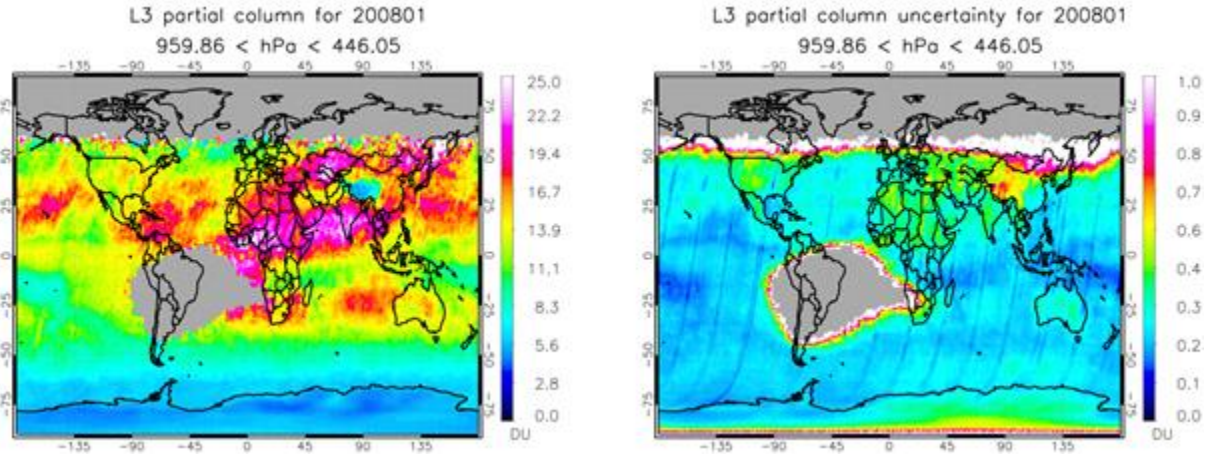
Figure 24: mean partial ozone column (left) and its uncertainty (right) for January 2008, based on L2 data provided in the first phase of the project.

3.1.3 **Ozone limb profiles Level 3 data**

Monthly zonal mean data from the individual limb instruments are computed in 10° latitude bands from 90°S to 90°N. For all sensors, the monthly zonal average is computed as the mean of ozone profiles. The uncertainty of the monthly mean is estimated as the standard error of the mean:

$$\sigma_\rho^2 = \frac{s^2}{N},$$ Eq. 2

where $s^2 = \left\langle (x_k - \rho)^2 \right\rangle$ is the sample variance and N is the number of measurements ($N>10$). A robust estimator for the sample variance is used: $s = 0.5 \cdot (P_{84} - P_{16})$, where P_{84} and P_{16} are the 84th and 16th percentiles of the distribution, respectively. Due to large number of data available for averaging, the standard error of the mean is usually less than 1% in the stratosphere for the limb instruments.

In addition, the inhomogeneity measures in latitude, H_{lat}, and in time, H_{time} (Sofieva et al., 2014b) are provided with the data. Each inhomogeneity measure H is the linear combination of two

classical inhomogeneity measures, asymmetry A and entropy E (for definition of these parameters, see Sofieva et al., 2014b):

$$H = \tfrac{1}{2}(A + (1 - E)) \hspace{3cm} \text{Eq. 3}$$

The inhomogeneity measure H ranges from 0 to 1 (the more homogeneous, the smaller H). The detailed description and illustrations of uncertainties of monthly zonal mean data can be found in (Sofieva et al., 2017b).

3.1.4 Tropospheric ozone column

3.1.4.1 Limb Nadir Matching (LMN)

The LNM technique is a residual approach that derive tropospheric ozone column (TOC) by subtracting stratospheric O_3 column (SOC), retrieved from the limb observations, from the total O_3 column (TOZ), derived from the nadir observations (Ebojie et al., 2014). This dataset is not included in the Ozone_cci+. The information about the dataset and error characterization can be found in Ozone_cci Phase 2 CECR.

3.1.4.2 Convective Cloud Differential (CCD) Method

The convective cloud differential (CCD) method to retrieve tropospheric ozone columns is applied to level 2 GOME data, i.e., ozone vertical columns and cloud data. This study was performed in Ozone_cci Phase 2 (CERC).

3.2 Merged data sets

3.2.1 Total ozone GTO-ECV

The merged Level 3 monthly gridded (1°x1°) mean total ozone product (GTO-ECV) incorporates measurements from six nadir-viewing satellite sensors: GOME/ERS-2, SCIAMACHY/ENVISAT, OMI/NASA-Aura, GOME-2/MetOp-A, GOME-2/MetOp-B, and TROPOMI/S5-P. Merging is performed on a daily basis using OMI as a reference sensor for inter-sensor calibration. Finally, monthly means are computed. The sample standard deviation and the standard error of the mean are provided. The latter takes into account spatial-temporal sampling errors inherent to the individual satellite data, which were obtained from an Observing System Simulation Experiment (OSSE).

Garane et al. (2018) presented a study about the comparison of the merged GOME-type Total Ozone ECV record with ground-based instruments. The authors report that the long-term drift of the data set w.r.t. ground-based observations is negligible in the Northern Hemisphere with −0.11±0.10 % per decade for Dobson and 0.22±0.08 % per decade for Brewer collocated

measurements. In the Southern Hemisphere the drift with respect to Dobson collocations was found to be 0.23±0.09 % per decade.

