# Ozone_cci+

# Algorithm Theoretical Basis Document (ATBD)

Date: 24/01/2023

Version: 2.1

**WP Manager:** R. Siddans

**WP Manager Organization:** STFC Rutherford Appleton Laboratory

**Other partners:** DLR-IMF, KNMI, RAL, ULB, UBR, FMI

# DOCUMENT PROPERTIES

| | |
|---|---|
| Title | Algorithm Theoretical Basis Document (ATBD) |
| Reference | Ozone_cci+_D2p1_ATBD_2.1 |
| Issue | 2 |
| Revision | 1 |
| Status | Draft |
| Date of issue | 24 January 2023 |
| Document type | D2.1 |

| | FUNCTION | NAME | DATE | SIGNATURE |
|---|---|---|---|---|
| LEAD AUTHOR | Project partner | Richard Siddans | | |
| CONTRIBUTING AUTHORS | Project partner | M. Coldewey-Egbers, C. Lerot, V. Sofieva, R. van der A, C. Wespes, C. Arioso | | |
| REVIEWED BY | Science leader | Daan Hubert | 21/12/2022 | |
| APPROVED BY | ESA | Christian Retscher | 20/01/2023 | |
| ISSUED BY | Project partner | Richard Siddans | 24/01/2023 | |

# DOCUMENT CHANGE RECORD

| Issue | Revision | Date | Modified items | Observations |
|---|---|---|---|---|
| 0 | 0 | 14/02/2020 | Initial template | Creation of document |
| 1 | 0 | 7/05/2020 | First issued version | |
| 1 | 1 | 8/3/2021 | Modifications in response to ESA comments on version 1.0:<br>Fig 3 (now 2.3) update (and now shows to TROPOMI - GOME2C, consistent with fig 2.2)<br>p23: Reference to the chorine activation term added.<br>Ch 3.2.7: Averaging kernel quality check clarified.<br>p32: Clarified text reference to x vs r.<br>p33: Clarification has been added regarding use of x and x-hat.<br>Ch 4.2.3 New section added.<br>Editorial:<br>Table of contents alignment fixed<br>MLS acronym added<br>Some minor typos fixed<br>Rationalised numbering of equations figures and tables throughout. | |
| 2 | 0 | 5/11/2021 | Updates to Sects. 2.1, 3.2, 4.2, 4.4. | |
| 2 | 1 | 24/01/2023 | Inclusion of GOP-ECV algorithm (Sect. 3.2) | |

## Executive Summary

The Algorithm Theoretical Basis Document (ATBD, deliverable D2.1 in Ozone_cci+) describes the algorithms used for total ozone columns, nadir-based ozone profiles, and limb-based ozone profiles.

A series of new algorithms will be developed with a focus on both level-2 and level-3 data products.

For total ozone, the following activities will be undertaken:

- Integration of Sentinel-5P and MetOp-C/GOME-2 in the GTO-ECV data record, based on the reference GODFIT level-2 algorithm
- Merging of GTO-ECV data record with historical NASA/SBUV/TOMS data records to create a 40-year CDR starting in 1979
- Tentative expansion of the MSR data record back to 1960 using ground-based total ozone measurements

For ozone profiles from nadir sensors:

- Integration of Sentinel-5P and MetOp-C/GOME-2 in CCI data sets based on the reference RAL level-2 retrieval algorithm
- Creation of a first European merged ozone profile and total ozone data record (GOP-ECV) starting with ERS-2/GOME in 1995 and consistent with the GTO-ECV total ozone data record
- Improvement of IASI level-2 retrieval algorithm (FORLI) to reduce uncertainties and systematic biases in the UTLS region
- Exploration of the potential to reduce uncertainties in nadir profile level-2 products in the lower troposphere and UTLS through exploitation of the synergy between UV, VIS and TIR retrievals

For ozone profiles from limb sensors:

- Integration of several new sensors (POAM-III on SPOT-4, SAGE-III on Meteor-3M, and SAGE-III on ISS) in the HARMOZ data record
- Intercomparison of different competing OMPS-LP level-2 retrieval algorithms. This activity might lead to the identification of a need for a dedicated round-robin exercise.
- Improvement of OMPS-LP level-2 algorithm in the UTLS and in polar regions
- Extension of the merged SAGE-II/CCI/OMPS long-term zonal mean data record with additional sensors to improve its accuracy, especially in lower stratosphere
- Improvement of latitude-longitude gridded level-3 data products based on limb-type sensors

# Table of Contents

# 1 Purpose and scope

## 1.1 Purpose

This document describes the algorithms used in the Ozone_cci+, addressing total ozone columns, nadir-based ozone profiles, and limb-based ozone profiles.

## 1.2 Scope

The scope of the ATBD is to document algorithms used in the CCI+ project, with particular emphasis on algorithm features which are new in CCI+. Where algorithms are already described by existing ATBDs or similar documents, references are made to these documents and only a brief overview of the algorithmic basis is given.

## 1.3 Applicable documents

[AD-1] Data Standards Requirements for CCI Data Producers. Latest version at time of writing is v1.2: ref. CCI-PRGM-EOPS-TN-13-0009, 9 March 2015, available online at: http://cci.esa.int/sites/default/files/CCI_Data_Requirements_Iss1.2_Mar2015.pdf

[AD-2] CCI Data Policy v1.1. Available online at: https://earth.esa.int/documents/10174/1754357/RD-7_CCI_Data_Policy_v1.1.pdf

## 1.4 References documents

[RD-1] GCOS Climate Monitoring Principles, November 1999, reproduced in [RD-5], p48.

[RD-2] Guideline for the Generation of Satellite-based Datasets and Products meeting GCOS Requirements, GCOS Secretariat, GCOS-128, March 2009 (WMO/TD No. 1488). Available online at: http://library.wmo.int

[RD-3] Quality assurance framework for earth observation (QA4EO): http://qa4eo.org

[RD-4] EU Research Programmes on Space and Climate: Horizon 2020 (H2020), (http://ec.europa.eu/programmes/horizon2020/en/h2020-section/space, https://ec.europa.eu/programmes/horizon2020/en/h2020-section/climate-action-environment-resource-efficiency-and-raw-materials) and Copernicus (http://www.copernicus.eu/).

[RD-5] The Global Observing System for Climate: Implementation Needs, GCOS-200, October 2016. Available online at: http://library.wmo.int

[RD-6] Status of the Global Observing System for Climate, GCOS-195, October 2015. Available online at: http://library.wmo.int

[RD-7] Joint Committee for Guides in Metrology, 2008, Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM), JGCM 100: 2008. Available online at http://www.bipm.org/en/publications/guides/gum.html.

[RD-8] Copernicus Space Component: http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Space_Component

[RD-9] User requirements for monitoring the evolution of stratospheric ozone at high vertical resolution (Operoz), 2015, ESA Expro contract 4000112948/14/NL/JK. Available online at: http://projects.knmi.nl/capacity/Operoz/Operoz_final_report_with_exec_summary_1mar2015.pdf

[RD-10] K. Bogumil, J. Orphal, T. Homann, S. Voigt, P. Spietz, O.C. Fleischmann, A. Vogel, M. Hartmann, H., Bovensmann, J. Frerick, and J.P. Burrows, "Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: Instrument characterization and reference data for atmospheric remote sensing in the 230-2380 nm region", J. Photochem. Photobiol. A: Chem. 157, 167-184 (2003); DOI: 10.1016/S1010-6030(03)00062-5

[RD-11] Boynard, A., et al.: Validation of the IASI FORLI/EUMETSAT O3 products using satellite (GOME-2), ground-based (Brewer-Dobson, SAOZ, FTIR) and ozonesonde measurements, Atmos. Meas. Tech.,https://doi.org/10.5194/amt-11-5125-2018 , 2018.

[RD-12] Cariolle, D. and Teyssèdre, H.: A revised linear ozone photochemistry parametrization for use in transport and general circulation models: Multi-annual simulations, Atmos. Chem. Phys. Discuss., 7, 1655–1697, 2007.

[RD-13] Chance, K., Burrows, J et al., Satellite measurements of atmospheric ozone profiles, including tropospheric ozone, from ultraviolet/visible measurements in the nadir geometry: a potential method to retrieve tropospheric ozone, J. Quant. Spectrosc. Radiat. Transfer Vol. 51. No. 4, pp. 461476, 1997

[RD-14] Chen, Z., DeLand, M., and Bhartia, P. K.: A new algorithm for detecting cloud height using OMPS/LP measurements, Atmos. Meas. Tech., 9, 1239–1246, https://doi.org/10.5194/amt-9-1239-2016, 2016.

[RD-15] Chinaud, J., et al.: IASI-C L1 Cal/Val System performance synthesis. EUMETSAT Technical Note IA-NT-0000-4477-CNES, https://www.EUMETSAT.int/website/home/News/DAT_4439637.html, 2019.

[RD-16] Coldewey-Egbers, M., Loyola, D. G., Koukouli, M., Balis, D., Lambert, J.-C., Verhoelst, T., Granville, J., van Roozendael, M., 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, Atmos. Meas. Tech., 8, 3923–3940, https://doi.org/10.5194/amt-8-3923-2015, 2015.

[RD-17] Coldewey-Egbers, M., Loyola, D., Labow, G., and Frith, S.: Comparison of GTO-ECV and Adjusted-MERRA total ozone columns from the last two decades and assessment of interannual variability, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-297, in review, 2019.

[RD-18] Eskes, H. J., van Velthoven, P. F. J., Valks, P. J. M. and Kelder, H. M.: Assimilation of GOME total ozone satellite observations in a three-dimensional tracer transport model, Q. J. R. Meteorol. Soc., 129, 1663-1681, 2003.

[RD-19] Garane, K., Lerot, C., Coldewey-Egbers, M., Verhoelst, T., Koukouli, M.E., Zyrichidou, I., Balis, D.S., Danckaert, T., Goutail, F., Granville, J., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., Pommereau, J.-P., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 1: Ground-based validation of total ozone column data products, Atmos. Meas. Tech., 11, 1385-1402, doi:10.5194/amt-11-1385-2018, 2018.

[RD-20] Garane, K., Koukouli, M.-E., Verhoelst, T., Lerot, C., Heue, K.-P., Fioletov, V., Balis, D., Bais, A., Bazureau, A., Dehn, A., Goutail, F., Granville, J., Griffin, D., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., McLinden, C., Pazmino, A., Pommereau, J.-P., Redondas, A., Romahn, F., Valks, P., Van Roozendael, M., Xu, J., Zehner, C., Zerefos, C., and Zimmer, W.: TROPOMI/S5P total ozone column data: global ground-based validation and consistency with other satellite missions, Atmos. Meas. Tech., 12, 5263–5287, https://doi.org/10.5194/amt-12-5263-2019, 2019.

[RD-21] Hermans, C.: O4 absorption cross-sections at 296K (335.59–666.63 nm), online available from: http://spectrolab.aeronomie.be/index.htm (last accessed 15 January 2011), 2011.

[RD-22] Hollmann, R., et al., The ESA climate change initiative: Satellite data records for essential climate variables. American Meteorological Society. Bulletin, Vol. 94, No. 10, 2013, p. 1541-1552.

[RD-23] Hurtmans, D,. et al.: FORLI radiative transfer and retrieval code for IASI, Journal of Quantitative Spectroscopy and Radiative Transfer, https://doi.org/10.1016/j.jqsrt.2012.02.03, 2012.

[RD-24] Keppens, A., J.-C. Lambert, J. Granville, D. Hubert, T. Verhoelst, S. Compernolle, B. Latter, B. Kerridge, R. Siddans, A. Boynard, J. Hadji-Lazaro, C. Clerbaux, C. Wespes, D. R. Hurtmans, P. F. Coheur, J. van Peet, R. van der A, K. Garane, M. E. Koukouli, D. S. Balis, A. Delcloo, R. Kivi, R. Stübi, S. Godin-Beekmann, M. Van Roozendael, C. Zehner: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 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.

[RD-25] Kerr, J. B.: New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct sun spectra, J. Geophys. Res., 107(D23), 4731, doi:10.1029/2001JD001227, 2002.

[RD-26] Koukouli, M. E., Lerot, C., Granville, J., Goutail, F., Lambert, J.-C., Pommereau, J.-P., Balis, D., Zyrichidou, I., Van Roozendael, M., Coldewey-Egbers, M., Loyola, D., Labow, G., Frith, S., Spurr, R. and Zehner, C.: Evaluating a new homogeneous total ozone climate data record from OME/ERS-2, SCIAMACHY/Envisat, and GOME-2/Metop-A, J. Geophys. Res. Atmos., 120(23), 12,296-2,312, doi:10.1002/2015JD023699, 2015.

[RD-27] Kramarova, N. A., Bhartia, P. K., Jaross, G., Moy, L., Xu, P., Chen, Z., DeLand, M., Froidevaux, L., Livesey, N., Degenstein, D., Bourassa, A., Walker, K. A., and Sheese, P.: Validation of ozone profile retrievals derived from the OMPS LP version 2.5 algorithm against correlative satellite measurements, Atmos. Meas. Tech., 11, 2837–2861, https://doi.org/10.5194/amt-11-2837-2018, 2018.

[RD-28] Krol, M., S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener and P. Bergamaschi: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmos. Chem. Phys., 5, 417-432, 2005.

[RD-29] Lerot, C., et al. "Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat and GOME-2/Metop-A." J. Geophys. Res., 119, 3, 1639-1662, doi:10.1002/2013JD020831, 2014.

[RD-30] Loughman, R., Bhartia, P. K., Chen, Z., Xu, P., Nyaku, E., and Taha, G.: The Ozone Mapping and Profiler Suite (OMPS) Limb Profiler (LP) Version 1 aerosol extinction retrieval algorithm: theoretical basis, Atmos. Meas. Tech., 11, 2633–2651, https://doi.org/10.5194/amt-11-2633-2018, 2018.

[RD-31] Merchant, C., et al., 2017, Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data Discuss., vol. 9, p511-527.

[RD-32] Miles, G.M., R. Siddans, B. J. Kerridge, B. G. Latter, and N. A. D. Richards, Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, AMT, doi:10.5194/amt-8-385-2015

[RD-33] Munro, R., R. Siddans, W.J. Reburn, and B.J. Kerridge. "Direct measurement of tropospheric ozone distributions from space." Nature 392 (1998): 168-171.

