ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product Quality Assurance Document (PQAD) Ozone products

Version 2.1

Issued by: BIRA-IASB / Jean-Christopher Lambert

Date: 29/04/2021

Ref: C3S_D312b_Lot2.2.1.1_202102_PQAD_O3_v2.1

Official reference number service contract: 2018/C3S_312b_Lot2_DLR/SC1

Contributors

ROYAL BELGIAN INSTITUTE FOR SPACE AERONOMY (BIRA-IASB)
Daan Hubert
Arno Keppens
Tijl Verhoelst
José Granville
Jean-Christopher Lambert

Document history

Version	Date	Description of modification	Chapters / Sections
1.1	28/02/2019	C3S_312a_Lot4 Ozone PQAD (v2.0) updated	All
1.1a	16/04/2019	Incorporation of review comments	Table RD, 4.4
1.2	28/02/2020	Updated related documents, overviews available data records & validation results	Sects. 2, 3 & 6
2.0	28/02/2021	Updated related documents, overviews available data records, methodology & validation results	All
2.1	29/04/2021	Incorporation of review comments	All

Related documents

Reference ID	Document
RD1	C3S_312b_Lot2 – Atmospheric composition Final project proposal (Annex 2 to the Framework Agreement), 07/05/2018
RD2	C3S_312b_Lot1 – Target Requirements and Gap Analysis Document – Ozone ECV (TRGAD) v3.11, 09/12/2020, C3S_312b_Lot2.1.0_2020(O3)_TRD-GAD_v3.11
RD3	C3S_312b_Lot2 – Ozone Algorithm Theoretical Basis Document (ATBD) v2.1, 25/03/2021, C3S_312b_Lot2.1.1.2_202102_ATBD_v2.1
RD4	C3S_312b_Lot2 – Ozone Product User Guide and Specification (PUGS) v2.1, 25/03/2021, C3S_312b_Lot2.3.2.1_202102_PUGS_O3_v2.1
RD5	C3S_312b_Lot2 – Product Quality Assessment Report (PQAR) v1.2b, 22/06/2020, C3S_D312b_Lot2.2.1.2_202005_PQAR_O3_v1.2b
RD6	C3S_312b_Lot2 – System Quality Assurance Document : Atmospheric Composition ECVs – Annex A : Ozone Products (SQAD) v1.0, issue 5, 29/03/2021, C3S_312b_Lot2.2.0-v5_SQAD_202103_Annex_A_Ozone_v1.0
RD7	Ozone_cci – User Requirements Document (URD) v3.1, 01/09/2020, Ozone_cci_URD_3.1 (https://climate.esa.int/en/projects/ozone/key-documents)
RD8	Ozone_cci – Product Validation and Intercomparison Report (PVIR) v2.0, 30/06/2016, Ozone_cci_Phase-II_PVIR_2.0 (https://climate.esa.int/en/projects/ozone/key-documents)
RD9	van Weele, M., Müller, R., Riese, M., Engelen, R., Parrington, M., Peuch, V.-H., Weber, M., Rozanov, A., Kerridge, B., Waterfall, A., and Reburn, J., 2015, User requirements for monitoring the evolution of stratospheric ozone at high vertical resolution, 'Operoz', Operational ozone observations using limb geometry, ESA report, Expro Contract: 4000112948/14/NL/JK
RD10	GCOS, 2016, The Global Observing System for Climate: Implementation Needs. GCOS-200, WMO, Geneva, October 2016

Acronyms

Acronym	Definition
ACE-FTS	Atmospheric Chemistry Experiment – Fourier Transform Spectrometer
ASOPOS	Assessment of Standard Operating Procedures for OzoneSondes
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
C3S	Copernicus Climate Change Service (EU)
CCI	Climate Change Initiative
CDR	Climate Data Record
CEOS	Committee on Earth Observation Satellites
CMM	Cell Monthly Mean data
CMMa	Cell Monthly Mean anomaly data
DLR	Deutsche Luft- und Raumfahrt
DS	Direct Sun
DU	Dobson Unit
ECMWF	European Centre for Medium-range Weather Forecasts
ECV	Essential Climate Variable
FORLI	Fast Optimal Retrievals on Layers for IASI
GAW	Global Atmosphere Watch
GCOS	Global Climate Observing System
GOME	Global Ozone Monitoring Experiment
GOMOS	Global Ozone Monitoring by Occultation of Stars
GPS	Global Positioning System
GTO	GOME-type Total Ozone
HALOE	HALogen Occultation Experiment
IASI	Infrared Atmospheric Sounding Interferometer
ICDR	Intermediate Climate Data Record
ISO	International Organization for Standardization
KNMI	Koninklijk Nederlands Meteorologisch Instituut
KPI	Key Performance Indicator
LMZ	Limb Monthly Zonal profile average
LP	Limb Profile
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder
MSR	Multi-Sensor Reanalysis
Multi-TASTE	Versatile multi-satellite validation system
NDACC	Network for the Detection of Atmospheric Composition Change
NP	Nadir Profile
OMI	Ozone Monitoring Instrument
OSIRIS	Optical Spectrograph and InfraRed Imager System
OSSSMOSE	Observing System of Systems Simulator for Multi-Observations Synergy Exploration
PQAD	Product Quality Assurance Document