3.2.2 Ozone profiles from nadir sensors

This product is under development.

3.2.3 Merged SAGE-CCI-OMPS limb dataset

The description of the merged SAGE II - CCI - OMPS_LP data set can be found in (Sofieva et al., 2017b). The merged SAGE-CCI-OMPS dataset consists of deseasonalized anomalies of ozone in 10° latitude bands from 90S to 90N and from 10 to 50 km in steps of 1 km covering the period from October 1984 to present. In addition, merged monthly zonal mean number density profiles are also included.

The merging is performed via taking the median of deseasonalized anomalies. Each data in the merged SAGE-CCI-OMPS dataset is provided with estimated uncertainty, which is estimated as follows. First, uncertainties of individual deseasonalized normalized anomalies for each month and each latitude-altitude bin $\sigma_{\Delta i}$ are evaluated as.

$$\sigma_{\Delta i} = \frac{1}{\rho_m} \sqrt{\sigma_{\rho,i}^2 + \sigma_m^2} , \qquad \text{Eq. 4}$$

where $\sigma_{\rho,i}$ is the uncertainty of the monthly zonal mean value from Eq. (2), and σ_m is uncertainty of the seasonal cycle (Sofieva et al., 2017b).

The uncertainties of the merged deseasonalized anomalies (which correspond to median value) are estimated as

$$\sigma_{\Delta,merged} = \min\left(\sigma_{\Delta,j_{med}} , \sqrt{\frac{1}{N}\sum_{j=1}^{N}\sigma_{\Delta,j}^2 + \frac{1}{N^2}\sum_{j=1}^{N}\left(\Delta_j - \Delta_{merged}\right)^2} \right), \qquad \text{Eq. 5}$$

where $\sigma_{\Delta,j_{med}}$ is the uncertainty of the anomaly of the instrument corresponding to the median value. Eq. (5) can be interpreted as follows. If individual anomalies are significantly different, i.e., the corresponding error bars do not intersect, the uncertainty of the merged anomaly is the uncertainty corresponding to the median value. In case several instruments report a similar anomaly (intersecting error bars), this provides more confidence of this anomaly value, and the resulting uncertainty of the merged anomaly is approximated as $\sqrt{\frac{1}{N}\sum_{j=1}^{N}\sigma_{\Delta,j}^2 + \frac{1}{N^2}\sum_{j=1}^{N}\left(\Delta_j - \Delta_{merged}\right)^2}$

The uncertainty of the merged dataset is illustrated in Figure 1Figure 25 for several latitude bands. As expected, the uncertainties in the time period when only SAGE II data were available are larger

than uncertainties for time periods when several instruments have contributed. The average uncertainty is usually less than 4% before 2001 and below 1% for the years 2002-2017. In the UTLS, uncertainties are larger than in the stratosphere and are in the range of 3-9 %. At mid-latitudes, uncertainties are larger in winter than in summer due to larger ozone variability during winter; this is observed clearly in the period before 2001.
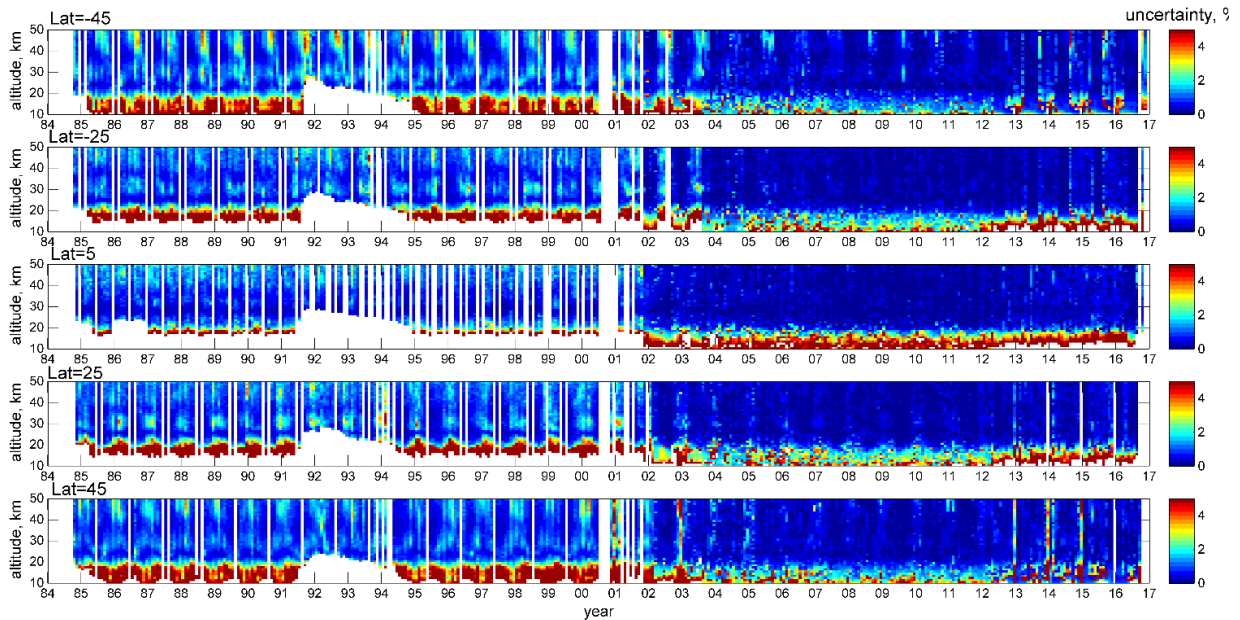
Figure 25: Uncertainties of the merged deseasonalized anomalies in %, Eq. (5), for several 10° latitude bands, centres of which are specified in the legend.