[RD-34] Newman, P. A., Daniel, J. S., Waugh, D. W., and Nash, E. R.: A new formulation of equivalent effective stratospheric chlorine (EESC), Atmos. Chem. Phys., 7, 4537–4552, https://doi.org/10.5194/acp-7-4537-2007, 2007.

[RD-35] Ohring, G., 2007: Achieving Satellite Instrument Calibration for Climate Change. Eos, Transactions, American Geophysical Union, Vol. 88, Issue 11.

[RD-36] Pedergnana, M., D. Loyola, A. Apituley, M. Sneep, P. Veefkind, Sentinel-5 precursor/TROPOMI L2 Product User Manual O3 Total Column, S5P-L2-DLR-PUM-400A, Issue 01-01-01, 16-07-2018

[RD-37] Petropavlovskikh, I., Godin-Beekmann, S., Hubert, D., Damadeo, R., Hassler, B., and Sofieva, V.: SPARC/IO3C/GAW report on Long-term Ozone Trends and Uncertainties in the Stratosphere, SPARC/IO3C/GAW, SPARC Report No. 9, WCRP-17/2018, GAW Report No. 241, https://doi.org/10.17874/f899e57a20b, 2019. a, b, c, d, e, f, g, h

[RD-38] Prather, M. J.: Numerical advection by conservation of ozone data in weather-prediction models, Q.J.R.Meteorol. Soc., 122, 1545-1571, 1986.

[RD-39] Rahpoe, N., A. Laeng, G. Stiller, M. Weber *(eds)* and R. van der A, C. Adams, P. Bernath, T. von Clarmann, M. Coldewey-Egbers, D. Degenstein, A. Dudhia, R. Hargreaves, C. Lerot, D. Loyola, J. van Peet, V. Sofieva, J. Tamminen, J. Urban, M. Van Roozendael, T. Danckaert, R. Astoreca, K.-P. Heue, P. Sheese, K. Walker and S. Tukiainen, Ozone_cci Phase-II Algorithm Theoretical Basis Document (ATBD), Version 2, Issue 0, Revision 0, Ozone_cci_ATBD_Phase2_V2, December 2017.

[RD-40] Redondas, A. and Cede, A.: Brewer algorithm sensitivity analysis, SAUNA workshop, Puerto de la Cruz, Tenerife, November, 2006.

[RD-41] Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence, Atmos. Meas. Tech., 7, 625–636, https://doi.org/10.5194/amt-7-625-2014, 2014

[RD-42] Siddans, R. Height Resolved Ozone Retrievals from Global Ozone Monitoring Experiment. PhD Thesis, University of Reading, 2003

[RD-43] Siddans, R. et al, Sentinel 5 L2 Prototype Processors, Algorithm Theoretical Basis Document: Ozone Profile. RAL-ESA-S5L2PP-ATBD-001, Version 3.1, 17 May 2019

[RD-44] Siddans, R., CCI+ Water Vapour: ATBD Part 2 - IMS L2 Product, Ref: CCIWV.REP.005, Issue 1, 2019

[RD-45] Siddans, R. et al, Sentinel 4 Level 2 ATBD: Tropospheric Ozone, Document ID S4-L2-RAL-ATBD-2003, Issue 2.2 , Date: 2020-04-21

[RD-46] Sofieva, V. F., Kyrölä, E., Laine, M., Tamminen, J., Degenstein, D., Bourassa, A., Roth, C., Zawada, D., Weber, M., Rozanov, A., Rahpoe, N., Stiller, G., Laeng, A., von Clarmann, T., Walker, K. A., Sheese, P., Hubert, D., van Roozendael, M., Zehner, C., Damadeo, R., Zawodny, J., Kramarova, N., and Bhartia, P. K.: Merged SAGE II, Ozone_cci and OMPS ozone profile dataset and evaluation of ozone trends in the stratosphere, Atmos. Chem. Phys., 17, 12533–12552, https://doi.org/10.5194/acp-17-12533-2017, 2017.

[RD-47] Spurr, R. "LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems." In Light Scattering Reviews, Volume 3, by A. A. Kokhanovsky (ed.). Springer, 2008.

[RD-48] Spurr, R., V. Natraj, C. Lerot, M. Van Roozendael, and D. Loyola. "Linearization of the Principal Component Analysis method for radiative transfer acceleration: Application

to retrieval algorithms and sensitivity studies." J. Quant. Spectrosc. Radiat. Transfer 125 (2013): 1-17.

[RD-49] Spurr, R., D. Loyola, M. Van Roozendael, C. Lerot, K.-P. Heue, J. Xiu, S5P/TROPOMI Total Ozone ATBD, S5P-L2-DLR-ATBD-400A, Issue 2.0, 23/08/2019.

[RD-50] van der A, R.J., M.A.F. Allaart and H.J. Eskes, Multi sensor reanalysis of total ozone Atm. Chem. Phys., 2010, 10, 11277-11294, doi:10.5194/acp-10-11277-2010.

[RD-51] 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, doi:10.5194/amt-8-3021-2015, 2015.

[RD-52] Van Roozendael, M., et al. "Fifteen years of GOME/ERS2 total ozone data: the new direct-fitting GOME Data Processor (GDP) Version 5: I. Algorithm Description." J. Geophys. Res. 117 (2012): D03305

[RD-53] Van Rozendael, M., et al., C3S Ozone Algorithm Theoretical Basis Document (ATBD), Version 1.1a, 3S_312b_Lot2.1.1.2_201902_ATBD_v1.1a, 8/4/2019.

[RD-54] Weber, M et al, Ozone CCI ATBD, Issue 0 - Revision 00 - Status: Final, Date of issue: Dec 7, 2017

[RD-55] Weber, M., Coldewey-Egbers, M., Fioletov, V. E., Frith, S. M., Wild, J. D., Burrows, J. P., Long, C. S., and Loyola, D.: Total ozone trends from 1979 to 2016 derived from five merged observational datasets – the emergence into ozone recovery, Atmos. Chem. Phys., 18, 2097–2117, https://doi.org/10.5194/acp-18-2097-2018, 2018.

[RD-56] WMO (World Meteorological Organization), Scientific Assessment of Ozone Depletion: 2014, World Meteorological Organization, Global Ozone Research and Monitoring Project-Report No. 55, 416 pp., Geneva, Switzerland, 2014.

[RD-57] WMO (World Meteorological Organization), Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project – Report No. 58, 588 pp., Geneva, Switzerland, 2018.

[RD-58] Zawada, D. J., Rieger, L. A., Bourassa, A. E., and Degenstein, D. A.: Tomographic retrievals of ozone with the OMPS Limb Profiler: algorithm description and preliminary results, Atmos. Meas. Tech., 11, 2375–2393, https://doi.org/10.5194/amt-11-2375-2018, 2018.

[RD-59] Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarmann, T., and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb scattering observations, Atmos. Meas. Tech., 11, 2135–2149, https://doi.org/10.5194/amt-11-2135-2018, 2018.

[RD-60] D. Hurtmans, P.-F. Coheur, C. Wespes, L. Clarisse, O. Scharf, C. Clerbaux, J. Hadji-Lazaro, M. George, S. Turquety, FORLI radiative transfer and retrieval code for IASI , March 2012. https://doi.org/10.1016/j.jqsrt.2012.02.036

[RD-61] Thomas August, Dieter Klaes, PeterSchlüssel, TimHultberg, MarcCrapeau, ArlindoArriaga, AnneO'Carroll, Dorothée Coppens. RoseMunro, XavierCalbet, IASI on Metop-A: Operational Level 2 retrievals after five years in orbit, https://doi.org/10.1016/j.jqsrt.2012.02.028

[RD-62] 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., von Clarmann, T., Froidevaux, L., Livesey, N., van Roozendael, M., and Retscher, C.: Measurement report: regional trends of stratospheric ozone evaluated using the MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP), Atmos. Chem. Phys., 21, 6707–6720, https://doi.org/10.5194/acp-21-6707-2021, 2021

[RD-63] Xu, J., Schüssler, O., Loyola, D.G., Romahn, F., and Doicu, A.: A Novel Ozone Profile Shape Retrieval Using Full-Physics Inverse Learning Machine (FP-ILM), IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 5442-5457, doi:10.1109/JSTARS.2017.2740168, 2017.

## 1.5 Acronyms

| | |
|---|---|
| ACE-FTS | Atmospheric Chemistry Experiment – Fourier Transform Spectrometer |
| ATBD | Algorithm Theoretical Basis Document |
| BIRA-IASB | Belgian Institute for Space Aeronomy |
| CCI | Climate Change Initiative |
| CDR | Climate Data Record |
| C3S | Copernicus Climate Change Service |
| DLR | German Aerospace Centre |
| ECMWF | European Centre for Medium-range Weather Forecast |
| ECV | Essential Climate Variable |
| EVISAT | Environmental Satellite (ESA) |
| EO | Earth Observation |
| ESA | European Space Agency |
| EU | European Union |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| FMI | Finnish Meteorological Institute |
| GAW | Global Atmosphere Watch |
| GCOS | Global Climate Observation System |
| GOME | Global Ozone Monitoring Experiment (aboard ERS-2) |
| GOME-2 | Global Ozone Monitoring Experiment – 2 (aboard MetOp-A) |
| GOMOS | Global Ozone Monitoring by Occultation of Stars |
| GOP | GOME-type Ozone Profile |
| GTO | GOME-type Total Ozone |
| IASI | Infrared Atmospheric Sounding Interferometer |
| KNMI | Royal Netherlands Meteorological Institute |
| MetOp | Meteorological Operational Platform (EUMETSAT) |
| MIPAS | Michelson Interferometer for Passive Atmospheric Sounding |
| MLS | Microwave Limb Sounder |

| | |
|---|---|
| NASA | National Aeronautics and Space Administration |
| NDACC | Network for the Detection of Atmospheric Composition Change |
| OMI | Ozone Monitoring Instrument (aboard EOS-Aura) |
| OSIRIS | Optical and Spectroscopic Remote Imaging System (aboard Odin) |
| RAL | Rutherford Appleton Laboratory |
| SCIAMACHY | Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (aboard Envisat) |
| TOMS | Total Ozone Mapping Spectrometer |
| UV | Ultraviolet |

# 2 Total Ozone ECV Retrieval Algorithms

## *2.1 GODFIT (BIRA)*

The GOME-type Direct Fitting (GODFIT) algorithm version 4 [RD-29][RD-19] relies on a direct-fitting approach to retrieve in a one-single step total ozone columns from satellite nadir UV hyperspectral measurements. A non-linear least squares minimization of differences between measured and simulated reflectances is performed in the Huggins bands (fitting window: 325-335 nm) which provides high sensitivity to ozone absorption down to the surface. In addition to total ozone, a number of other parameters form the state vector, including the effective temperature, an effective albedo for the observed scene, and the amplitude of the inelastic structures (Ring effect). Simulations are performed on the fly with the radiative transfer model LIDORT for which the computational performance has been enhanced by application of Principal Component Analysis of the optical properties [RD-36][RD-49]. The RT model also provides the Jacobians required for the inversion. Alternatively, in order to further accelerate the retrievals, the simulated data can be extracted from precomputed look-up tables, e.g. for sensors providing large amount of data.

The algorithm originally developed for analysing spectra recorded by the GOME/ERS-2 instrument [RD-52] has been further developed during the two first phases of the CCI programme to produce consistent level-2 data sets from GOME/ERS-2, SCIAMACHY/Envisat, GOME-2/Metop-A, GOME-2/Metop-B, OMI:Aura and OMPS/Suomi-NPP [RD-29][RD-19]. One particular aspect of the CCI algorithm is that it includes an optional soft-calibration procedure of the L1b data, allowing to further reduce possible systematic biases in the L2 retrievals attributed to limitations in the L1 calibration. This procedure is currently applied to the two GOME-2 instruments and to SCIAMACHY. Thanks to the application of one common retrieval approach and to this soft-calibration procedure, it has been shown that all individual L2 data sets agree with ground-based reference measurements at the percent level [RD-26][RD-19]. The high maturity of the total ozone L2 retrievals developed within CCI allows producing and extending operationally the different level-2 data sets as part of the Copernicus Climate Change Service (C3S) activities. Figure 2-1 illustrates the current total ozone time series obtained when the different data sets available are combined together. The algorithm has been extensively described in previous documents ([RD-36] [RD-53]) and the reader is invited to refer to them for more details. Additional information related to the required input data as well as output variables, including diagnostic metrics, are also provided there.

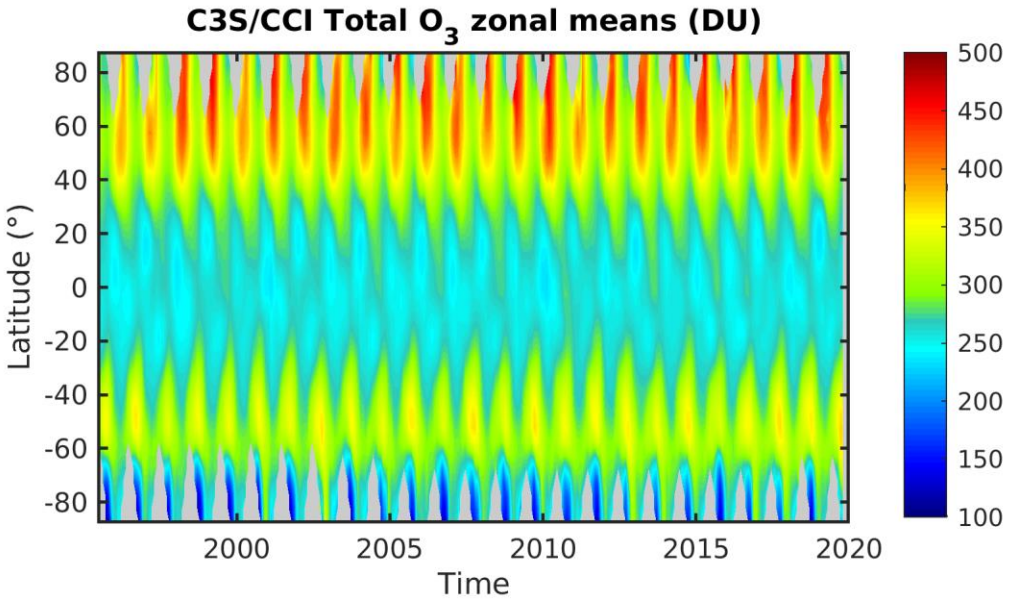


**Figure 2-1: Illustration of the total ozone time series as produced as part of the CCI and C3S activities and obtained by combining different sensors (GOME, SCIAMACHY, GOME-2A/B, OMI, OMPS).**

In the context of CCI+, no new algorithm development is planned but the suite of the L2 total ozone products will be extended by integrating records from the new sensors TROPOMI/S5P and GOME2/Metop-C.