Acronym	Definition
PQAR	Product Quality Assessment Report
PVIR	Product Validation and Intercomparison Report
RAL	Rutherford Appleton Laboratory
SABER	Sounding of the Atmosphere using Broadband Emission Radiometry
SAGE	Stratospheric Aerosol and Gas Experiment
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SHADOZ	Southern Hemisphere ADditional Ozonesondes
SMM	Station Monthly Mean data
SMMa	Station Monthly Mean anomaly data
SMR	Sub-Millimeter Radiometer
SSC	Station Seasonal Cycle data
SZA	Solar Zenith Angle
T(O)C	Total (Ozone) Column
TRGAD	Target Requirements and Gap Analysis Document
URD	User Requirement Document
UTLS	Upper Troposphere to Lower Stratosphere
VMR	Volume Mixing Ratio
WMO	World Meteorological Organization
WOUDC	World Ozone and Ultraviolet Radiation Data Centre

Table of Contents

Executive summary / Scope of the document

The C3S_312b_Lot2 service procures several atmospheric ozone climate data records (CDR) for the Climate Data Store (CDS, https://cds.climate.copernicus.eu) of EU's Copernicus Climate Change Service (C3S). It performs the production, quality assurance, and delivery to the CDS of gridded (level-3) and assimilated (level-4) satellite measurements of the total column, the tropospheric column and the vertical profile of atmospheric ozone, acquired by current and historical satellite sounders. This document describes the methodology adopted in the C3S_312b_Lot2 procurement service for the quality assurance of these ozone CDRs, with details on the ground-based measurements used as a reference for validation, the specific technical solutions implemented to enable meaningful level-3 and level-4 data comparisons, and the metrics developed to link validation results to the user requirements (assessment of compliance, or of fitness-for-purpose). The results of the validation analyses and the assessment of the quality of the different C3S data products are reported in the Product Quality Assessment Report (RD5).

The different sections describe:

1) The general validation principles;
2) The products to be validated;
3) The validation (reference) data sets;
4) The details of the methodology;
5) The linkage between validation results and user requirements.

Level-3 and level-4 data products are the result of a complex processing chain, starting from raw (uncalibrated) satellite measurements which are subsequently calibrated and geolocated (level-1b data), and then analysed spectrally to retrieve the geophysical quantities (level-2 data) that finally are averaged over grid cells (level-3 data) or ingested in a data assimilation system coupling chemical-transport modelling with observations (level-4 data). A prerequisite of level-3 and level-4 quality assurance is the existence of an appropriate quality assurance process of the underlying level-1 and level-2 data (see Section 1). The validation of lower-level products being outside of the scope of the C3S_312b_Lot2 service, the quality assurance undertaken here builds upon the validation work carried out on the level-2 data products within ESA's Climate Change Initiative – Ozone project (RD7) up to the end of its Phase 2 (December 2017). Validation of newly acquired level-2 data products continue as part of CCI+ projects that kicked off in March 2019.

While there exists a separate System Quality Assurance Document (RD6), the successful processing of the data files by the validation system described here also constitutes a verification of the file formats and product coverage.

Tables

Figures

1. General validation principles

The Committee on Earth Observation Satellites (CEOS) and the International Organization for Standardization (ISO) define validation as the process of assessing, by independent means, the quality of the data products derived from the system outputs. The validation of an atmospheric ozone data product aims at verifying that the data produced respond to predefined quality requirements (fitness for purpose of the data). Validation generally involves the assessment of the closeness of the data to the geophysical reality, and of its sources of uncertainty, over the spatial and temporal domains of relevance. Uncertainty estimates can include, but are not restricted to, estimates of the bias and precision of the data with respect to reference data, and identification of the temporal and spatial domains over which those estimates remain valid.

Fundamental is therefore the comparison with reference measurements representing the atmospheric “truth”. A key aspect of any comparison for validation purposes is thus the selection of the reference data sets. The quality, traceability and suitability of the latter are essential to allow proper, unbiased and independent validation. Reference measurements must be well documented and procedures must exist to ensure adequate quality control on the long term, as it is the case e.g. within international ground-based networks. This is discussed in detail in Section 3.