3.2.4 **MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP)**

The merging in LAT-LON merged monthly mean dataset is performed in the same way as for SAGE-CCI-OMPS dataset, thus the uncertainties are evaluated in a similar way. The detailed information about uncertainty estimates, which is analogous to that provided in Sect. 3.2.3 is provided in Sofieva et al., (2021a).

The average estimated uncertainty of the merged ozone is usually less than 2% before 2012 and below 1% after 2012. In the UTLS, uncertainties are larger than in the stratosphere; they are typically in the range of 2-12 % before 2012 and 2-6 % after 2012.

3.2.5 **Multi Sensor Reanalysis (MSR-2) of Total ozone**

The uncertainty of the MSR-2 data set is given in the data and has been evaluated in van der A et al. (2015) for the period after 1970. Below follows the summary.

To evaluate the quality of the MSR-2 data, the observation-minus-forecast (OmF) and the observation-minus-analysis (OmA) statistics have been analysed. The OmA of this dataset is less than 1%, which is better than for the assimilation of observations of a single sensor and is improved

as compared to the MSR-1. 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 percent. This holds also for the period with only sparse BUV observations, although model biases of several percent as a function of latitude become visible. The RMS errors are around 2-3 percent between 1979 and 2012, which is small given that the RMS errors contain contributions from the representativity errors, forecast errors and instrumental noise. For very long time periods without any data (e.g. in 1977), longer than several months, the error becomes more than 20%. These cases may be efficiently excluded from the dataset by filtering with the forecast error estimate provided in the ozone data product, which correctly indicates large model forecast errors during these periods.

It has been shown that the MSR-2 level 2 data show an insignificant drift and SZA and effective ozone temperature dependence as compared to the ground-based observations. The fitted offset, trend and seasonality in the comparison between the MSR-2 level 4 ozone fields and the average of the ground-based observations were negligible. The maximum fitted offset is 0.2 DU.

3.3 Ozone nadir profiles Level 4 data

This dataset is not included in the Ozone_cci+. The information about the dataset and error characterization can be found in Ozone_cci Phase 2 CECR.

4 Overview of the methods for uncertainty validation

4.1 Validation of random uncertainties

In remote-sensing measurements, random uncertainty is usually estimated via propagation of instrumental noise through the inversion algorithm. These estimates can be imperfect due to different approximation used in retrievals. They are sometimes referred to as "ex-ante" errors (von Clarmann, 2006). Below we present an overview of the methods for validation of random error component. Some of the methods were discussed in (Sofieva et al., 2014b).

In the laboratory, the experimental random uncertainty estimates can be obtained using repeated measurements under the same conditions: the sample variance $s^2 = \mathrm{var}(x)$ approaches the variance of random error distribution σ^2 when the size of sample N tends to infinity. The sample variance has a χ^2 distribution with $N-1$ degrees of freedom. It can be approximated for large N by a Gaussian distribution with variance

$$\mathrm{var}(s^2) \approx \sigma^4 \frac{2}{N}, \qquad\qquad Eq.\ 6$$

giving the uncertainty of the experimentally estimated random uncertainty.

Contrary to laboratory experiments, geophysical observation conditions cannot be kept exactly constant for atmospheric measurements. Therefore, the sample variance contains a contribution due to the natural variability σ_{nat}^2: $s^2 = \sigma^2 + \sigma_{nat}^2$. For validation of uncertainty estimates, σ_{nat}^2 should be minimized by selecting collocated measurements or it should be known from independent sources (for example, from a chemistry-transport model, CTM). In this section, we overview the methods for random uncertainty validation based on measurements. The use of CTMs in random uncertainty validation is discussed in Section 4.3.

Approaches to validation of random error estimates usually rely on the variance of the difference s_{12}^2 in a set of collocated measurements x_1 and x_2:

$$s_{12}^2 = \left\langle \left(x_1 - x_2 \right)^2 \right\rangle - \left\langle x_1 - x_2 \right\rangle^2 = \sigma_{0,nat}^2 + \sigma_1^2 + \sigma_2^2 \qquad\qquad Eq.\ 7$$

In **Error! Reference source not found.**, $\sigma_{0,nat}^2$ stands for the natural variability within a space-time collocation window (note that $\sigma_{0,nat}^2$ is different from σ_{nat}^2), and the angular brackets denote the mean.

The methods for random uncertainty validation can be divided into 2 groups depending on what kind of data are used: (1) from the same instrument and (2) from different instruments.