### 2.1.1 Total ozone retrievals from TROPOMI/S5P

TROPOMI aboard the Sentinel-5 Precursor platform is a nadir-viewing instrument and has been launched in October 2017 on a sun-synchronous orbit at 824 km and an equator crossing time of 13:30 local solar time. It records earthshine radiances in spectral ranges from the ultraviolet to the shortwave infrared regions at an unprecedented spatial resolution (3.5 x 7 km² and 3.5 x 5.5 km² after August 2019) and aims at providing key information for the understanding and monitoring of the Earth-atmosphere system, and more particularly of aspects related to ozone layer protection, air quality and climate change.

Although the primary focus of the mission is the tropospheric composition, it will also contribute to extending the time series of European total ozone measurements from space, which have been initiated in the nineties with GOME/ERS-2. For the total ozone, the TROPOMI near-real time (NRTI) and offline (OFFL) products are derived using two different retrieval algorithms. On one hand, ozone columns are produced with the DOAS-type GDP algorithm, which follows the requirements of a near-real time processing. On the other hand, the offline product is based on the direct-fitting algorithm GODFITv4 in order to ensure consistency with long-term total ozone Climate Data Records (CDR). During pre-launch activities, GODFIT has been successfully

implemented within the operational ground-segment, which provides offline total ozone retrievals about 4 days after sensing. Corresponding documentation is given in the total ozone ATBD [RD-49] and the Product User Manual [RD-54]. From the level-2 perspective, the integration of the TROPOMI product within the GTO-ECV data record requires only a verification of its consistency with the existing CCI/C3S data sets.

Figure 2-2 shows comparisons of the TROPOMI offline total ozone product v01.01 with the CCI/C3S OMI, OMPS, GOME-2A and GOME-2B data sets. For each of those sensors, relative differences are shown as a function of altitude and time. The differences are also shown for all sensors when averaged either in time or in latitude (50°S-50°N). The level of consistency with the other data sets is excellent at low and mid-altitudes with differences less than 0.5% and no visible drift. At larger latitudes, differences slightly increase but remain within GCOS requirements (+/-3%). Also those high latitudes differences significantly depend on the reference instrument and no clear pattern can be directly attributed to TROPOMI. First validation results of the TROPOMI total ozone offline product have also shown that it is of high quality, with bias compared to ground-based instruments less than 1% and that no significant dependence with respect to key quantities could be identified [RD-20]. Those different findings confirm that the TROPOMI L2 total ozone data set can be combined with the other instruments to produce and extend the L3 GTO-ECV record.

However, a new TROPOMI L1 version has been recently released including an improved radiometric calibration of the irradiances. The latter directly influences total ozone retrievals, via the effective albedo fit. This change leads to an increase of the offline total ozone columns v02.02 in the range 0.5-2% (low/high latitudes). This is illustrated in Figure 2-3 where a clear discontinuity is visible in July 2021 when the new level-1 version has been activated. Although the aforementioned latitude dependence of the product seems to be improved, the discontinuity is an issue for long-term analysis. This calls for a reprocessing of the entire time series of the operational product based on the newest level-1 data. Note that further developments in level-1 data are on-going, including improvement of the radiance radiometric calibration, wich will also impact the offline total ozone product. This will need to be monitored as well.

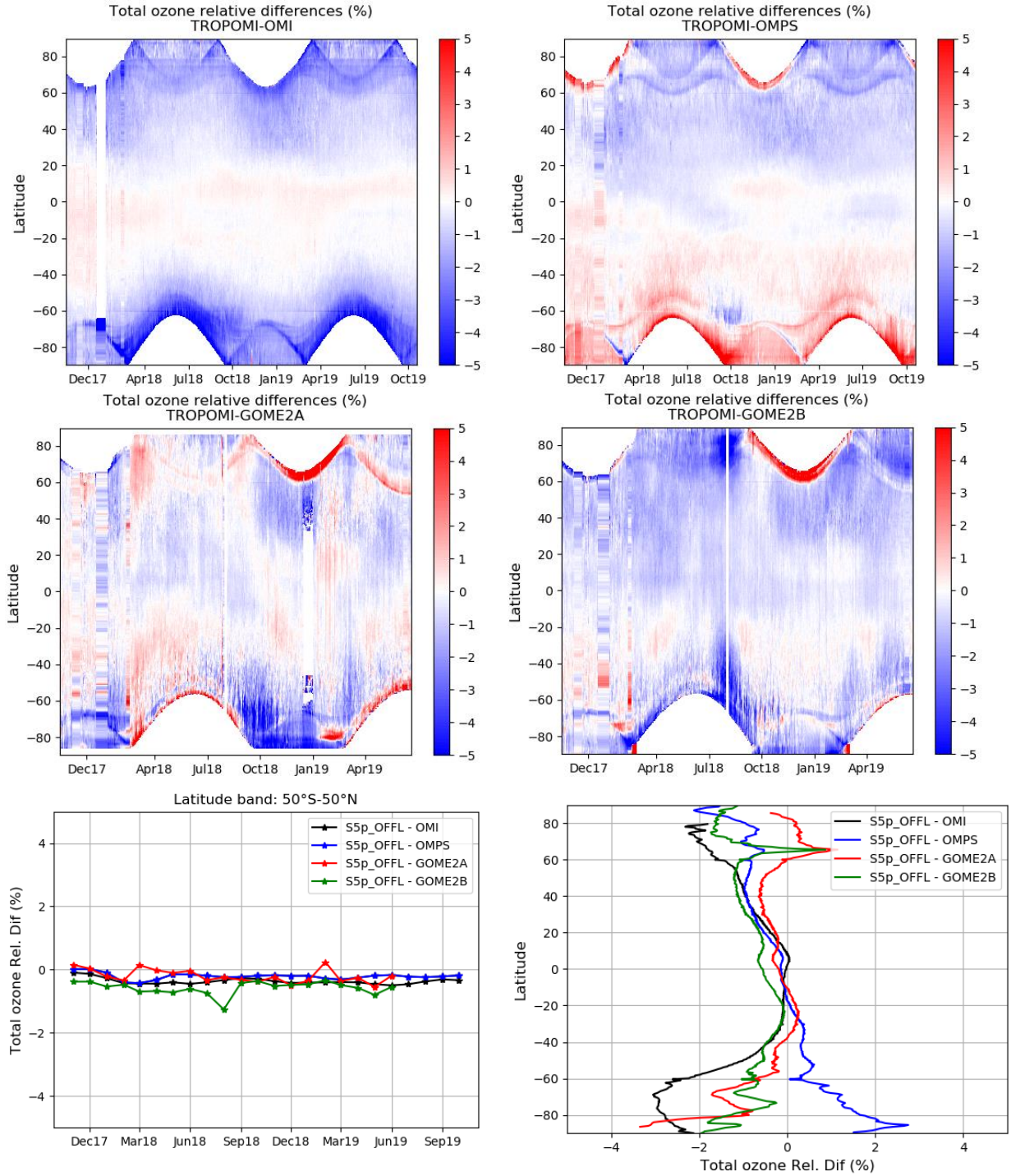


**Figure 2-2: Comparison of the TROPOMI OFFL total product v01.01 with other CCI/C3S data sets. The 4 upper panel show mean relative differences with respect to OMI, OMPS, GOME-2A and GOME-2B as a function of the time and latitude. The two lowest panels summarizes those comparisons by showing the time dependence of the mean differences in the latitude band [50°S-50°N] and the latitude dependence of those mean differences by agglomerating the full time series.**

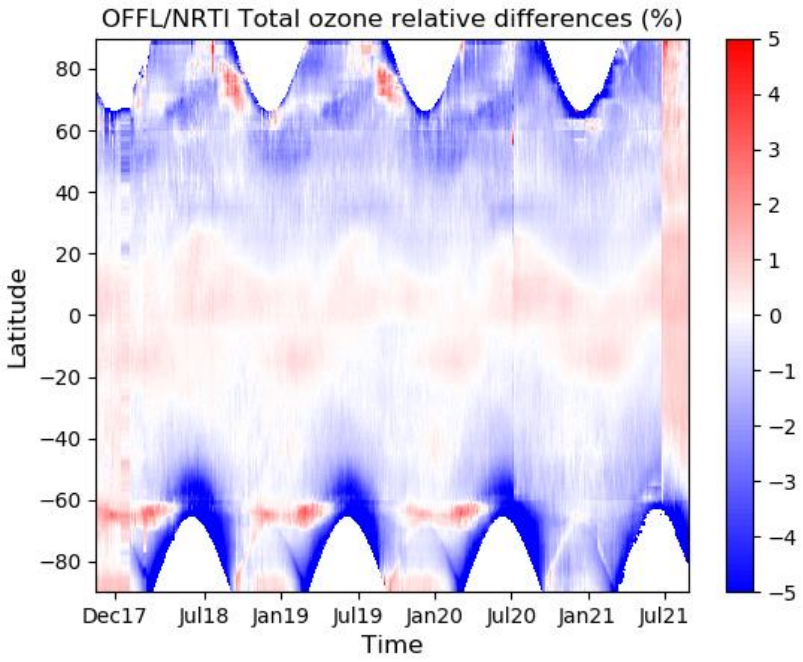


**Figure 2-3: Time series of the zonal mean relative differences between OFFL and NRTI S5p total ozone products.**

### 2.1.2 Total ozone retrievals from GOME-2/Metop-C

The third GOME-2 instrument aboard the Metop-C platform has been launched in November 2018 to continue the series of mid-morning instruments flying on polar sun-synchronous orbit. As it is conceptually the same as its two predecessors, producing total ozone data from GOME-2C mostly requires adaptation of input data (e.g. cross-section) and possibly fine-tuning of the retrieval settings to optimize the consistency with other data sets. As aforementioned, a soft-calibration procedure is applied to the spectra of GOME-2A and -2B. Therefore, the need for the application of this procedure on the GOME-2C spectra has been investigated and turned out not to be necessary. Figure 2-4 shows relative total ozone differences with respect to the columns measured by TROPOMI as a function of time and latitude in the left panel. The right panel shows the time dependence of those differences binned in different latitude bands. Overall, the differences are quite small (<1%) for most latitudes. Nevertheless, they slightly increase at higher latitudes during the local winter season. Although such larger differences are usually observed in those regions, the low bias observed for GOME-2C might be a bit more important at high solar zenith angles. This concerns a limited range of geophysical conditions and may be corrected by the level-3 merging algorithm.

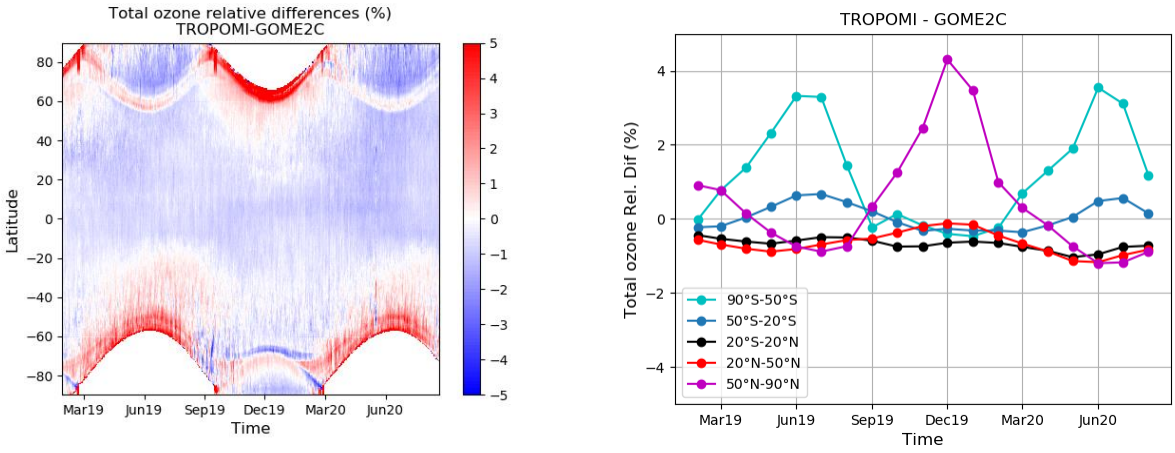


**Figure 2-4: Comparison of the GOME-2C GODFITv4 total ozone product with the TROPOMI OFFL total product v01. The left panel shows mean relative differences as a function of time and latitude. The right panel shows the time dependence of mean differences in different latitude bands.**

## 2.2 GOME-type Total Ozone - Essential Climate Variable (GTO-ECV) (DLR)

### 2.2.1 Total ozone L3 algorithm

The individual level-2 total ozone observations as described in Sect. 2.1 are converted into a level-3 product, i.e. they are mapped onto a regular global grid of 1°x1° in latitude and longitude to construct daily averages for each sensor. Each grid cell contains an average of all level-2 data from the same day, whose center coordinates fall in the respective grid cell.

### 2.2.2 Total ozone merging algorithm

Before combining the individual gridded daily level-3 data into a merged product, adjustments are made in order to account for possible biases and drifts between the instruments. OMI measurements serve as a baseline for the inter-sensor calibration. Their long-term stability with respect to ground-based observation data is remarkable [RD-19] and the periods of overlap with the other sensors are sufficiently long, i.e. at least 5 years. The calculation of the correction factors for GOME, SCIAMACHY, GOME-2A, and GOME-2B with respect to OMI is based on a comparison of 1° zonal monthly means during their overlap periods. The correction factors depend on latitude and time and they are applied to the daily level-3 data. Subsequently, the adjusted data sets are combined into one single record. All available daily measurements (weighted by the number of measurements per day and grid box for the corresponding sensor) are averaged. GOME and SCIAMACHY data are included until December 2004. Finally, monthly means are computed taking into account the latitudinal constraints as defined in Table 2-1, in order to provide representative averages that contain a sufficient number of measurements equally distributed over time (see also [RD-16]).