Level-3 and level-4 atmospheric ECV data products are the result of a complex processing chain, starting from raw (uncalibrated) satellite measurements which are subsequently calibrated and geolocated (level-1 data), enabling the retrieval of geophysical quantities (level-2 data) that can be aggregated, averaged (level-3 data) or assimilated (level-4 data) into a gridded product. Each step corresponds to a particular set of challenges and potential error sources, all of which require detailed quality assurance. For detailed descriptions of the validation needs for ozone products up to level-2, we refer to Balis et al. (2007) and Verhoelst et al. (2015) for total ozone column data, to Keppens et al. (2015, 2019) and Hubert et al. (2016) for nadir and limb ozone profile data. The protocols described therein include analyses essential to ensure the product quality at level-2 but not all are applicable to level-3 or level-4 data; e.g. the assessment of the dependence on influence quantities such as solar zenith angle, cloud and/or surface properties become irrelevant for gridded data. An example of tying level-3 to level-2 validation results can be found in Coldewey-Egbers et al. (2015). Since no lower-level validation is foreseen within C3S_312b_Lot2 – Atmospheric composition (RD1), the validation performed here builds heavily upon the validation of the underlying level-2 products carried out within ESA’s Climate Change Initiative – Ozone project (CCI+ Phase I, and earlier).

A significant challenge in validating level-3 or level-4 satellite products is to deal with potential differences in spatial and temporal representativeness between the (assimilated) satellite product and the ground-based reference measurements. A crucial part of the validation methodology is therefore the construction of a ground-based level-3 product mimicking the satellite product, and the co-location between satellite and ground-based data. These are key topics in Section 4, which furthermore describes the quality indicators derived from the comparison of level-3 data.

A final topic that needs to be elaborated is the translation between the metrics that can be derived from the validation exercise and the quantified user requirements compiled to ensure fitness-for-purpose of the product. This is addressed in Section 5.

2. Validated products

2.1 Total ozone column products

In Table 1, the main characteristics of the total ozone column products to be validated are summarized, as derived from the product data files delivered in February 2021. Note that for some products, these represent advances w.r.t. the characteristics described at the onset of the project. For instance, the individual sensor products are based on a GODFITv4 rather than GODFITv3 processing, and the horizontal resolution of the L4 product (TC_MSR), after January 1979, is better than originally foreseen: 0.5°x0.5° instead of 1.0°x1.0°.

Table 1. Overview of the main characteristics of the total ozone column products to be validated, as of February 2021.

Product tag	Input sensor	Proc. level	Product type	Processors, file format convention, and TC units	Temporal coverage	Spatial resolution (lat x lon)	Uncertainty information	Provision and provenance
TC_GOME	GOME	3	CDR	GODFITv4 (level-2) DLRv1 (level-3) NETCDF-CF 1.5 mol/m^2	06/1995-07/2011	1°x1°	Standard error and standard deviation	BIRA/DLR
TC_SCIA	SCIAMACHY	3	CDR		08/2002-04/2012	1°x1°		BIRA/DLR
TC_GOME2A	GOME-2A	3	ICDR		01/2007-10/2020	1°x1°		BIRA/DLR
TC_GOME2B	GOME-2B	3	ICDR		01/2013-10/2020	1°x1°		BIRA/DLR
TC_OMI	OMI	3	ICDR		10/2004-10/2020	1°x1°		BIRA/DLR
TC_OMPS	OMPS	3	ICDR		01/2012-10/2020	1°x1°		BIRA/DLR
TC_GTO-ECV	GOME, SCIA, GOME-2A/B, OMI	3	ICDR		07/1995-10/2020	1°x1°		BIRA/DLR

Product tag	Input sensor	Proc. level	Product type	Processors, file format convention, and TC units	Temporal coverage	Spatial resolution (lat x lon)	Uncertainty information	Provision and provenance
TC_IASI-A	IASI-A	3	ICDR	V0001, FORLI v20151001 NETCDF-CF 1.6 mol/m^2	10/2007-01/2021	1°x1°	Standard error	ULB/LATMOS
TC_IASI-B	IASI-B	3	ICDR		05/2013-01/2021			
TC_MSR	(1)	4	ICDR	TM3-DAM v3.3 NETCDF-CF 1.4 Dobson Units	01/1979-12/2020	Up to 01/1979: 1.0°x1.5° From 01/1979: 0.5°x0.5°	Standard deviation	KNMI

(1) Merged/assimilated product based on GOME, SCIAMACHY, OMI, GOME-2A/B, BUV-Nimbus4, TOMS-Nimbus7, TOMS-EP and SBUV-7, -9, -11, -14, -16, -17, -18, -19

2.2 Nadir ozone profile and tropospheric ozone column products

The C3S nadir ozone profile level-3 data as summarized in Table 2 consist of monthly averages on a 1x1 degree latitude-longitude grid. The nadir products, produced by KNMI from RAL's level-2 satellite retrievals (GOME, GOME-2A, SCIAMACHY, and OMI level-2 data), contain 19 layers between 20 fixed pressure levels at each grid-point. For the IASI data, only a level-3 tropospheric ozone column is made available, generated by ULB/LATMOS from its own FORLI v20151001 level-2 ozone profile retrievals. The KNMI and IASI level-3 algorithms are described in the C3S ATBD (RD3).