4.1.1 **Using collocated measurements from the same instrument**

For perfectly collocated measurements ($\sigma_{0,nat}^2 \approx 0$) from the same instrument with the same precisions $\sigma_1 = \sigma_2 = \sigma$, Eq. (7) is reduced to $s_{12}^2 \approx 2\sigma^2$, thus allowing validation of the uncertainty estimate σ. In this estimate, measurements x_1 and x_2 are assumed to be independent. This uncertainty validation method was realized, for example, for closely collocated MIPAS and OSIRIS measurements (Bourassa et al., 2012; Piccolo and Dudhia, 2007). The uncertainty of this experimental precision estimate $\hat{\sigma}^2 = s_{12}^2/2$ is defined by the uncertainty of sample variance s_{12}^2.

There are several limitations associated with this method. First, natural variability $\sigma_{0,nat}^2$ is not necessarily small, even in a tight spatio-temporal window. In such cases s_{12}^2 will be larger than a combined uncertainty $2\sigma^2$, thus the estimate $\hat{\sigma}^2 = s_{12}^2/2$ will be overestimated. Second, number of self-collocated measurements for limb satellites is limited (self-collocated measurements are usually around the Poles, where variability is extremely high in winter season).

Provided many collocated measurements from the same instrument are available (self-collocations), the precision of the dataset can be estimated also by computing a structure function, or the rms difference as a function of increasing separation in time and in space:

$$D(\boldsymbol{\rho}) = D(\mathbf{r}_1 - \mathbf{r}_2) = \frac{1}{2}\left\langle \left[f(\mathbf{r}_1) - f(\mathbf{r}_2) \right]^2 \right\rangle \qquad\qquad Eq.\ 8$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are two locations and $\boldsymbol{\rho} = \mathbf{r}_1 - \mathbf{r}_2$. When using experimental (noisy) data for evaluation of variogram/structure function, the difference of an atmospheric parameter in two locations is defined not only by the natural variability of this atmospheric parameter, but also by uncertainty of measurements. Therefore, with the spatio-temporal separation $\rho \to 0$, $D(\boldsymbol{\rho})$ tends to the random uncertainty variance σ^2_{noise} (the offset at zero is called "nugget" in geostatistics). Figure 26 illustrates the structure function method, which is discussed in details in (Sofieva et al., 2021b) and applied to TROPOMI total ozone measurements.
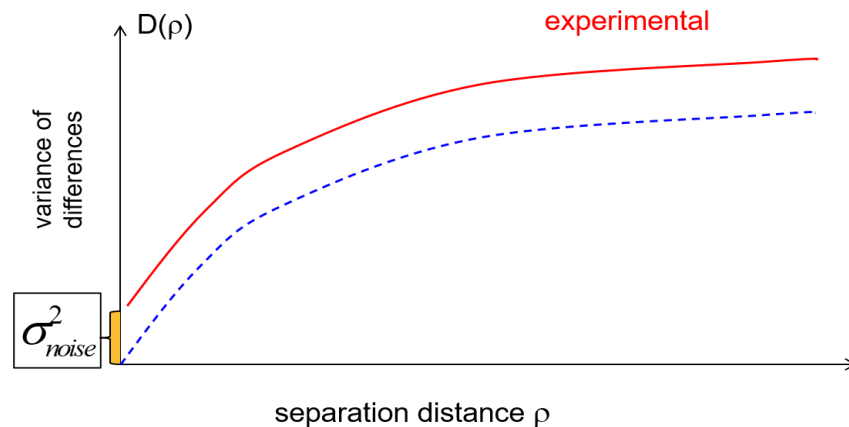

Figure 26. From (Sofieva et al., 2021b): The schematic representation of the structure function estimated from noisy measurements.

An analogous method - evaluation of the one-dimensional structure function in polar regions (with transformation of temporal mismatch to spatial separation) - has been applied for validation of random uncertainty estimates of the MIPAS and GOMOS ozone profiles (Laeng et al., 2015; Sofieva et al., 2014a).

4.1.2 **Using measurements from different instruments**

Fioletov et al., (2006) have proposed estimating simultaneously the measurement precision and natural variability from sample variances of two perfectly collocated datasets and the variance of their difference. Since the precision estimates by the Fioletov method are linear combinations of three sample variances, they can have large uncertainty if one of the sample variances is large and/or the number of collocated measurements is limited. The natural variability within the space-time collocation window is small but not zero. This results in additional difficulties in the application of this method, as observed by Bourassa et al., (2012).

Sofieva et al., (2014a) applied comparison of natural variability from GOMOS different stars for validation of random uncertainties. This method allows detection of problems with GOMOS uncertainty estimates for dim stars. The same method, but with using datasets from different instruments, has been applied in Sofieva et al., (2014a) for GOMOS and MIPAS, and for different pairs of limb ozone datasets in CCI Phase 2 CECR.

4.2 Validation of systematic uncertainties

The systematic uncertainties in the data are investigated via comparison with the reference instruments. The biases between the datasets give the a posteriori estimates of systematic uncertainties.

4.3 Using CTM simulations to estimate co-location mismatch uncertainty

4.3.1 **Observing System Simulation Experiment (OSSE)**

Methods based on self-colocations allow us to quantify the uncertainty due to random errors in the satellite data set. However, this technique is blind to any potential systematic error and therefore cannot be used to assess the full measurement uncertainty. For this reason, validation with independent reference measurements remains crucial. However, as illustrated by Eq. (7), the differences between satellite and reference measurements cannot be confronted directly with the reported measurement uncertainties. For all reasonable co-location criteria, i.e. those resulting in a sufficiently large number of comparison pairs, one must also take into account the additional differences due to co-location mismatch, i.e. differences in spatio-temporal sampling and smoothing of the variable and inhomogeneous ozone field.