**Table 2-1: Latitude constraints used in the total ozone merging algorithm**

| Month | Latitudes | Month | Latitudes |
|---|---|---|---|
| January | 60.0°N – 90.0°S | July | 90.0°N – 57.5°S |
| February | 70.0°N – 90.0°S | August | 90.0°N – 62.5°S |
| March | 80.0°N – 80.0°S | September | 82.5°N – 72.5°S |
| April | 90.0°N – 65.0°S | October | 72.5°N – 85.5°S |
| May | 90.0°N – 60.0°S | November | 65.0°N – 90.0°S |
| June | 90.0°N – 57.5°S | December | 60.0°N – 90.0°S |

### 2.2.3 GTO-ECV extension backward in time

For extending the GTO-ECV data record backward in time, we select the Adjusted-MERRA-2 reanalysis data set as described in [RD-17]. It covers the period from 1980 to 2018. We compare 5° zonal monthly means of GTO-ECV and Adjusted-MERRA-2 during the overlap period 1995-2018 and compute corrections factors for the Adjusted-MERRA-2 data in order to calibrate them

with respect to GTO-ECV. The correction reduces potential biases between both data records. It depends on latitude and month (January, February, ...) and is applied to the entire Adjusted-MERRA-2 data set starting in 1980. Finally, both data records are combined; we use the calibrated Adjusted-MERRA-2 from 1980 through June 1995 and GTO-ECV from July 1995 to present.

## 2.3 Multi-Sensor-Reanalysis scheme (KNMI)

### 2.3.1 Introduction MSR algorithm

A single coherent total ozone dataset, called the Multi Sensor Reanalysis (MSR), has been created from all available ozone column data measured by polar orbiting satellites in the near-ultraviolet Huggins band in the last thirty years. All available total ozone satellite retrieval datasets have been used in the MSR. As first step a bias correction scheme is applied to all satellite observations, based on independent ground-based total ozone data from the World Ozone and Ultraviolet Data Center. The correction is a function of solar zenith angle, viewing angle, time (trend), and effective ozone temperature. As second step data assimilation was applied to create a global dataset of total ozone analyses. The data assimilation method is a sub-optimal implementation of the Kalman filter technique, and is based on a chemical transport model driven by ECMWF meteorological fields. The chemical transport model provides a detailed description of (stratospheric) transport and uses parameterisations for gas-phase and ozone hole chemistry. The MSR dataset is available on a grid of a 0.5x0.5 degrees with a sample frequency of 6 hours for the complete time period.

*Constructing the MSR level 2 data set*

Creating a consistent and coherent assimilated dataset of use for trend studies requires that systematic offsets of each one of the satellite retrieval products is small. A practical way to accomplish this is to choose a reference dataset which is available for the full reanalysis period, and subsequently correct the systematic effects in the satellite datasets to bring them in line with this reference. As reference we use the ground measurements from Brewer and Dobson monitoring stations, which are present for the full 30-year period.

First of all, these ground measurements will also show biases, depending on geometry, meteorological variability and ozone profile. The direct Sun measurement method used by the Brewer instruments is very sensitive to small details in the ozone absorption cross section, and the various available laboratory measurements of the ozone absorption coefficients give totally different dependencies of the retrieved total ozone values as function of the effective ozone temperature [RD-40]. Kerr (2002) [RD-25] has developed a new methodology for deriving total ozone and effective ozone temperature values from the observations made with a Brewer instrument. He concludes that the effective ozone temperature has little effect on the amount of ozone derived with the standard algorithm. So in this study the data from the Brewer network has been adopted as a primary reference. The Dobson data compared to Brewer data show a temperature dependence. Therefore, the Dobson data has been corrected for effective ozone temperature and added to this reference data set.

For each satellite product an “overpass” dataset has been created for each ground instrument in our list. The overpass value for an orbit is the satellite observation that has the centre of its footprint closest to the ground station. For each satellite product a maximum allowed distance between the centre of the ground pixel and the ground station was defined. This number is typically 50-200 km depending on the ground pixel size. Apart from the local date/time and the total ozone value, auxiliary data is also recorded, like the measurement error, the Solar Zenith Angle (SZA), the Viewing Zenith Angle (VZA), cloud properties and the distance from the centre of the footprint to the ground station. From all the overpasses each day only one is selected and used. This is the one with the smallest reported observation error or the one closest to the ground station if the observation error is not available.

For the purpose of data assimilation it is relevant to reduce offsets, trends and long-term variations in the satellite data, so that the data can be used as input to the assimilation scheme without biases and with known standard deviations. The satellite data set corrections are based on a few relevant regression coefficients fitted for the overpass time series of all stations together. By fitting all data together regional biases that may be caused by offsets of individual ground instruments are avoided.

The ozone differences (satellite minus ground observation) show a clear seasonal cycle. This led to the choice of SZA and effective ozone temperature as fit parameter, as these imply a clear seasonal component. Some of the satellite products show a clear trend in time, so date/time is another obvious choice. The Viewing Zenith Angle (VZA) is also used as fit parameter. These are all critical parameters in the retrieval schemes and therefore constitute a satisfying choice to estimate systematic biases. A ground-station dependent offset was allowed when the regression coefficients were computed. This has been done to reduce the effect (e.g. spurious trends) of “appearing” and “disappearing” ground stations during the lifetime of the satellite instrument from the results. Thus, the total number of fit parameters is in the order of 150 per satellite dataset. A basic assumption is that all the corrections are additive to the total ozone amount.

Based on the calculated corrections the merged MSR level 2 dataset has been created. The original satellite datasets were read, filtered for bad data and corrected, and finally merged into a single time ordered dataset. Essential information in the MSR level 2 dataset is time, location, satellite product index and ozone.

*Data assimilation*

The satellite instrument observations are combined with meteorological, chemical and dynamical knowledge of the atmosphere by using data assimilation. The data assimilation scheme used here is called TM3DAM and is described in Eskes et al. (2003). The chemistry-transport model used in this data assimilation is a simplified version of TM5 [RD-28], which is driven by ECMWF analyses of wind, pressure and temperature fields. As input the MSR ozone values and the estimates of the measurement uncertainty are used.

The three-dimensional advection of ozone is described by the flux-based second order moments scheme of Prather et al. (1986) [RD-38]. The model is driven by 6-hourly meteorological fields (wind, surface pressure, and temperature) of the medium-range meteorological analyses of the ECMWF. The assimilation is using the ERA interim or ERA5 reanalysis whenever available. The ECMWF hybrid layers between 0.01 hPa and the surface have been converted into the 44 layers used in TM3DAM. The horizontal resolution of the model version used in this study is 1 x 1 degrees. This resolution is compensated by the practically non-diffusive Prather scheme (with 10 explicit ozone tracers for each grid cell), which allows the model to produce ozone features with a fair amount of detail.

Ozone chemistry in the stratosphere is described by the Cariolle version 2.9 parameterisation [RD-12]. This consists of a linearization of the gas-phase chemistry with respect to production and loss, the ozone amount, temperature and UV radiation. In addition, a second parameterization scheme accounts for heterogeneous ozone loss. This scheme introduces a three-dimensional chlorine activation tracer, which is formed when the temperature drops below the critical temperature of polar stratospheric cloud formation. Ozone breakdown occurs in the presence of the chlorine activation tracer, depending on the presence of sunlight. The rate of ozone decrease is described by an exponential decay, with a rate proportional to the amount of activation tracer below the critical temperature and with a minimal decay time of 12 days. The cold tracer is deactivated when light is present with a time scale of respectively 5 and 10 days on the Northern and Southern hemisphere.

The total ozone data are assimilated in TM3DAM by applying a parameterized Kalman filter technique. In this approach the forecast error covariance matrix is written as a product of a time independent correlation matrix and a time-dependent diagonal variance. The various parameters in the approach are fixed and are based on the forecast minus observation statistics accumulated over the period of one year (2000) using GOME observations. This approach produces detailed and realistic time- and space-dependent forecast error estimates, which is included in the MSR level 4 product.

More details of the MSR algorithm can be found in [RD-51] and [RD-50].

### 2.3.2 Algorithm update to extend the MSR data set into the past

*Heterogeneous chemistry*

An EESC dependent heterogeneous chemical destruction of stratospheric ozone factor has been be incorporated within the exponent of the ozone depletion term within the Braesicke scheme.

The EESC correction factor $f_{CFC}$ is calculated according

$$F_{CFC} = (EESC/EESC_{max})^2 \text{ where } EESC_{max} = 4092.8508$$

**Equation 2-1**

The EESC values are derived in [RD-34] and are calculated for the period 1957-2020.

*Solar Zenith Angle dependence of photolysis factor*

A linear correction factor $f_{SZA}$ for photolysis during twilight has been derived. The twilight is the zone with a Solar Zenith Angle (SZA) between 85 and 94 degree.

$$f_{SA} = (94.0 – SZA)/(94.0 – 85.0) \text{ for SZA between 84.0-95.0 degree}$$

**Equation 2-2**

This factor has been multiplied with the chlorine activation term [RD-18].

*Error correlation length*

For sparse observations it is likely that the error correlation length in the spatial distribution of assimilated ozone is different. Therefore the correlation length is determined by assimilating ozone using only sparse Dobson observations in 2017. The resulting ozone has been compared with the MSR2 ozone distribution that represent the “true” ozone distribution. From analysing these datasets we conclude that the correlation length is about 800 km.

# 3 Nadir Profile ECV Retrieval Algorithms

## *3.1 RAL nadir profile ECV retrieval algorithms (RAL)*

The RAL profile scheme applied to GOME, GOME-2, SCIAMACHY and OMI within the CCI and C3S is described in ATBDs [RD-54], [RD-53], as well as [RD-32], [RD-42] and [RD-32]. The scheme is also currently being developed for application to Sentinel 4 and 5 via ESA contracts and documented via related ATBDs [RD-45] and [RD-43]. In general, the scheme applies the optimal estimation approach to retrieve the ozone profile in two main steps:

1. "B1 fit": Fit ozone profile to the sun-normalised radiance in the Harley band (in GOME Band 1) from 265-307nm. This provides information mainly on the stratospheric profile and requires good absolute calibration of the sun-normalised radiance spectra. Soft calibration of the B1 L1 radiances in usually required (i.e. radiances are corrected based on differences between L1 radiances and calculated spectra based on radiative transfer calculations from prior knowledge of the ozone distribution).

2. "B2 fit": Add information on tropospheric ozone from the Huggins bands (320-340nm) using the result from step 1) as prior constraint. This requires fitting of differential structure to precision better than 0.1% (close to the noise level) to allow the ozone absorption cross-section temperature dependence to be exploited for tropospheric information (though the requirements on absolute radiometic calibration are less stringent). Extremely good knowledge of the instrument spectral response function is critical for this step.

Work within CCI+ focuses on the application of the algorithm to new sensors GOME-2 Metop C and S5P. For S5P several advances made for Sentinel 4 and 5 will be exploited, in particular:

- The core radiative transfer model is LIDORT [RD-47].

- The speed of the forward model is greatly increased using a "PCFM" approach which requires the core RTM to be run for only a few monochromatic spectral points.

- The B1 uv soft-calibration approach will be based on that developed for S5 [RD-43], based on calculations based on ERA5 meteorology and ozone climatology.

- A new approach to mitigate scene inhomogeneity effects on the spectral response function will be tested using an approach developed for S4 [RD-45].

## 3.2 GOME-type Ozone Profile - Essential Climate Variable (GOP-ECV) (DLR)

Work within CCI+ Phase-I focuses on the development of a merged ozone profile climate data record based on the same satellite sensors as the well-established total column product GTO-ECV (see Sect. 2.2). The approach, that is applied to the ozone profiles, consists of two main steps: (1) merging the time series of the individual sensors into a combined product and (2) harmonization of the combined record with respect to the GTO-ECV total columns in order to achieve broad consistency between both data sets.

### 3.2.1 Ozone profile merging algorithm

At first, the individual gridded monthly level-3 ozone profile data are merged into a combined record. The level-3 partial column data are based on the level-2 data retrieved with the RAL nadir profile retrieval scheme as described in the previous section (Sect. 3.1). Data from five sensors are included: GOME (1995-2011), SCIAMACHY (2002-2012), OMI (2004-present), GOME-2A (2007-2021), and GOME-2B (2013-present). The values given in parenthesis denote the time periods which are covered by the sensors. These sensors are also used to create the GTO-ECV record. Note that data from TROPOMI and GOME-2C are not yet included in this first version of GOP-ECV. The merging is based on deseasonalized anomalies and similar to the method proposed by Sofieva et al. (2021) [RD-62]. Deseasonalized anomalies $\Delta O_{3,i}(\varphi,\lambda,z,t)$ at latitude $\varphi$, longitude $\lambda$, altitude z, and month t are computed for the time series of all individual instruments *i* following:

$$\Delta O_{3,i}(\varphi,\lambda,z,t) = O_{3,i}(\varphi,\lambda,z,t) - O_{3,i,c}(\varphi,\lambda,z,m),$$

**Equation 3-1**

where $O_{3,i}(\varphi,\lambda,z,t)$ is the monthly mean ozone profile. $O_{3,i,c}(\varphi,\lambda,z,m)$ is the corresponding climatological mean value for the month m with m=January,…, December. Since the individual sensors cover only partially the same periods, the corresponding seasonal cycles have to be computed for different reference periods as follows: GOME (1996-2002), SCIAMACHY (2005-2010), OMI (2005-2010), GOME-2A (2007-2016), GOME-2B (2015-2020). The anomalies for OMI and SCIAMACHY are already aligned and can be merged directly. For the other sensors (GOME, GOME-2A, and GOME-2B) additive offsets w.r.t. OMI anomalies are computed from the overlap periods. These offsets depend on latitude, longitude, and altitude. After the application of the offsets, the merged ozone profile for each spatio-temporal bin is obtained from the median of all available instruments as :

$$\Delta O_{3,mer}(\varphi,\lambda,z,t) = \text{median}(\Delta O_{3,i}(\varphi,\lambda,z,t))$$

**Equation 3-2**

Figure 3-1 shows the anomalies 1995-2022 for all sensors for the latitude/longitude bin 12.5°N/0.5°E and for altitude 20km. In addition, the adjusted anomalies for GOME, GOME-2A,

and GOME-2B are shown. Finally, in order to reconstruct absolute ozone values, the seasonal cycle obtained from OMI measurements is used and added back to the anomalies:

$$O_{3,mer}(\varphi,\lambda,z,t) = \Delta O_{3,mer}(\varphi,\lambda,z,t) + O_{3,omi,c}(\varphi,\lambda,z,m).$$

**Equation 3-3**

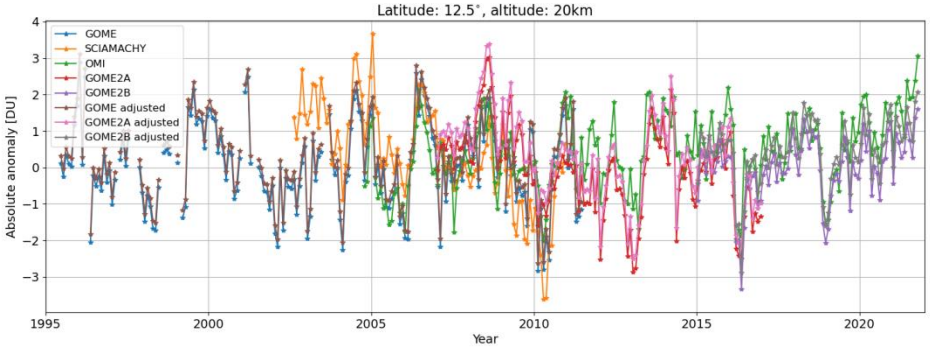


Figure 3-1 Deseasonalized anomalies [DU] 1995-2022 for all satellite sensors for the latitude/longitude bin 12.5°N/0.5°E and for the altitude of 20km. In addition, for GOME, GOME-2A, and GOME-2B the adjusted anomalies w.r.t. OMI anomalies are shown.