Table 2. Overview of the main characteristics of the nadir ozone profile and tropospheric column products to be validated, as of February 2021.

Product tag	Level-2 version	Level-3 version	Temporal coverage
NP_GOME	RAL v 3.01	KNMI v0006	06/1995-06/2011
NP_SCIAMACHY	RAL v 3.00	KNMI v0006	08/2002-04/2012
NP_GOME2A	RAL v 3.00	KNMI v0006	01/2007-08/2019
NP_GOME2B	RAL v 3.02	KNMI v0006	04/2013- 08/2019

Product tag	Level-2 version	Level-3 version	Temporal coverage
NP_OMI	RAL v2.15	KNMI v0006	10/2004-08/2019
06TC_IASI-A	FORLI v20151001	V0001	10/2007-01/2021
06TC_IASI-B	FORLI v20151001	V0001	05/2013-01/2021

2.3 Limb ozone profile products

Table 3. Overview of the main characteristics of the limb profile products to be validated, as of February 2021.

Product tag	Sensor	Data version		Space-time binning	Vertical range	Period	Representation† profile (native)	Comments
		Level-2	Level-3					
LMZ_SAGE2	SAGE II	LaRC v7.0	FMI v1	month x 10° lat	5-65 km	1984-2005	alt, ndens	
LMZ_HALOE	HALOE	GATS v19			0.05-500 hPa	1991-2005	pres, ndens (VMR)	
LMZ_SMR	SMR	UChalm v2.1			10-85 km	2001-2014	alt, ndens (VMR)	501.8 GHz data
LMZ_OSIRIS	OSIRIS	USask v5.10			10-59 km	2001-2020	alt, ndens	
LMZ_SABER	SABER	GATS v2.0			1-500 hPa	2002-2020	pres, ndens (VMR)	9.6 μm data
LMZ_GOMOS	GOMOS	FMI ALGOM2s v1			10-105 km	2002-2011	alt, ndens	Occultation data
LMZ_MIPAS	MIPAS	IMK/IAA V7			6-70 km	2002-2012	alt, ndens (VMR)	NOM mode data
LMZ_SCIA	SCIAMACHY	UBr v3.5			5-65 km	2002-2012	alt, ndens	
LMZ_ACE	ACE-FTS	UoT v3.6			6-94 km	2004-2020	alt, ndens (VMR)	
LMZ_MLS	Aura MLS	JPL v4.2			0.02-500 hPa	2004-2020	pres, ndens (VMR)	
LMZ_OMPS	OMPS-LP	USask-2D v1.1.0			6-59 km	2012-2020	alt, ndens	
LMZ_MERGED	MIPAS, GOMOS, SCIAMACHY, ACE, OSIRIS, SAGE II and OMPS-LP.	As above.			10-50 km	1984-2020	alt, ndens	

Product tag	Sensor	Data version		Space-time binning	Vertical range	Period	Representation† profile (native)	Comments
		Level-2	Level-3					
LP_MERGED	MIPAS, GOMOS, SCIAMACHY and OSIRIS.	As above.		month x 10° lat x 20° lon	10-50 km	2001-2020	alt, ndens	

† The profile representation is defined by vertical coordinate: alt(itude) or pres(sure); and the ozone data unit : number density (ndens) or volume mixing ratio (VMR). The retrieved ozone data unit is given between parentheses in case it deviates from that provided in the data files.

3. Description of validation datasets

3.1 Total ozone column: Dobsons and Brewers

Dobson and Brewer ultraviolet spectrophotometers rely on the method of differential absorption in the Huggins band where ozone exhibits strong absorption features of the ultraviolet part of the solar spectrum. This technique has been described in detail in several reference papers (Kerr et al., 1988) and references therein. The Dobson spectrophotometer measures TOC values with a total uncertainty of 2–3% for solar zenith angles smaller than 75°. Since the International Geophysical Year in 1957, Dobson instruments have been deployed in a worldwide network. The Brewer grating spectrophotometer is in principle similar to the Dobson. However, it has an improved optical design and is fully automated. The ozone column abundance is determined from a combination of five wavelengths between 306 nm and 320 nm. Since the 1980s, Brewer instruments are part of the ground-based network as well. Most Brewers are single monochromators, but a small number of systems are double monochromators with improved stray light performance.