A possible approach to quantify these additional uncertainty terms, e.g. $\sigma^2_{0,\text{var}}$ in Eq. (7) when looking at the spread on the differences, is by performing an Observing System Simulation Experiment (OSSE). This method consists in (1) the creation of appropriate observation operators, which quantify the actual 4-D extent of measurement sensitivity of each measurement, (2) application of these observation operators on high-resolution global gridded fields, e.g. from an ozone reanalysis, to simulate the individual measurements, and (3) quantifying the differences between these simulated measurements in an exact copy of the actual validation exercise. When no measurement errors are included in the simulated measurements, this allows an estimate of the differences due solely to co-location mismatch, from which for instance $\sigma^2_{0,\text{var}}$ in Eq. (7) can be derived.

This approach was explored for the validation of Ozone_cci total ozone column products by (Verhoelst et al., 2015), and their Figure 12 is reproduced here as Figure 27.
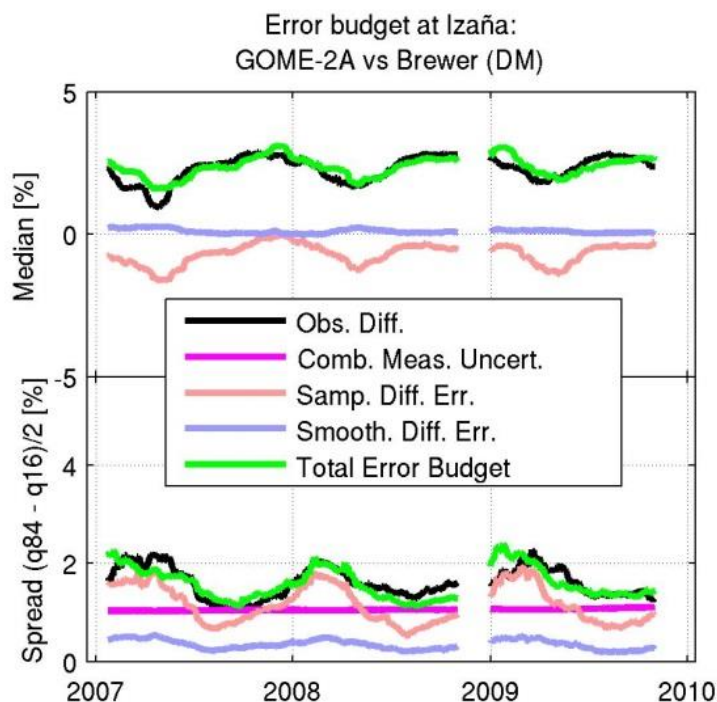
Figure 27: Error budget of comparisons between GODFITv3 GOME-2/MetOp-A total ozone columns and co-located Brewer (daily mean) measurements from the NDACC station at Izana, Canary Islands. The upper panel contains the 3-month median on the differences, the bottom panel the spread (as an interpercentile). Black lines show the statistics of the observed differences, the coloured lines the different components of the OSSE. While the combined measurement uncertainty (magenta) does not account for the observed comparison spread, the error budget can be closed by including the errors due to co-location mismatch. Note that the large median error of approx. 3% is due to the station's mountain top location, due to which the Brewer misses part of the column seen by the larger satellite pixel measuring down to sea level.

4.3.2 **A posteriori assessment of random errors**

If the natural variability is known from an external source, a posteriori (ex-post in von Clarmann et al., (2020) terminology) uncertainties can be estimated as $\sigma^2_{ex-post} = s^2 - \sigma^2_{nat}$, where s^2 is the sample variance in a set of measurements and σ^2_{nat} is the estimated of the natural variability. This approach has been successfully applied for data homogenization of ozone profiles from limb satellite measurements in (Sofieva et al., 2021b: Synergy of Using Nadir and Limb Instruments for Tropospheric Ozone Monitoring, submitted to AMTD)

4.4 A combined validation of measurement and ex-ante uncertainty

While techniques such as self co-location allow the quantification of random uncertainty in the satellite measurements, they do not test whether the measurement is consistent with the "truth" (or a proxy thereof in the form of a reference measurements) to within the claimed ex-ante uncertainty. In the CCI Product Validation and Intercomparison Report, a consistency test is introduced, quantifying the agreement between satellite and ground-based reference measurements w.r.t to their reported ex-ante uncertainties. The mean difference between satellite and co-located reference measurements (i.e. the bias) is compared to the combined systematic uncertainty, and the non-systematic part is assessed with a reduced Chi Square, χ_r^2, defined as:

$$\chi^2{}_r = \frac{1}{N-1} \Sigma \frac{(SAT - GND - mean(SAT - GND))^2}{\sigma_{SAT}{}^2 + \sigma_{GND}{}^2}$$