### 3.2.2 Ozone profile merging algorithm

The aim of the harmonization of the merged ozone profiles, generated as described in the previous section, is to achieve consistency between the GTO-ECV total column product and the new GOP-ECV ozone profile record with respect to the total column. The harmonization consists of the following steps:

1. A clustering algorithm is applied to a subset of 80,000 profiles randomly selected from the entire merged data record. We use the same *k*-means clustering procedure as described in Xu et al. (2017) [RD-63] and applied to a set of ozone profiles. From this study we also adopt the optimal number of clusters: K=11. Figure 3-2 shows the output of the clustering: the ozone profiles in each class. Colors denote the total column of each profile. Profiles of similar shapes are grouped into the same class. Class 03 contains mainly profiles from the Northern Hemisphere middle latitudes in boreal spring (maximum total ozone columns >400DU) and class 05 contains mostly profiles from high latitudes in the Southern Hemisphere from September to November (ozone hole season).
2. For each class of ozone profiles a separate Neural Network is trained using as input the total ozone column and as output the corresponding ozone profile shape. The NNs comprise two hidden layers. From the NN training we extract the derivatives for each class, which provide the information about the altitude-dependent change of the profile, when the

total column changes. Figure 3-3 exemplarily shows the derivatives for class 03. The shape of the derivatives depend on the total column, and derivatives can be negative.

3. In order to assign one of the 11 classes to each of the remaining profiles, which were not used for the clustering procedure (step (1)), a *k*-nearest neighbors classifier method is applied. The training data set for the classifier contains the profiles of the 11 clusters and the class label (1, …, 11). The fitted classifier is then applied to the remaining profiles in order to predict the corresponding class label, which is needed to select the correct set of derivatives for the altitude-dependent scaling.
4. The last step is the application of the scaling procedure to all merged profiles. The difference ΔTOZ between the total column from the integrated profile and the corresponding total column from the GTO-ECV data record is computed for each spatiotemporal bin and the adjusted profile is computed as follows:

$$O_{3,new}(\varphi,\lambda,z,t) = O_{3,mer}(\varphi,\lambda,z,t) + \Delta TOZ(\varphi,\lambda,t) * d/dTOZ(z),$$

**Equation 3-4**

$O_{3,mer}(\varphi,\lambda,z,t)$ denotes the merged profile from Sect. 3.2.1, $O_{3,new}(\varphi,\lambda,z,t)$ is the adjusted profile which is consistent with GTO-ECV, and d/dTOZ(z) is the derivative derived from the NN in step (2).

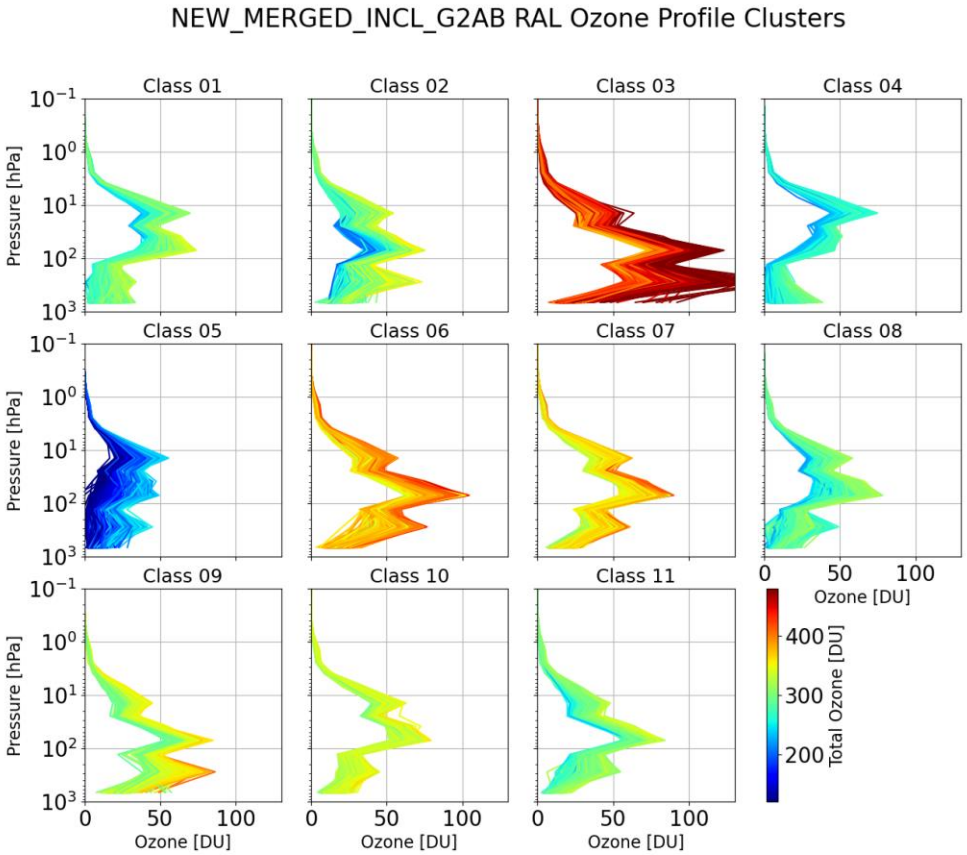


Figure 3-2 Merged ozone profiles grouped in 11 classes using a *k*-means clustering procedure. The total number of profiles is 80,000 randomly selected from the entire data record. Colors denote the total column of each profile.

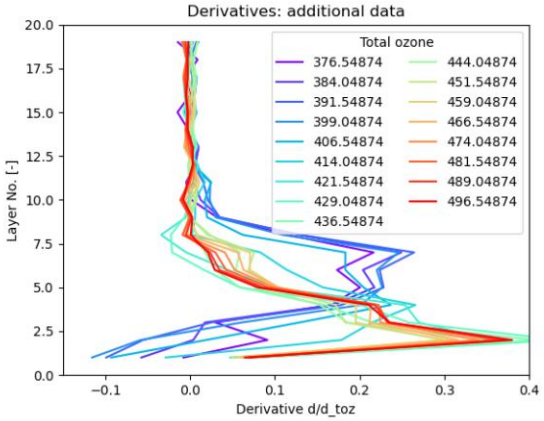


Figure 3-3 Derivatives extracted from the NN training for class 03 (see Figure 3-2). Colors denote the total column.

## 3.3 IASI FORLI Ozone profile retrieval algorithm (ULB)

The IASI ozone profile data product is based on the FORLI (Fast Optimal/Operational Retrieval on Layers for IASI) algorithm. FORLI is a line-by-line radiative transfer model capable of processing in near-real-time the numerous radiance measurements made by the high-spatial and high-spectral resolution IASI, with the objective to provide global concentration distributions of atmospheric trace gases.

The FORLI-O3 v20151001 product was developed and validated during the Ozone_cci Phase-II and constitutes the reference product from 1rst October 2007 to 11 December 2019. That IASI/Metop-A FORLI-$O_3$ dataset has been extensively validated in [RD-11] and [RD-24]. The FORLI software v20151001 [RD-60] was implemented at EUMETSAT in 2019 and the IASI $O_3$ product is operational at EUMETSAT and distributed via Eumetcast in BUFR format since 4 December 2019. The Eumetsat IASI O3 BUFR files are reformatted in netcdf format by LATMOS and are now distributed by AERIS. They constitute the new CCI database from 4 December 2019. Hence, we recommend to use the ULB-LATMOS FORLI-v20151001 dataset from 20071001 to 20191204 and the EUMETSAT FORLI-v20151001 dataset afterwards, both being available on the Aeris portal at: http://iasi.aeris-data.fr/O3/

An updated FORLI-$O_3$ v20191122 is now used for the processing of the IASI dataset from 12 December 2019 till present. The backprocessing of the whole IASI dataset with the last v20191122 is ongoing and will be archived on the Aeris portal in the future.

Finally, a dedicated version of FORLI (FORLI-climate) that uses the ERA-5 analysis as input for temperature and humidity profiles to improve the climate record is being implemented on the French CICLAD-IPSL computing center and tested.

This part describes the methods used for FORLI (most is extracted from [RD-23]) and is an update of the Ozone_cci ATBD Phase-II. We refer the readers to the Ozone_cci ATBD Phase-II when the section remains unchanged.

### 3.3.1 Basic retrieval equations

Refer to the Ozone_cci_ATBD_Phase2_V2

### 3.3.2 Assumptions, grid and sequence of operations

*Spectral ranges*

FORLI-$O_3$ (v20151001 and v20191122) uses the Level1C radiances disseminated by EumetCast. A subset of the spectral range, covering 1025–1075 $cm^{-1}$, is used for the $O_3$ retrieval. The spectral range used in the forward model is 960-1075 $cm^{-1}$ and the spectral oversampling is 100.

*Vertical grid*

FORLI-$O_3$ uses a vertical altitude grid in km with 1-km tick layers as discretization of the atmosphere

*Ozone state vector*

The ozone product from FORLI is a profile retrieved on 40 1km-thick layers between surface and 40 km, with an extra layer from 40 to TOA considered at 60 km.

The *a priori* profile $\mathbf{x_a}$ covariance matrix $\mathbf{S}_a$ are constructed from the McPeters/Labow/Logan climatology of ozone profiles (McPeters et al., 2007), which combines long term satellite limb measurements (from the Stratospheric Aerosol and Gas Experiment II and the Microwave Limb Sounder) and measurements from ozone sondes. The *a priori* profile $\mathbf{x}_a$ is the mean of the ensemble. Fig. 1 illustrates this *a priori* information: the *a priori* profile $\mathbf{x}_a$ has values slowly increasing from around 25 ppbv at the surface to 100 ppbv at 10km, reaching a maximum of 7.3 ppmv in the middle stratosphere. The variability (taken hereafter as the square root of the variance, i.e. of the diagonal elements of $\mathbf{S}_a$) is below 30% in the boundary layer and the free troposphere; it is maximum in the upper troposphere–lower stratosphere, between 10 and 20 km, where it is of the order of 60%. There is significant correlations between the concentrations in the layers 0–10, 10–25 and 25–40 km, but weak correlation between these three (Fig.8).

*Other state vector elements*

Besides the ozone profile, surface temperature and the water vapour column are retrieved.

*Measurement covariance matrix*

$\mathbf{S}_\eta$ is taken diagonal. The value of the noise is wavenumber dependent in the spectral range used for the retrieval, varying around 2 x $10^{-8}$ W/($cm^2$ .$cm^{-1}$ .sr).

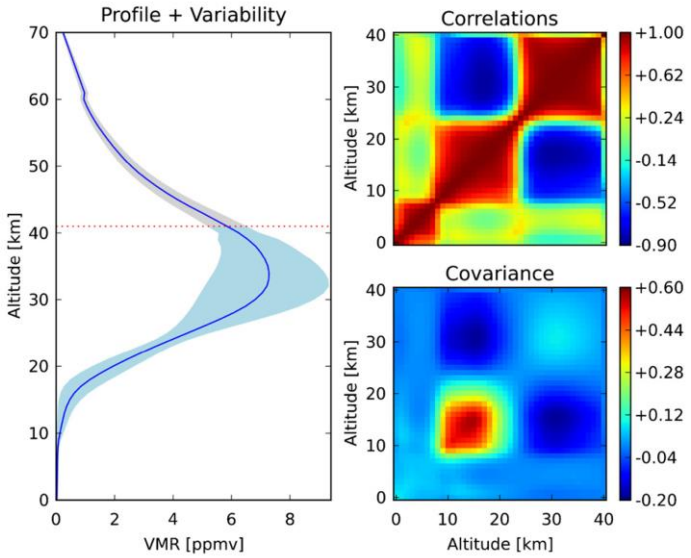


**Figure 3-4. Left: *xa* (ppmv, blueline) and associated variance (shaded blue) for the FORLI-O3. The dashed red line indicates the top altitude of the last retrieved layer. Right: correlations and *Sa* variance–covariance matrices in unitless multiplicative factor. (from Hurtmans et al. 2012)**

### 3.3.3 Iterations and convergence

Refer to Ozone_cci_ATBD_Phase2_V2

### 3.3.4 Forward model

***Atmospheric state input to the RTM***

L1C radiances

FORLI-$O_3$ uses the Level1C radiances disseminated by EumetCast. A subset of the spectral range, covering 1025–1075 $cm^{-1}$, is used for the $O_3$ retrieval.

Temperature and humidity profiles

Profiles of temperature and humidity are from the IASI L2 PPF [RD-61] for both FORLI-v20151001 and FORLI-v20191122, while they are taken from the ERA-5 reanalysis for FORLI-Climate . The atmospheric temperatures are kept fixed whereas the water profile is used as a priori and further adjusted.