The uncertainty on Direct Sun (DS) total ozone measurements by a well maintained Brewer instrument is about 1% (e.g., Kerr et al., 1988). When Brewer spectrophotometers are regularly calibrated and maintained, the DS TOC records can potentially maintain a stability of 1% over long time intervals (WMO, 2006). Despite similar performance, small differences within ±0.6% on average are introduced between the Brewer and Dobson data because of the use of different wavelengths and a different temperature dependence of the ozone absorption coefficients (Staehelin et al., 2003). The seasonal cycle in atmospheric temperature results in a seasonal variation of the Brewer ozone data, where the contribution of the systematic offset is less than 1% (van Roozendael et al., 1998). Dobson and Brewer instruments might also suffer from long-term drift associated with calibration changes. Additional problems arise at solar elevations lower than 15°, for which diffuse and direct radiation contributions can be of the same order of magnitude. Therefore, we limit the use of measurements by Brewer and Dobson ultraviolet spectrophotometers, to the data acquired up to 80° SZA for Brewer of the MK-III and MK-IV series (double monochromators), and up to 70-75° of SZA for Dobsons and single monochromator Brewers.

The data used for C3S_312b_Lot2 total ozone validation are collected from WOUDC (World Ozone and Ultraviolet radiation Data Centre), hosted by Environment and Climate Change Canada (ECCC) in Toronto, Canada, where they are publicly available (http://www.woudc.org). WOUDC is a World Meteorological Organization (WMO) data centre supporting its Global Atmosphere Watch (GAW). The contributing stations used here are listed in Table 7 in Appendix A and their geographical distribution is visualized in Figure 1.

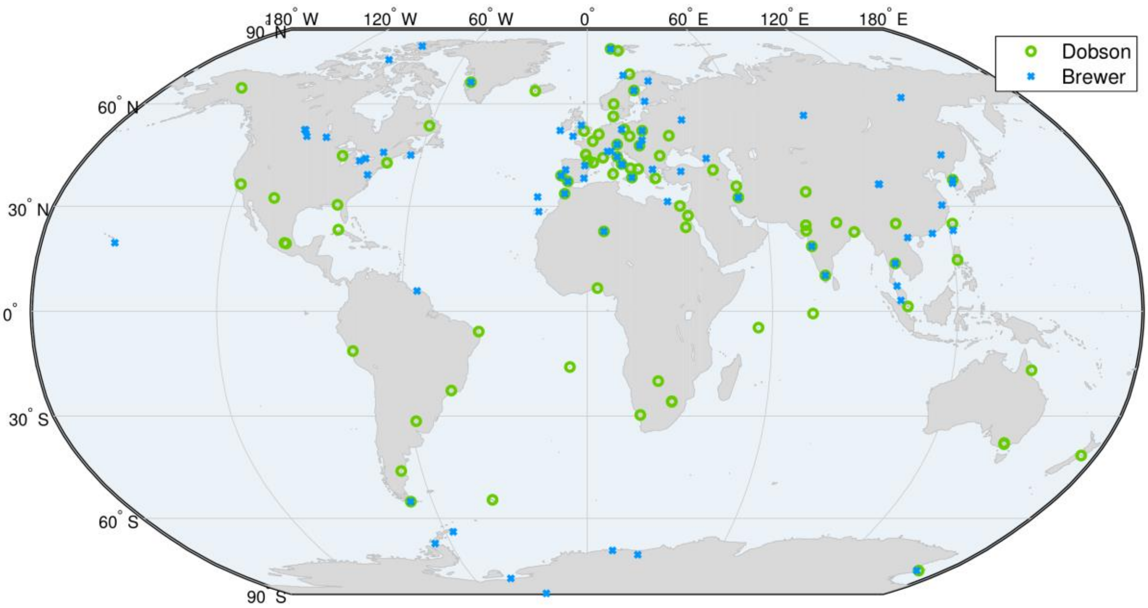

Figure 1. Geographical distribution of the Dobson and Brewer stations used in this study. A detailed listing can be found in Table 7 in Appendix A.

3.2 Vertical ozone profile and tropospheric ozone column: Ozonesondes

In-situ measurements by balloon-borne electrochemical ozonesondes are widely used as a reference for the validation of satellite ozone profile data. The following C3S products are validated using ozonesonde data:

- Gridded tropical tropospheric column data by IASI-A and IASI-B;
- Gridded nadir profile data by GOME, SCIAMACHY, GOME-2A, GOME-2B and OMI;
- Zonal mean limb profile data by MIPAS, GOMOS, SCIAMACHY, SAGE II, HALOE, OSIRIS, SMR, ACE, MLS, SABER, OMPS-LP and a merged data set;
- Latitude-longitude gridded limb profile data for the merged data set.

A specific validation method was optimized for each product family, which relies on a differentiated approach to processing the ground-based ozone profile data obtained from the data archives to a higher level product. The description of these post-processing steps is described in the appropriate subsections of Sect. 4. Here, we describe the measurement principle and the data quality of single-profile measurements by ozonesonde.