where the satellite and reference uncertainties in the denominator exclude the systematic uncertainties. For properly estimated Gaussian uncertainties in both data sets, this estimator has an expectation value of 1. Its use is illustrated for a particular case study in **Error! Reference source not found.**. The χ_r^2 of 1.3 indicates excellent consistency within the reported uncertainties. Note that in some cases, the denominator must also include a variance related to co-location mismatch, e.g. derived using the methods described in 4.3.
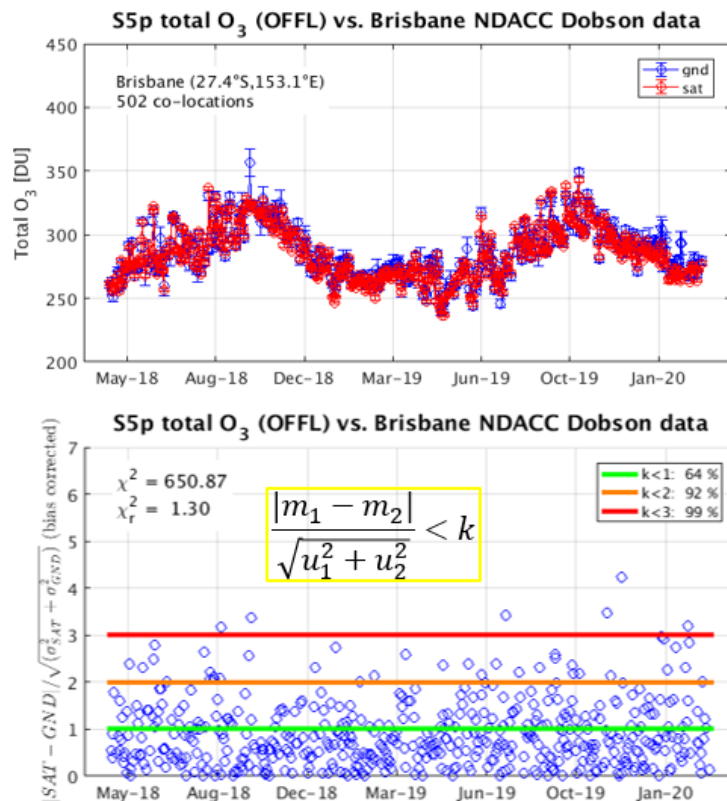

Figure 28: Illustration of a consistency test between S5P-TROPOMI and Dobson total ozone columns above Brisbane, Australia. Top panel: time series of measurements, both satellite and ground-based, with their ex-ante uncertainties. Bottom panel: Time series of differences,

normalized by the combined (random) measurement uncertainties, and the derived metrics for consistency (e.g. the reduced Chi Square). The distribution of the normalized differences w.r.t. the different coverage factors "k" is very close to the numbers expected for a Gaussian error distribution.

5 References

Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarmann, T., & Burrows, J. P. (2018). Retrieval of ozone profiles from OMPS limb scattering observations. *Atmospheric Measurement Techniques, 11*(4), 2135-2149.

Boone, C., Nassar, R., Walker, K., Rochon, Y., McLeod, S., Rinsland, C., & Bernath, P. (2005). Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer. *Applied Optics, 44*(33).

Bourassa, A. E., Roth, C. Z., Zawada, D. J., Rieger, L. A., McLinden, C. A., & Degenstein, D. A. (2018). Drift-corrected Odin-OSIRIS ozone product: algorithm and updated stratospheric ozone trends. *Atmospheric Measurement Techniques, 11*(1), 489-498.

Bourassa, A., McLinden, C., Bathgate, A., Elash, B., & Degenstein, D. (2012). Precision estimate for Odin-OSIRIS limb scatter retrievals. *Journal of Geophysical Research Atmospheres, 117*(4).

Boynard, A., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Turquety, S., George, M., . . . Meyer-Arnek, J. (2009). Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations. *Atmospheric chemistry and physics, 9*(16), 6255-6271.

Boynard, A., Hurtmans, D., Garane, K., Goutail, F., adji-Lazaro, J., Koukouli, M. E., . . . Pommereau, J.-P. (2018). Validation of the IASI FORLI/EUMETSAT ozone products using satellite (GOME-2), ground-based (Brewer--Dobson, SAOZ, FTIR) and ozonesonde measurements. *Atmospheric Measurement Techniques, 11*(9), 5125--5152.

Brühl, C., Drayson, S. R., Russell III, J. M., Crutzen, P. J., McInerney, J. M., Purcell, P. N., . . . McDermid, I. S. (1996). Halogen Occultation Experiment ozone channel validation. *Journal of Geophysical Research: Atmospheres, 101*(D6), 10217-10240.

Chinaud, J. L. (2019). *IASI-C L1 Cal/Val System performance synthesis.*

Coldewey-Egbers, M., G. Loyola, D., Labow, G., & M. Frith, S. (2020). Comparison of GTO-ECV and adjusted MERRA-2 total ozone columns from the last 2 decades and assessment of interannual variability. *Atmospheric Measurement Techniques, 13*(3).

Coldewey-Egbers, M., Loyola, D. G., Koukouli, M., Balis, D., Lambert, J. C., Verhoelst, T., . . . Zehner, C. (2015). The GOME-type Total Ozone Essential Climate Variable (GTO-ECV)

data record from the ESA Climate Change Initiative. *Atmospheric Measurement Techniques, 8*(9).

Damadeo, R., Zawodny, J., Thomason, L., & Iyer, N. (2013). SAGE version 7.0 algorithm: Application to SAGE II. *Atmospheric Measurement Techniques, 6*(12).