### Surface temperature

Surface temperatures (land and sea) are from the IASI L2 PPF (FORLI-v20151001 and –v20191122) or from the ERA-5 reanalysis (FORLI-Climate). Surface temperature is part of the parameters to be retrieved.

### Cloud fraction

FORLI-$O_3$ (v20151001 and v20191122) uses the cloud fraction from the IASI L2 PPF. All pixels with a cloud fraction equal to or lower than 13 % are processed.

### $CO_2$ profile

FORLI-$O_3$ v20151001 assumes a constant vertical profile at 380 ppm for $CO_2$, while the new FORLI-$O_3$ v20191122 and FORLI-Climate considers time-varying $CO_2$ concentrations according to the Keeling curve.

### Orography

Orography is from the GTOPO30 global digital elevation model and is integrated in the entire IASI FOV (http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info).

### Emissivity

A wavenumber-dependent surface emissivity above continental surfaces is used while for ocean a single standard emissivity is considered. For continental surfaces, it relies on the climatology of Zhou et al. (2011). In cases of missing values in the Zhou et al. climatology, the MODIS climatology of Wan (2008) is used. It is available on a finer 0.05° × 0.05° grid but is restricted to only 12 channels in the IASI spectral range. In order to deal with this, the spectrally resolved mean emissivity of the Zhou climatology is scaled to match as closely as possible the values in these 12 channels and it is this resulting emissivity that is considered. Finally when there is no correspondence between the IASI FOV and either climatologies, then the mean emissivity of the Zhou climatology is used.

### Lookup-tables

Tabulated absorption cross-sections at various pressures and temperatures are used to speed up the radiative transfer calculation. The spectral range for the LUTs used in FORLI-O3 is 960-1075 $cm^{-1}$ and the spectral oversampling is 100. The absorption cross-sections are computed on a logarithmic grid for pressure from $4.5\times10^{-5}$ to 1 atm with a grid step of 0.2 for the logarithm of pressure, and on a linear grid for temperature (162.8–322.6 K with a grid step of 5K). Relative humidity is also introduced in the LUT, varying linearly between 0 and 100%, by steps of 10%.

The FORLI-$O_3$ v20191122 has undergone corrections in the computation of the look-up tables.

### Spectroscopy

Line integrated absorption cross section, air broadening, self-broadening, line shifting and absorption cross section data are taken from the widely used HITRAN spectroscopic database

version 2008 (Rothman et al., 2009). Continuum formulations are taken from MT-CKD (Clough et al., 2005).

The last FORLI versions (v20191122 and FORLI-Climate) uses the Hitran database update to the latest available version with largely corrected CO line intensities and positions and also updates for $HNO_3$, as well as MT_CKD update and the related use of Line-Mixing for $CO_2$ lines.

Note that other corrections have also been implemented:  altitudes computation, correct usage of humidity …

#### *Radiative Transfer Model (RTM)*

Refer to Ozone_cci_ATBD_Phase2 for a full description of the RTM (general formulations and numerical approximations).

### 3.3.5 Error description

Refer to Ozone_cci_ATBD_Phase2.

### 3.3.6 Output product description

#### *Formats*

The FORLI-$O_3$ retrieval datasets for IASI-A, -B and -C processed with the ULB-LATMOS FORLI-$O_3$ v20151001 from 1rst October 2007 to 4 December 2019 and with the EUMETSAT FORLI-O3 v20151001 dataset afterwards till present are delivered in NetCDF (v4) format and can be downloaded from the Aeris portal at: http://iasi.aeris-data.fr/O3/.

#### *Ozone profile and characterization*

The ozone product from FORLI is a profile retrieved on 40 layers between surface and 40 km, with an extra layer from 40 km to the top of the atmosphere (TOA) considered at 60 km. The dataset includes $O_3$ total columns along with vertical profiles. It also includes other relevant information such as the a priori profile, the total error profile and the averaging kernel (AK) matrix, on the same vertical grid.

### 3.3.7 Retrievals and Quality flags

A series of Quality input and processing flags were applied to the FORLI-$O_3$ datasets to exclude bad quality data, specifically when:

(i) The input values (T, Q, Cloud) are missing, or in case of negative surface altitudes or unrealistic skin temperature

(ii) The spectral fit residual root mean square error (RMS) is higher than $3.5\times10^{-8}$ W/(cm$^2$ sr cm$^{-1}$), reflecting a too large difference between observed and simulated radiances
(iii) The spectral fit residual bias is lower than -0.75 x $10^{-9}$ W/(cm$^2$ sr cm$^{-1}$) or higher than 1.25 x $10^{-9}$ W/(cm$^2$ sr cm$^{-1}$)
(iv) The spectral fit residual is sloped
(v) The partial $O_3$ column or the humidity is negative
(vi) There were abnormal averaging kernel values. These are identified checking for unrealistic values – larger than 2 – or unrealistic strong gradient/oscillation on the averaging kernel profiles)
(vii) The spectral fit diverged or reached a fixed maximum number of iterations without converging, or the Chi-Square value is too high
(viii) The total error covariance matrix is ill conditioned.
(ix) No retrieval is done due to incorrect inputs or other reasons

For an optimal use of the data (e.g. for validation application), users should also exclude data when:

(i) The $O_3$ profiles have an unrealistic C-shape (i.e. abnormal increase in $O_3$ at the surface, e.g. over desert due to emissivity issue), with a ratio of the surface – 6 km column to the total column higher or equal to 0.085
(ii) The DOFS is lower than 2, which are mostly associated with bad quality data in the Antarctic region.

## *3.4 Combined uv/vis/thermal-ir retrieval algorithm (RAL)*

### 3.4.1 Overview

Within CCI+ (mainly in year 2) RAL will investigate improve the information content of the GOME-2 nadir ozone profiles by adding information from (i) the visibile Chappuis band (also measured by GOME-2) and (ii) thermal IR measurements by IASI (ii). The approach is to combine information from indendent level 2 (L2) retrievals from the different spectral ranges, i.e. adopting a so-called “L2-L2” combination approach. Sections below described the Chappuis and IR retrieval schemes to be employed, following by an overview of the L2 combination methodology.

### 3.4.2 Chappuis retrieval scheme

The potential advantages of using the Chappuis bands (440-700nm) are well known [RD-13]: The Chappuis absorption is generally optically thin with very little temperature dependence. There is consequently little sensitivity to vertical profile shape (only total column information can be retrieved from the Chappuis band alone). However it can complement the uv measurements of total ozone because near-ground sensitivity to ozone in the Huggins bands is limited by the typically low surface reflectance and Rayleigh / aerosol scattering. In the Chappuis range, the albedo over land is typically larger and Rayleigh scattering much reduced, so that a larger proportion of the observed photons will have passed through the lowest atmospheric layers. Extracting the signal from the Chappuis bands is complicated by the fact that the amplitude of differential structures (which are exploited by DOAS techniques) are relatively low and have a relatively broad band structure compared to the features in the Huggins bands. Such features may be easily confused with a number of other effects e.g. instrumental artefacts (such as polarisation sensitivity, aliasing of spatial/spectral structure due to non-synchronous detector pixel read-out) and spectral variations in surface reflectance. In addition, there are contaminating spectral features from water vapour (H2O), oxygen dimer (O4) and nitrogen dioxide (NO2). Although Rayleigh scattering is reduced, the Ring effect (due to inelastic scattering by air molecules) is still significant. Inelastic scattering at the ocean surface is also important over sea, although in conditions where this is important, the information content of the Chappuis bands is less useful (as the ocean surface reflection is very low).

Via UK National Centre for Earth Observation (NCEO) funded work (to be consolidated in the coming months) two approaches have been developed to extract the total (slant) column of ozone from the Chappuis: (1) A DOAS based approach, which has a directly physical basis. This involves fitting absorption cross sections for all relevant gases and scale factors for terms related to known surface spectral variability and other known effects (Ring effect, spectral shifts, instrument polarisation features etc); (2) A statistical approach based on regressing the Chappuis band measurements to the ozone slant column retrieved from the Huggins bands. This involves determine singular vectors of the Chappuis band spectral variability which is uncorrelated with the total ozone variability enabling total ozone to be retrieved along with spatial patterns which are related to other physical variables. In both cases the slant column of ozone is estimated. The sensitivity of the slant column to height resolved perturbations in ozone can be calculated, forming effectively the averaging kernel of the retrieved slant column with respect to changes in ozone profile. This information can then be combined with the uv retrieval using the L2-L2 approach outlined below.

Further details of the consolidated approach will be given in the next version of this ATBD.

### 3.4.3 RAL Infra-red / Microwave sounder retrievals

Also within NCEO and supported by Eumetsat studies, RAL have developed an infra-red microwave sounder (IMS) scheme which retrieves ozone profiles along with temperature,

humidity and other atmospheric variables from Metop IASI, AMSU and MHS. Products from the scheme have been validated in CCI+ Water Vapour and will be used to generate a combined limb/nadir sounder product in that project. The scheme is described in the CCI+ water vapour ATBD [RD-44], though the current version of the scheme uses an extended set of channels to improve the sensitivity to the ozone profile. The scheme is based on optimal estimation and provides all the necessary information to enable L2-L2 combination with the uv profiles as described below.

### 3.4.4 L2-L2 Combination

The L2-L2 combination is implemented by posing the problem as a linear retrieval in which we wish to optimally estimate a profile by combining the information contained in different retrievals. The approach should work for a retrieval of any property linearly related to the profile e.g. (sub-)column amounts or mixing ratio profiles on a coarsely sampled vertical grid (provided the transformation to a fine grid is clearly defined). For the approach to work, it is assumed that averaging kernels are provided (or can be constructed) for all input retrievals, with respect to fine scale perturbations in the true profile (defined in the same units on the same finely resolved grid).

We choose to represent the (output, optimised) profile in a flexible way on a fine vertical grid using N basis functions (this can be different from the way the input retrievals represent the profile):

$$\boldsymbol{r}(z) = o(z) + \sum_{i=1,N} x_i B_i(z)$$

**Equation 3-5**

Where $o(z)$ is an "offset" profile (which can be considered as the new prior profile on the fine grid). $B(z)$ are a set of suitable basis functions (e.g. triangular functions representing linear interpolation from a coarse vertical grid, or principal components of the assumed profile variability etc). The N element vector $\boldsymbol{x}$ contains state vector elements to be retrieved.

Alternatively, considering vector $\boldsymbol{r}$ to describe the profile on a finely resolved vertical grid:

$$\boldsymbol{r} = \boldsymbol{o} + \boldsymbol{B}\boldsymbol{x}$$

**Equation 3-6**

Where B is a matrix containing the N basis functions.

The optimal $\boldsymbol{x}$ (and hence $\boldsymbol{r}$) can be obtained using the input retrievals as the "measurements" for an OEM retrieval. These are contained in a measurement vector, $\boldsymbol{y}$. We can relate each to $\boldsymbol{r}$, using the averaging kernels as the forward model function:

$$y_j = \boldsymbol{A}_j\left(\boldsymbol{r} - \boldsymbol{r_{aj}}\right) + a_j$$

**Equation 3-7**

Where index j represents a specific sub-column amount (from TIR or SWIR). $\boldsymbol{A_j}$ is the averaging kernel (vector) for this sub-column (from the TIR or SWIR retrieval), describing the derivative of the sub-column with respect to perturbations *on the fine grid*. $\boldsymbol{r_{aj}}$ is the (finely resolved) prior profile used in the previous (input) retrieval and $a_j$ is the corresponding prior sub-column amount.

Substituting for $\boldsymbol{r}$ gives

$$y_j = \boldsymbol{A}_j\left(\boldsymbol{o} + \boldsymbol{B}\boldsymbol{x} - \boldsymbol{r_{aj}}\right) + a_j$$

**Equation 3-8**

Using this (linear) forward model for $F(x)$, $\boldsymbol{x}$ can be estimated via minimisation of the usual cost function, $\chi^2$, i.e.

$$\chi^2 = (\boldsymbol{x} - \boldsymbol{x_a})^T \boldsymbol{S_a^{-1}}(\boldsymbol{x} - \boldsymbol{x}_a) + \left(\boldsymbol{y} - F(\boldsymbol{x})\right)^T \boldsymbol{S_e^{-1}}\left(\boldsymbol{y} - F(\boldsymbol{x})\right),$$

**Equation 3-9**

Where the measurement covariance $\boldsymbol{S_e}$ constains the estimated error covariances of the (TIR+SWIR) subcolumns in $\boldsymbol{y}$. *A priori* covariance $\boldsymbol{S_a}$ describes the estimated prior errors on the state vector (i.e. the basis function weights). In principle these are defined to represent realistic prior knowledge in the profile, though tuning to match the information content of the joint retrieval is likely to be needed in practise. The prior state itself, $\boldsymbol{x}_a$ , is typically a vector of 0s, since the state describes increments to the offset profile, $\boldsymbol{o}$.

Since the forward model is linear, the solution which minimises the cost function is given by

$$\hat{\boldsymbol{x}} = \boldsymbol{x_a} + \left(\boldsymbol{K^T S_e^{-1} K} + \boldsymbol{S_a^{-1}}\right)^{-1} \boldsymbol{K}^T \boldsymbol{S_e^{-1}} \left(\boldsymbol{y} - F(\boldsymbol{x_a})\right)$$

**Equation 3-10**

where weighting function matrix $\boldsymbol{K}$ is the derivative of the forward model with respect to the state parameters, i.e.

$$\boldsymbol{K} = \boldsymbol{A}\,\boldsymbol{B}$$

**Equation 3-11**

Where **A** is the matrix containing the averaging kernels of the input retrieved amounts with respect to the fine scale profile (columns are $\boldsymbol{A}_j$ in Equation 3-8). Note that (assuming linearity), using the averaging kernel equation as the forward model operator, effectively removes the influence of the original retrievals prior constraint on the joint retrieval (this is effectively replaced by the new prior).