Numerous sites around the world launch ozonesonde instruments attached to small meteorological balloons. They measure the vertical profile of ozone partial pressure with 100-150 meter vertical resolution from the ground to the burst point of the balloon, usually between 30 and 35 km. An interfaced radiosonde provides the pressure, temperature and –in recent years– GPS data necessary to geolocate each measurement and to convert the ozone partial pressure to other units such as ozone volume mixing ratio and ozone number density.

Different types of ozonesonde were developed over the years. Those still in use today are mostly based on the electrochemical reaction of ozone with a potassium-iodide sensing solution. Laboratory tests and field campaigns indicate that between the tropopause and about 28 km altitude all sonde types produce consistent results when the standard operating procedures are followed (Smit and the ASOPOS panel, 2014). The bias is smaller than ±5% and the precision is about 3%. Above 28 km the bias increases for all sonde types. Below the tropopause, due to lower ozone concentrations, the precision degrades slightly from 3 to 5%, depending on the sonde type. The tropospheric bias also becomes larger, between ±5 to ±7%. Other factors besides ozonesonde type influence the data quality as well. A detailed overview can be found in Smit and the ASOPOS panel (2014). Recently, Stauffer et al. (2020) reported 5-10% low biases in recent stratospheric ozone measurements (since around 2013) at a number of sites, mainly in the Canadian and SHADOZ networks. The affected sonde data, mainly contributed to the SHADOZ and Canadian networks, will impart a change in the satellite-sonde comparison time series. The timing of the appearance and magnitude of the low bias vary by site. The cause of the issue and a possible correction are under investigation by the sonde community.

The present work relies on the ozonesonde data archived by the Network for the Detection of Atmospheric Composition Change (NDACC), by the Southern Hemisphere Additional Ozonesonde network (SHADOZ) and at the WOUDC archive that receives contributions from WMO's Global Atmosphere Watch (GAW). Together these three data sources collect observations at stations from 82.5°N to 90.0°S, many of which launch at least two to four sondes per month. Stations contributing to the C3S validation studies are shown in Figure 2. Table 8 in Appendix A lists the location of each site, the responsible institute and the archive from which data were taken.

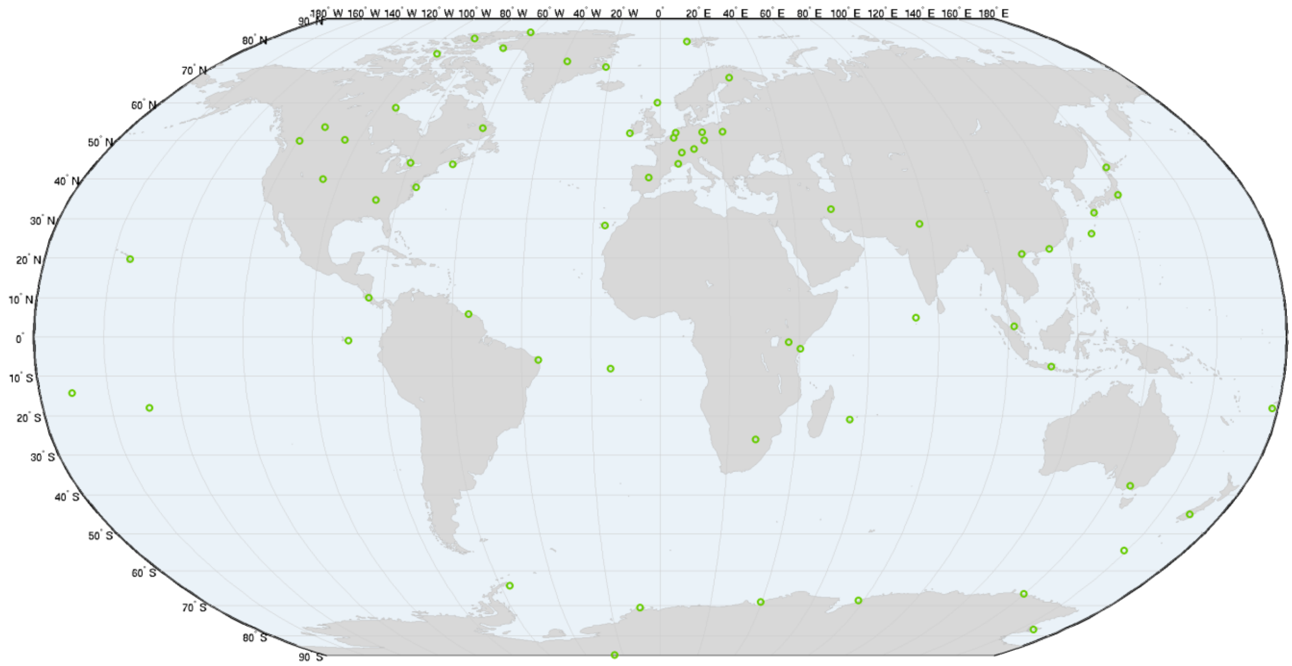

Figure 2. Geographical distribution of ozonesonde launching stations having archived regularly ozone profile data to the NDACC Data Host Facility, the SHADOZ archive and/or the WOUDC. A detailed listing can be found in Table 8 in Appendix A.