Dufour, G., Eremenko, M., Griesfeller, A., Barret, B., Leflochmoën, E., Clerbaux, C., . . . Hurtmans, D. (2012). Validation of three different scientific ozone products retrieved from IASI spectra using ozonesondes. *Atmospheric Measurement Techniques, 5*(3).

Dupuy, E., Walker, K., Kar, J., Boone, C., McElroy, C., Bernath, P., . . . Zawodny, J. (2009). Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE). *Atmospheric Chemistry and Physics, 9*(2).

Ebojie, F., Von Savigny, C., Ladstätter-Weißenmayer, A., Rozanov, A., Weber, M., Eichmann, K., . . . Burrows, J. (2014). Tropospheric column amount of ozone retrieved from SCIAMACHY limb-nadir-matching observations. *Atmospheric Measurement Techniques, 7*(7).

Fioletov, V., Tarasick, D., & Petropavlovskikh, I. (2006). Estimating ozone variability and instrument uncertainties from SBUV(/2), ozonesonde, Umkehr, and SAGE II measurements: Short-term variations. *Journal of Geophysical Research Atmospheres, 111*(2).

Froidevaux, L., Jiang, Y., Lambert, A., Livesey, N., Read, W., Waters, J., . . . Wagner, P. (2008). Validation of Aura Microwave Limb Sounder stratospheric ozone measurements. *Journal of Geophysical Research, 113*(D15).

Garane, K., Lerot, C., Coldewey-Egbers, M., Verhoelst, T., Elissavet Koukouli, M., Zyrichidou, I., . . . Zehner, C. (2018). Quality assessment of the Ozone-cci Climate Research Data Package (release 2017)-Part 1: Ground-based validation of total ozone column data products. *Atmospheric Measurement Techniques, 11*(3).

Hubert, D., Lambert, J., Verhoelst, T., Granville, J., Keppens, A., Baray, J., . . . Zawodny, J. (2016). Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records. *Atmospheric Measurement Techniques, 9*(6).

Hurtmans, D., Coheur, P., Wespes, C., Clarisse, L., Scharf, O., Clerbaux, C., . . . Turquety, S. (2012). FORLI radiative transfer and retrieval code for IASI. *Journal of Quantitative Spectroscopy and Radiative Transfer, 113*(11).

Keppens, A., Lambert, J.-C., Granville, J., Hubert, D., Verhoelst, T., Compernolle, S., . . . Boynard, A. (2018). Quality assessment of the Ozone_cci Climate Research Data Package (release 2017)--Part 2: Ground-based validation of nadir ozone profile data products. *Atmospheric Measurement Techniques, 11*(6), 3769--3800.

Kerridge, B., Siddans, R., Latter, B., Burrows, J., Weber, M., De Beek, R., . . . Hartman, W. (2002). *GOME-2 error assessment study.* Final Report EUMETSAT Contract No EUM/CO/01/901/DK.

Laeng, A., Hubert, D., Verhoelst, T., von Clarmann, T., Dinelli, B., Dudhia, A., . . . Zehner, C. (2015). The ozone climate change initiative: Comparison of four Level-2 processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS). *Remote Sensing of Environment, 162*.

Lerot, C., van Roozendael, M., Spurr, R., Loyola, D., Coldewey-Egbers, M., Kochenova, S., . . . Zehner, C. (2014). Homogenized total ozone data records from the european sensors

gome/ers-2, SCIAMACHY/envisat, and GOME-2/MetOp-A. *Journal of Geophysical Research, 119*(3).

Livesey, N. J., Read, W. G., Wagner, P. F., & Lambert, A. (2020). *Version 4.2x Level 2 and 3 data quality and description document.* available at: https://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf.

Livesey, N., Filipiak, M., Froidevaux, L., Read, W., Lambert, A., Santee, M., . . . Webster, C. (2008). Validation of Aura Microwave Limb Sounder O 3 and CO observations in the upper troposphere and lower stratosphere. *Journal of Geophysical Research, 113*(D15).

Lumpe, J., Bevilacqua, R., Hoppel, K., & Randall, C. (2002). POAM III retrieval algorithm and error analysis. *Journal of Geophysical Research Atmospheres, 107*(21).

McCormick, M., Lei, L., Hill, M., Anderson, J., Querel, R., & Steinbrecht, W. (2020). Early results and validation of SAGE III-ISS ozone profile measurements from onboard the International Space Station. *Atmospheric Measurement Techniques, 13*(3).

Miles, G., Siddans, R., Kerridge, B., Latter, B., & Richards, N. (2015). Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation. *Atmospheric Measurement Techniques, 8*(1).

Piccolo, C., & Dudhia, A. (2007). Precision validation of MIPAS-Envisat products. *Atmospheric Chemistry and Physics, 7*(8).

Rahpoe, N., Von Savigny, C., Weber, M., Rozanov, A., Bovensmann, H., & Burrows, J. (2013). Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model. *Atmospheric Measurement Techniques, 6*(10).

Rault, D. (2005). Ozone profile retrieval from Stratospheric Aerosol and Gas Experiment (SAGE III) limb scatter measurements. *Journal of Geophysical Research D: Atmospheres, 110*(9).

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice.* World scientific.

Rong, P., Russell, J., Mlynczak, M., Remsberg, E., Marshall, B., Gordley, L., & Lopez-Puertas, M. (2009). Validation of thermosphere ionosphere mesosphere energetics and dynamics/sounding of the atmosphere using broadband emission radiometry (TIMED/SABER) vl.07 ozone at 9.6 µ in altitude range 15-70 km. *Journal of Geophysical Research Atmospheres, 114*(4).