Given the solution state, it is trivial to compute the corresponding high-resolution version of the profile, $\boldsymbol{r}$ (Equation 3-6). Sub-column amounts for specific layers can then be calculated from that profile. This step can be carried out by matrix multiplication by matrix, $\boldsymbol{M}$, which contains the weights needed to integrated the profile to a set of sub-column amounts, such that

$$\boldsymbol{s = M\,r}$$

**Equation 3-12**

Via the usual OEM equations, the total a posteriori errors on $\hat{\boldsymbol{x}}$ are described by covariance

$$\boldsymbol{S_x = (\,K^T S_e^{-1} K + S_a^{-1}\,)^{-1}}$$

**Equation 3-13**

The (non-square) averaging kernel for $\hat{\boldsymbol{x}}$, giving derivatives of the solution with respect to *fine-scale* perturbations in the profile is given by

$$\boldsymbol{A_x = S_x K^T S_e^{-1} A}$$

**Equation 3-14**

($\boldsymbol{A}$, without subscript, is the derivative of the input retrievals with respect to fine-scale perturbations in the true profile).

The (square) averaging kernel for the output fine scale profile is:

$$\boldsymbol{A_r = B A_x}$$

**Equation 3-15**

The (non-square) averaging kernel for the derived sub-columns (with respect to fine scale perturbations in the profile) is

$$\boldsymbol{A_s = M A_r}$$

**Equation 3-16**

rrors on the retrieval profile and sub-columns are described by covariances:

$$\boldsymbol{S_r = B S_x B^T}$$

**Equation 3-17**

$$\boldsymbol{S_s = (MB) S_x (MB)^T}$$

**Equation 3-18**

(The noise and smoothing error covariances can be similarly derived starting from the usual OEM expression for these matrices for the state vector $\hat{\boldsymbol{x}}$**).**

# 4 Limb profile ECV Retrieval / Merging Algorithm

## 4.1 HARMonized dataset of OZone profiles (HARMOZ) (Bremen)

The concept of the HARMonized dataset of OZone profiles (HARMOZ) is based on limb and occultation measurements has been developed in the Ozone_CCI project (Sofieva et al., 2013). HARMOZ consists of original retrieved ozone profiles from each satellite instrument, which are screened for invalid data by the instrument teams. While the original ozone profiles are presented in different units and on different vertical grids, the harmonized dataset is given on a common vertical grid in netcdf-4 format. The vertical range of the ozone profiles is specific for each instrument, thus all information contained in the original data is preserved. Provided altitude and temperature profiles allow the representation of ozone profiles in number density or mixing ratio on a pressure or altitude vertical grids. Geolocation, uncertainty estimates and vertical resolution are provided for each profile. For each instrument, optional parameters, which are related to the data quality, are also included.

In the CCI project, two versions of the HARMOZ datasets are developed and created: altitude-gridded (HARMOZ_ALT) and pressure-gridded (HARMOZ_PRS) ozone concentration datasets. The vertical sampling of HARMOZ_ALT profiles is 1 km. For HARMOZ_PRS, the pressure grid corresponds to vertical sampling of ~1 km below 20 km and 2-3 km above 20 km. The information about the available HARMOZ datasets is collected in Table 4-1. The datasets, which are developed or extended in CCI + project, are highlighted.

**Table 4-1:** Information about the HARMOZ_ALT and HARMOZ_PRS dataset

| Instrument/ satellite | Level 2 processor | Years | Vertical range | Retrieval vertical coordinate |
|---|---|---|---|---|
| MIPAS/Envisat | KIT/IAA V7R_O3_240 | 2002-2012 | 6-70 km/400- 0.05 hPa | altitude |
| SCIAMACHY/Envisat | UBr v3.5 | 2002-2012 | 5-65 km/250 – 0.05 hPa | altitude |
| GOMOS/Envisat | ALGOM2s v1 | 2002-2011 | 10-105 km/250- $10^{-4}$ hPa | altitude |
| GOMOS bright limb/ Envisat | GBL v1.2 | 2002-2011 | 10-59 km/70 – 0.2 hPa | altitude |
| SAGE II/ ERBS | NASA, v7.0 | 1984-2005 | Cloud top-70 km / 300-0.05 hPa | altitude |
| SMR/Odin | CUT, v2.1 | 2001-2014 | 18-60 / 100-0.1 hPa | altitude |
| HALOE | NASA, v19 | 1991-2005 | 15-80 km / 300-0.01 hPa | pressure |
| POAM III/SPOT 4 | NASA, v4 | 1998-2005 | 6-55 km /500 - 0.1 hPa | altitude |
| SAGE III Meteor 3M | NASA, AO3 v4 | 2002-2005 | 6-60 km /500 - 0.05 hPa | altitude |
| OSIRIS/Odin | USask v5.10 | 2001-2021 | 10-59 km/450 - 0.1 hPa | altitude |
| ACE-FTS/SCISAT | UoT v3.5/3.6 | 2004-2021 | 10-94 km/450 - 2·$10^{-4}$ hPa | altitude |
| OMPS-LP/Suomi-NPP | USask 2D v1.1.0 | 2012-2021 | 6-59 km/500 - 0.1 hPa | altitude |

| | | | | |
|---|---|---|---|---|
| MLS/Aura | NASA_JPL v4.2 | 2004-2021 | 6-75 km /500 - 0.02 hPa | pressure |
| SABER/Timed | NASA_GATS v2 | 2002-2021 | 12-105 km/400- $10^{-4}$ hPa | pressure |
| SAGE III ISS | NASA, AO3 v5.2 | 2017-2021 | 6-60 km /500 - 0.05 hPa | altitude |

The dataset is available at https://climate.esa.int/en/projects/ozone/data and ftp://cci_web@ftp-ae.oma.be/esacci

## *4.2 OMPS Retrieval Schemes (Bremen)*

### 4.2.1 OMPS-LP NASA retrieval algorithm version 2.5

The NASA Environmental Data Record algorithm to retrieve ozone profiles from OMPS-LP measurements is based on the optimal estimation approach. Several versions of the retrieval algorithm were released starting from April 2012; the most recent version 2.5 is described in Kramarova et al. (2018). In this version, to improve the stability of the retrieval, a correlation radius of 5 km was introduced in the a priori covariance matrix instead of the previously employed Tikhonov parameter. To simulate limb radiance, the Gauss-Siedel radiative transfer model is used and the Bass and Paur ozone cross section selected.

The algorithm is designed to retrieve ozone independently from two spectral ranges: the UV region between 28.5 and 52.5 km and the Chappuis band between 12.5 (or cloud top height) and 37.5 km. The doublet and triplet method is used to obtain the measurement vector, respectively, for the Hartley-Huggins and Chappuis bands. Wavelengths characterized by a strong ozone absorption are paired with weak absorbing ones. The normalization of the radiance is performed using an upper tangent height measurement: 55.5 km in the UV and 40.5 km in the Vis. An additional TH correction is applied by NASA on L1 gridded data, as described in [RD-27]. A new algorithm was developed to detect the cloud top height [RD-14] and accordingly filter out underneath THs. In addition, aerosol extinction profiles are currently derived for each measurement, replacing the climatology profiles used in the previous versions, as described in [RD-30].

Only the central slit of the instrument is considered and currently provided in L2 NASA product, because of the uncertainties in the pointing and stray light issues affecting the two side slits.

### 4.2.2 OMPS-LP IUP retrieval algorithm version 2

The UBR-IUP retrieval retrieval implements a first-order Tikhonov regularization within the SCIATRAN radiative transfer model. As described in [RD-59], five spectral segments are selected: three in the Hartley and Huggins bands and two in the visible spectral range. We have to take into account the presence of water vapor and $O_2$ absorption features in the Chappuis band used for the lowest altitude range (*) so that wavelengths in the intervals 585-605 nm and 620-635 nm are rejected. The treatment of these absorption features requires line-by-line calculations, which are computationally expensive to be implemented for the whole time series.

Considering the decreasing sensitivity above 55 km and the saturation of limb signal in the lower stratosphere, the retrieval of ozone profiles is performed over the altitude range between 12 and 60 km, with the lower boundary that can be higher in the presence of a cloud. An evenly spaced vertical grid spans this vertical range with steps every 1 km.

The measurement vector consists of the logarithms of the normalized limb radiances. In detail, OMPS-LP spectrum in the five spectral segments at each altitude is normalized by a limb measurement at an upper TH. This provides a self-calibration of the instrument, by removing the need for solar irradiance measurements, and reduces the effects of surface/cloud reflectance uncertainties. Table 4-2 lists the details about spectral segments and the used normalization altitudes.

For each TH a polynomial is subtracted from the logarithm of the normalized radiance to remove slowly variable spectral features, for example, related to Rayleigh or aerosol scattering. The last column of Table 4.2 provides information about the order of the polynomial subtracted: zeroth order or no polynomial in the UV region and first order for the Chappuis band.

| TH [km] | Spectral segment [nm] | Normalization TH [km] | Poly. order |
|---|---|---|---|
| 48-60 | 290-302 | 62.5 | - |
| 34-49 | 305-313 | 51.5 | - |
| 28-39 | 321-330 | 51.5 | 0 |
| 16-31 | 508-660 | 42.5 | 1 |
| 8-16 (*) | 508-670 | 42.5 | 1 |

**Table 4-2**: ***List of the spectral segments considered for the ozone retrieval with corresponding TH ranges, altitudes used for the normalization and order of the subtracted polynomial.***

As $O_3$ , $NO_2$ and $O_4$ have relevant spectral signatures in the selected spectral ranges, the radiation field in the forward model is calculated, taking into account these three gases. The respective cross-sections are taken from [RD-41], [RD-10] and [RD-21] and are beforehand convolved to the OMPS-LP spectral resolution.

Before the main retrieval, a spectral correction is applied in the Chappuis band to take into account issues related to the spectral calibration and a possible thermal expansion of the sensor. This correction consists in a shift and squeeze of the modeled spectrum relative to the measured one. This pre-processing procedure is applied independently for each observation at each TH between 12 and 31 km. The differential absorption structure in the Huggins band is mostly not resolved in OMPS-LP spectra, due to the relatively low spectral resolution of the sensor. As a consequence, the UV retrieval uses either normalized radiances or their slopes and the influence of a possible spectral misalignment is not expected to be significant; thus, the shift and squeeze algorithm is not applied. Typical values of the spectral shift are inside the range [+1,+4] nm for the first point of the interval and [-2,+1] nm for the last spectral point.

For all spectral ranges, the measurement noise covariance matrix is obtained at this stage from the fit residuals, after all the relevant gases in the selected spectral windows have been fitted. The noise to signal ratio is taken as the root mean square of the fit residuals and fed into SCIATRAN.

Surface albedo is retrieved simultaneously, by assuming a Lambertian surface and using the sun-normalized radiance. For the albedo retrieval, we selected two spectral ranges at TH around 38 km: 350-365 nm and 445-455 nm, where ozone absorption is weak.

### 4.2.3 OMPS-LP Usask retrieval algorithm

The USask implemented a 2-D or tomographic retrieval scheme to produce ozone profiles from OMPS-LP observations [RD-58]. The 2-D retrieval algorithm accounts for variations along the line-of-sight dimension, combining all measurements from the sunlit portion of the orbit and simultaneously fitting the ozone profile for the portion of the orbit with solar zenith angle less than 88˚. Ozone is retrieved from the thermal tropopause to 59 km on a 1 km grid and with a vertical resolution of approximately 2 km. The forward model used is SASKTRAN-HR and the regularization technique implemented consists in a Tikhonov second derivative constraint applied to the horizontal direction of the retrieval grid. The doublet/triplet technique is adopted to retrieve ozone, using the wavelengths as reported in Tab 4-3 together with the corresponding altitude range and normalization altitude.

**Table 3-3: Wavelength triplet/doublets used in the ozone retrieval**

| Ozone-sensitive wavelength [nm] | Reference wavelengths [nm] | Valid altitudes [km] | Normalization altitude [km] |
|---|---|---|---|
| 292.43 | 350.31 | 22-59 | 60 |
| 302.17 | 350.31 | 22-55 | 56 |
| 306.06 | 350.31 | 22-51 | 52 |
| 310.70 | 350.31 | 22-48 | 49 |
| 315.82 | 350.31 | 22-46 | 47 |
| 322.0 | 350.31 | 22-42 | 43 |
| 331.09 | 350.31 | 22-39 | 40 |
| 602.39 | 543.84, 678.85 | 0-30 | 31 |

The distribution of stratospheric aerosol is also retrieved using the 745 nm spectral point, which is also used to estimate the Lambertian surface reflectance. These two parameters are considered to be a second-order correction for the ozone retrieval and are derived independently in the following order: surface reflectance, aerosol and ozone. Two passes of this procedure are performed, allowing results from the ozone retrieval to couple back into the other retrievals and a fixed number of iterations is performed in each of the passes [RD-58]. In order to assure accurate pointing

corrections, a separate pointing analysis was conducted, by implementing the Rayleigh Scattering Attitude Sensor (RSAS) technique.

## 4.3 SAGE-CCI-OMPS Extension (FMI)

The merged monthly zonal mean dataset of ozone profiles, which is also referred to as the SAGE-CCI-OMPS dataset, is created using the data from several satellite instruments: SAGE II on ERBS, GOMOS, SCIAMACHY and MIPAS on Envisat, OSIRIS on Odin, ACE-FTS on SCISAT, and OMPS on Suomi-NPP. This dataset has been created in the framework of Ozone_cci project [RD-46]. Originally, the datasets covered the time period from 1984 to 2016, but now it is regularly extended.

The stability of the individual-instrument data records has been extensively studied; only stable data are used for the merged dataset.

All the data used for creating the merged dataset have a sufficiently good resolution of 2–3 km in the UTLS. For all instruments, ozone profiles are retrieved on the geometric altitude grid. The majority of the datasets -SAGE II, GOMOS, OSIRIS, SCIAMACHY and OMPS - provide number density ozone profiles; therefore this representation is used for the merged dataset. For ACE-FTS and MIPAS, the retrievals are in volume mixing ratio on altitude grid. Conversion to number density profiles is performed using temperature profiles retrieved by these instruments, thus providing consistent (i.e., without using external information about temperature and pressure profiles) representation of number density ozone profiles.

The information about individual datasets is collected in Table 4.3.1. For some instruments, the selected time period is shorter than the full operation period. The individual datasets have been compared with each other and with ground-based data, and only the time periods when the instruments were operating the best are selected.