4. Validation methodology

4.1 Total ozone column

4.1.1 Data preparation

This section deals with the pre-processing of the ground-based reference data so as to make them directly comparable to the satellite data set.

4.1.1.1 Unit conversions

As validation results are to be presented in the units of the validated product, the ground-based data are - for the validation of the level-3 products - converted from Dobson Units (DU) to mol/m^2. The conversion factor used is the one provided with the data files, in agreement with Basher (1982):

$$TC\left[\frac{mol}{m^2}\right] = \frac{1}{2241.339} TC[DU]$$

For validation of the TC_MSR level-4 product, provided in DU, no unit conversion is required.

4.1.1.2 Conversion to level-3 type data

To be able to compare as similar entities as possible, the ground-based total ozone records are averaged into monthly means one month at a time (i.e. no running mean), in accordance with the production of the satellite level-3 products. This is done under a set of strict criteria, which are the result of rigorous scientific analysis:

- Even though the following choice limits the absolute amount of common data points, for the Brewers and Dobsons only direct sun ground-based observations are included in the comparison to improve the quality of the results.
- A lower limit of 10 measurements per monthly mean is enforced.
- The effective day of the ground-based monthly mean is required to agree with the effective day of the C3S product to within 5 days. For the level-4 TC_MSR product, the effective day is always assumed to be the 15th (14th for February), for the level-3 products the effective day is computed from the actual sampling as provided in the “time_coverage_list” attribute of the files.

To be able to exploit our knowledge on the performance of individual ground-based instruments, records from different stations falling within a single 1°x1° grid cell are not combined. The potential gain in spatial representativeness does not weigh up against the loss in traceability, at least at this fine horizontal resolution of the level-3 and level-4 products validated here.

4.1.2 Co-location

The co-location criteria take into account that:

- In the level-3 satellite products, the TC field is reported on a latitude-longitude grid that represents the centres of the grid cells over which satellite data were averaged. Consequently, monthly station means are co-located with the nearest (lat, lon) coordinate (minimizing both the latitude and longitude difference) of the satellite level-3 product. In principle, if multiple stations co-locate with a single satellite grid point, i.e. they fall within the same satellite grid cell, they could be averaged to better represent the averaging behind the level-3 product. However, to keep track of station-to-station variations (due to ground-based instrument peculiarities), this avenue is not pursued (see also Sect. 4.1.1.2)

- In the level-4 product, the TC field is reported as point-like values on the (lat, lon) grid, and consequently also here co-location can be done by finding the minimum distance in latitude and longitude separately.

4.1.3 Estimation of data quality indicators

The baseline output of the validation exercise consists of time series of absolute and relative differences at individual stations, separated for different ground-based instruments. To investigate inter-product consistency, multiple products can be visualized together in as far as that is compatible with figure clarity. Perusal of these graphs by the validation experts can already reveal potentially complex data quality issues (outliers, sudden jumps,…) not always caught by the quantitative metrics discussed below.

4.1.3.1 Bias

For each station the median difference between C3S product and ground-based reference (both absolute and relative) is computed for the entire time series. This median difference is a robust (against outliers) estimator of the systematic error, i.e. the bias, of the satellite data product. Note that this median difference does also include contributions from representativeness (sampling and smoothing) differences, but thanks to the high horizontal resolution of the satellite products and our constraints on temporal representativeness of the ground-based level-3 product, these are kept to a minimum. It also includes any potential systematic error in the reference data, but this is unavoidable in any comparison with reference data. The biases for the entire list of stations are then visualized on a so-called "pole-to-pole" graph (bias vs. latitude), each station represented by a single marker, in order to reveal any latitudinal dependence of the systematic error.

4.1.3.2 Estimation of precision

Besides the median difference, also the q84-q16 interpercentile of the differences is calculated as a robust upper limit on the spread of the random errors in the satellite data product, i.e. the precision. As for the bias determination, contributions from representativeness differences are minimized as much as possible, and contributions from errors in the reference data are unavoidable. The q16 and q84 quantiles are added as (not necessarily symmetric) error bars to the "pole-to-pole" graph described above, in Sect. 4.1.3.1. Note again that the spread on differences will also include contributions from representativeness (sampling) differences, but thanks to the high horizontal resolution of the satellite products and our constraints on temporal representativeness of the ground-based level-3 product, these are kept to a minimum.

4.1.3.3 Estimation of stability

Long-term stability of the systematic errors in the ozone data products is a key requirement for C3S. Robust linear regressions[1] are performed on the satellite-ground differences at each station, including an uncertainty estimate on the derived drifts. To avoid spurious effects due to a seasonal cycle in the differences, only time series longer than 5 years are used for drift assessment. These results are again visualized as a function of latitude in a pole-to-pole graph.