Sheese, P. E., Walker, K. A., Boone, C. D., Bourassa, A. E., Degenstein, D. A., Froidevaux, L., . . . Zou, J. (2021). Assessment of the quality of ACE-FTS stratospheric ozone data. *Atmospheric Measurement Techniques Discussions*, 1-27.

Siddans, R. (2003). *Height Resolved Ozone Retrievals from Global Ozone Monitoring.* University of Reading.

Siddans, R., Kerridge, B., Reburn, W., & Munro, R. (1998). Height-resolved ozone retrievals in the troposphere and lower stratosphere from GOME. *Earth Observation Quarterly*(58).

Sofieva, V. F., Ialongo, I., Hakkarainen, J., Kyrölä, E., Tamminen, J., Laine, M., . . . Bertaux, J.-L. (2017a). Improved GOMOS/Envisat ozone retrievals in the upper troposphere and the lower stratosphere. *Atmospheric Measurement Techniques, 10*(1), 231-246.

Sofieva, V. F., Kalakoski, N., Päivärinta, S. M., Tamminen, J., Laine, M., & Froidevaux, L. (2014b). On sampling uncertainty of satellite ozone profile measurements. *Atmospheric Measurement Techniques, 7*(6).

Sofieva, V. F., Kyrölä, E., Laine, M., Tamminen, J., Degenstein, D., Bourassa, A., . . . Bhartia, P. K. (2017b). Merged SAGE II, Ozone-cci and OMPS ozone profile dataset and evaluation of ozone trends in the stratosphere. *Atmospheric Chemistry and Physics, 17*(20).

Sofieva, V. F., Shing Lee, H., Tamminen, J., Lerot, C., Romahn, F., & Loyola, D. G. (2021b). A method for random uncertainties validation and probing the natural variability with application to TROPOMI on board Sentinel-5P total ozone measurements. *Atmospheric Measurement Techniques, 14*(4).

Sofieva, V. F., Szelag, M., Tamminen, J., Kyrölä, E., Degenstein, D., Roth, C., . . . Retscher, C. (2021a). Measurement report: Regional trends of stratospheric ozone evaluated using the MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP). *Atmospheric Chemistry and Physics, 21*(9).

Sofieva, V. F., Tamminen, J., Kyrölä, E., Laeng, A., Von Clarmann, T., Dalaudier, F., . . . Vanhellemont, F. (2014a). Validation of GOMOS ozone precision estimates in the stratosphere. *Atmospheric Measurement Techniques, 7*(7).

Steck, T., Clarmann, T. v., Fischer, H., Funke, B., Glatthor, N., Grabowski, U., . . . Linden, A. (2007). Bias determination and precision validation of ozone profiles from MIPAS-Envisat retrieved with the IMK-IAA processor. *Atmospheric chemistry and physics, 7*(13), 3639--3662.

Tamminen, J., Kyrölä, E., Sofieva, V., Laine, M., Bertaux, J., Hauchecorne, A., . . . Fraisse, R. (2010). GOMOS data characterisation and error estimation. *Atmospheric Chemistry and Physics, 10*(19).

Urban, J., Lautie, N., Le Flochmoen, E., Jimenez, C., Eriksson, P., de La Noe, J., . . . Frisk, U. (2005). Odin/SMR limb observations of stratospheric trace gases: Level 2 processing of ClO, N2O, HNO3, and O3. *Journal of Geophysical Research: Atmospheres, 110*(D14).

Valks, P., Hao, N., Garcia, S., Loyola, D., Dameris, M., Jöckel, P., & Delcloo, A. (2014). Tropical tropospheric ozone column retrieval for GOME-2. *Atmospheric Measurement Techniques, 7*(8).

Van Der A, R. a. (2015). Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012. *Atmospheric Measurement Techniques, 8*(7), 3021-3035.

Verhoelst, T., Granville, J., Hendrick, F., Kohler, U., Lerot, C., Pommereau, J., . . . Lambert, J. (2015). Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons. *Atmospheric Measurement Techniques, 8*(12).

Von Clarmann, T. (2006). Validation of remotely sensed profiles of atmospheric state variables: Strategies and terminology. *Atmospheric Chemistry and Physics, 6*(12).

Von Clarmann, T. (2014). Smoothing error pitfalls. *Atmospheric Measurement Techniques, 7*(9).

Von Clarmann, T., Degenstein, D., Livesey, N., Bender, S., Braverman, A., Butz, A., . . . Zawada, D. (2020). Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature. *Atmospheric Measurement Techniques, 13*(8).

Wespes, C., Hurtmans, D., K Emmons, L., Safieddine, S., Clerbaux, C., Edwards, D. P., & Coheur, P. F. (2016). Ozone variability in the troposphere and the stratosphere from the first 6 years of IASI observations (2008-2013). *Atmospheric Chemistry and Physics, 16*(9).

Zawada, D. J., Rieger, L. A., Bourassa, A. E., & Degenstein, D. A. (2018). Tomographic retrievals of ozone with the OMPS Limb Profiler: algorithm description and preliminary results. *Atmospheric Measurement Techniques, 11*(4), 2375-2393.