**Table 4-4: Information about the datasets used in the merged dataset (ozone profiles from limb and occultation sensors).**

| Instrument/ satellite | Processor, data source | Time period | Local time | Vertical resolution | Estimated precision | Profiles per day |
|---|---|---|---|---|---|---|
| SAGE II/ ERBS | NASA V7.0, original files | Oct 1984 – Aug 2005 | sunrise, sunset | ~1-2 km | 0.5-5% | 14-30 |

| | | | | | | |
|---|---|---|---|---|---|---|
| OSIRIS/ Odin | USask v 5.10, HARMOZ_ALT | Nov 2011 – present | 6 a.m., 6 p.m. | 2-3 km | 2-10% | ~250 |
| GOMOS/ Envisat | ALGOM2s v 1.0, HARMOZ_ALT | Aug 2002 – Aug 2011 | 10 p.m. | 2-3 km | 0.5–5 % | ~110 |
| MIPAS/ Envisat | KIT/IAA v.7, HARMOZ_ALT | Jan 2005 – Apr 2012 | 10 p.m., 10 a.m. | 3-5 km | 1–4% | ~1000 |
| SCIAMACHY/ Envisat | UBr v3.5, HARMOZ_ALT | Aug 2003- Mar 2012 | 10 a.m. | 3-4 km | 1-7% | ~1300 |
| ACE-FTS/ SCISAT | V3.5/3.6, HARMOZ_ALT | Feb 2004 – present | sunrise, sunset | ~3 km | 1-3% | 14-30 |
| OMPS/ Suomi NPP | USask 2D, HARMOZ_ALT | Apr 2012- present | 1:30 p.m. | ~2 km | 2-10% | ~1600 |

The merged dataset is created in 10° latitude zones from 90°S to 90°N, in the altitude range 10 – 50 km.

The merging is performed on the deseasonalized anomalies computed from each individual dataset. The details of the merging procedure can be found in [RD-46].

The main dataset consists of the merged deseasonalized anomalies and their uncertainties. For the purpose of other applications (e.g., comparisons with models etc.), we presented the ozone profile also in number density. The computing of merged number density profiles from the merged deseasonalized anomalies is performed via restoring the seasonal cycle in the data. The best estimate of the amplitude of seasonal cycle is given by MIPAS measurements, because they provide all season pole-to-pole measurements with dense sampling. We take the absolute values

of the seasonal cycle from SAGE II and OSIRIS in the overlapping period (which are very close to each other and also with GOMOS measurements), thus preserving the consistency in the dataset through the whole observation period.

The merged SAGE-CCI_OMPS is actively used in ozone trends analyses and in ozone assessments ([RD-56], [RD-57], [RD-55], [RD-37]).

## 4.4 Gridded merged Level 3 dataset (FMI)

The MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP) is significantly developed further in the Ozone CCI+ project. The monthly zonal mean gridded ozone profile dataset is provided in the altitude range from 10 km to 50 km. It covers the time period from late 2001 until now. The data are gridded monthly in the 10° latitude x 20° longitude zones. Since the sampling of solar occultation measurements is rather low, they are not included. The gridded ozone profiles are computed first separately for GOMOS, MIPAS, SCIAMACHY, OSIRIS, OMPS-LP, MLS, and they are also merged into one dataset. The information about the individual datasets can be found Tables 4.1 and 4.3. The principle of creating the Level 3 gridded data for individual datasets, as well as data merging is the same as for the monthly zonal mean dataset, and it is described in detail in [RD-62].

The merging method used for creating MERGRIDOP is similar to that used in creating the merged SAGE-CCI-OMPS dataset [RD-46]. The deseasonalized anomalies of all instruments except OMPS are aligned, as the seasonal cycle was estimated using the same period. First, we offset the OMPS deseasonalized anomalies to the median of the deseasonalized anomalies from all other instruments. These additive offsets are computed using the data from years 2012-2018, and the offsetting procedure is illustrated in Figure 4-1.

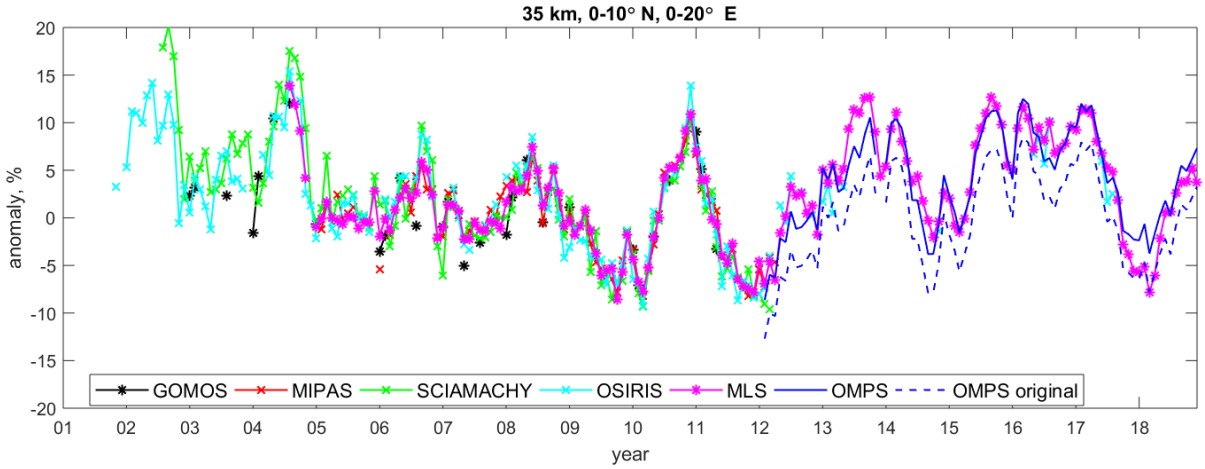


**Figure 4-1: From [RD-62]: Illustration of offsetting the OMPS deseasonalized anomalies. The data are shown for altitude 35 km and the latitude/longitude bin 0-10° N/ 0-20° E.**

After offsetting OMPS, the merged ozone profiles in each spatiotemporal bin and at each altitude level is obtained from the median of the deseasonalized anomalies corresponding to individual instruments:

$$\Delta_{merged}(z,b,t) = median(\Delta_i(z,b,t)) \qquad \textbf{Equation 4-1}$$

The advantage of using the median estimate is that the merged anomaly follows the majority of the data, and it is not very sensitive to exclusion/addition of an individual data record, in cases where there are several (and consistent) anomaly datasets available.

The uncertainties of the merged deseasonalized anomalies are computed similarly to those used for the merged SAGE-CCI-OMPS dataset [RD-46]. For each instrument, the uncertainty of the deseasonalized anomalies, $\sigma_{\Delta i}$, is estimated via Gaussian error propagation; it is given by

$$\sigma_{\Delta i} = \Delta_i \sqrt{\frac{\sigma_i^2}{\rho_i^2} + \frac{\sigma_{m,i}^2}{\rho_{m,i}^2}} \qquad \textbf{Equation 4-2}$$

where $\sigma_i$ is the uncertainty of the gridded ozone profiles (see Sect. 3.1.) and $\sigma_{m,i}$ is the uncertainty of the seasonal cycle $\rho_{m,i}$, which can be estimated via propagation of random uncertainties to the mean value:

$$\sigma_{m,i}^2 = \frac{1}{N_m^2} \sum_{j=1}^{N_m} \sigma_i^2(z,b,t_j) \qquad \textbf{Equation 4-3}$$

Analogously to (Sofieva et al., 2017), the uncertainties of the merged deseasonalized anomalies are estimated as:

$$\sigma_{\Delta,merged} = \min\left( \sigma_{\Delta,j_{med}}, \sqrt{\frac{1}{N} \sum_{j=1}^{N} \sigma_{\Delta,j}^2 + \frac{1}{N^2} \sum_{j=1}^{N} \left(\Delta_j - \Delta_{merged}\right)^2} \right), \qquad \textbf{Equation 4-4}$$

where $\sigma_{\Delta,i_{med}}$ is the anomaly uncertainty of the instrument corresponding to the median value. In cases where there are even number of measurements, the mean of two neighbours to the median is used.

Examples of deseasonalized anomalies and their estimated uncertainties are displayed in Figures Figure 4-2 and Figure 4-3, respectively.

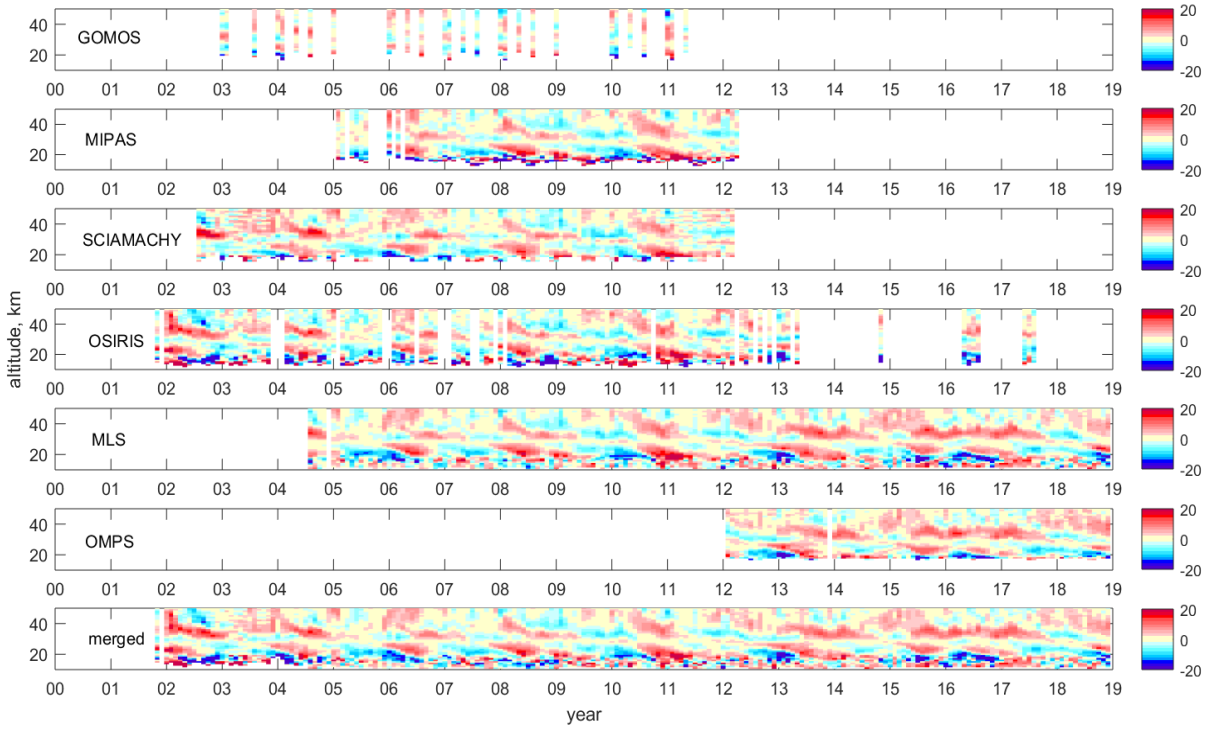


**Figure 4-2. An example of deseasonalized anomalies (in %) for individual instruments and the merged dataset in the spatial bin 0-10° N, 0-20° E.**

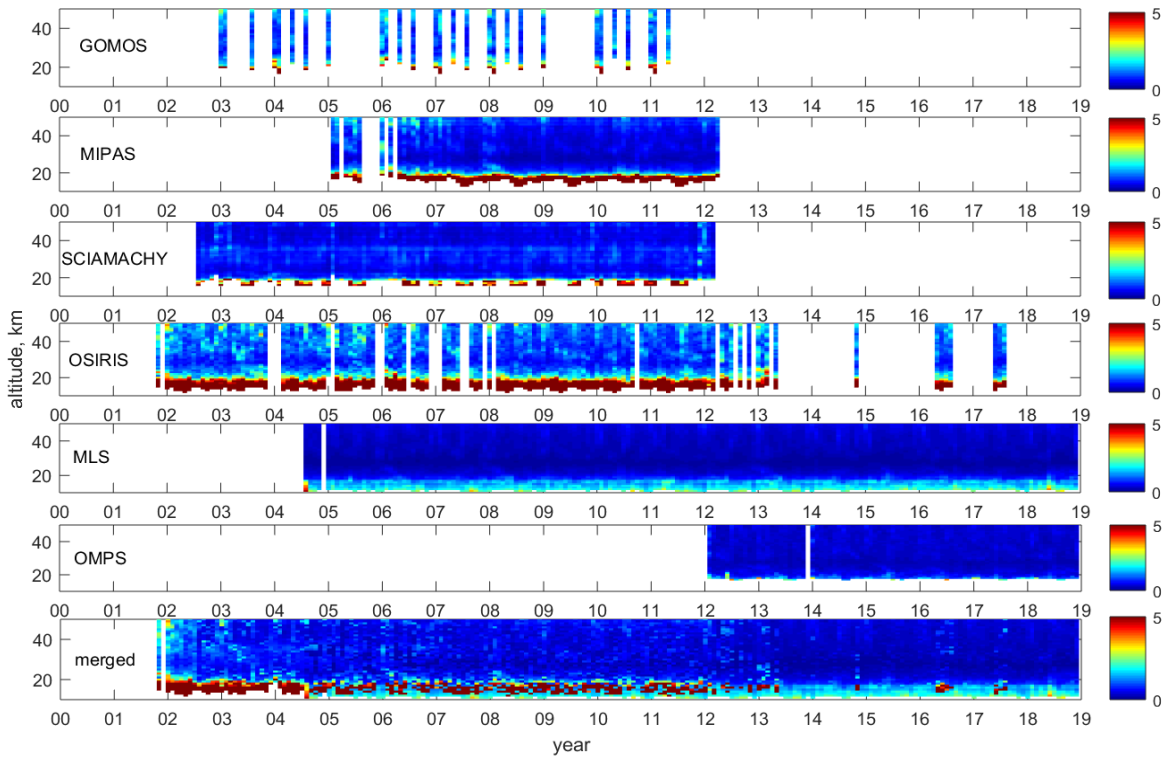


**Figure 4-3. From Sofieva et al. (2021): An example of uncertainties in deseasonalized anomalies (in %) for individual instruments and the merged dataset in the spatial bin 0-10° N, 0-20° E**

The merged deseasonalized anomalies can be used directly for evaluation of ozone trends. For other applications, the MEGRIDOP is also available in the form of ozone number density profiles. The dataset is available through open access at https://climate.esa.int/en/projects/ozone/data/ and at ftp://cci_web@ftp-ae.oma.be/esacci.

The MEGRIDOP dataset has been used for evaluation of regional trends of stratospheric ozone profiles [RD-62].