[1] As implemented in MATLABs "robustfit".

4.2 Nadir ozone profile

4.2.1 Validation approach

The satellite-based and 1x1 degree gridded nadir profile level-3 data x_s^{L3} can be compared with spatially co-located ground-based reference profiles x_r directly, or with monthly (gridded) averages of the latter $\langle x_r \rangle$ (i.e. a ground-based level-3-type dataset). Yet both approaches introduce spatial and temporal representativeness errors into the difference statistics, and upon taking averages of the differences $\langle \Delta x \rangle$ both methods yield comparable outcomes:

$$\langle \Delta x \rangle = \frac{1}{N_m} \left\{ \left(x_s^{L3} - x_{r,1} \right) + \left(x_s^{L3} - x_{r,2} \right) + \cdots + \left(x_s^{L3} - x_{r,m} \right) \right\} = x_s^{L3} - \langle x_r \rangle$$

For sufficiently fine-gridded level-3 data, the comparisons can therefore be limited to the direct level-3 to ozonesonde differences, if one additionally only considers ozonesonde launch stations with a sufficient number N_m of launches with valid measurements per month. This number has been set to four in the ground-based validation presented here. As such, an implicit averaging of at least four ozonesonde measurements per month is introduced in the comparison statistics. The 1x1 degree box that overlaps with the ground measurements is thereby taken as the co-located measurement, in agreement with the 100-150 km nadir ozone profile co-location criterion that has been applied to the Ozone-CCI level-2 data (Keppens et al., 2015).

4.2.2 Data preparation

Raw ozonesonde profiles retrieved from the public NDACC, WOUDC and SHADOZ data archives are screened according to the criteria outlined in Hubert et al. (2016). At several ozonesonde stations an unexplained drop-off was observed, which affects estimates of satellite stability in the middle stratosphere. We therefore discard, for these sites, all profiles after the drop-off dates reported by Stauffer et al. (2020, Table 1). If there is no GPS altitude data in the data files, the altitude scale is reconstructed via the hydrostatic equation from the pressure and temperature recordings by the radiosonde attached to the ozonesonde. Ozone number density and volume mixing ratio (VMR) are computed using the same auxiliary data. The number density ozonesonde profiles are converted into partial column profiles by use of their corresponding altitude grids. These partial column profiles are then converted to the fixed 19-layer satellite grid by use of mass-conserved regridding, meaning that the integrated ozone column between the outer vertical edges is conserved (Langerock et al., 2015).

4.2.3 Estimation of data quality indicators

The baseline output of the validation exercise consists of median absolute and relative nadir ozone profile differences at individual stations or within latitude bands. Perusal of these graphs by the validation experts can already reveal potentially complex data quality issues (outliers, sudden jumps, …) not always caught by the quantitative metrics discussed below.

4.2.3.1 Bias

For each station the median difference profile (both absolute and relative) is computed for the entire time series. This median difference is a robust (against outliers) estimator of the vertically dependent systematic error, i.e. the bias, of the satellite data product. The bias profiles for the entire list of stations are then visualized for five latitude bands (with sections at -60, -20, 20, and 60 degrees latitude) in order to reveal any meridian dependence of the systematic error.

4.2.3.2 Estimation of precision

Besides the median difference, also the q84-q16 interpercentile of the differences is calculated as a robust upper limit on the spread of the random errors in the satellite data product, i.e. the precision profile. Note again that the spread on differences will also include contributions from representativeness (sampling) differences, but thanks to the high horizontal resolution of the satellite products and our constraints on temporal representativeness of the ground-based level-3 product, these are kept to a minimum.

4.2.3.3 Estimation of stability

Long-term stability of the systematic errors in the ozone data products is a key requirement for C3S. Robust linear regressions including an uncertainty estimate based on a bootstrapping approach (Hubert et al., 2016 are performed on the satellite-ground difference profiles at each station. To avoid spurious effects due to a seasonal cycle in the differences, only time series longer than 5 years are used for drift assessment.

4.3 Tropospheric ozone column

4.3.1 Validation approach

The tropospheric ozone column validation approach for the TCC_IASI products is analogous to the approach for the nadir ozone profile validation as described in Section 4.2.1.

4.3.2 Data preparation

The tropospheric ozone column data preparation for the TCC_IASI products is analogous to the approach for the nadir ozone profile validation as described in Section 4.2.2, with the single difference that the ozonesonde partial column profiles are integrated –again mass-conserved– from the ground up to the tropopause provided in the satellite product.

4.3.3 Estimation of data quality indicators

The estimation of data quality indicators for the tropospheric ozone column products is analogous to the description provided for the total ozone column products in Section 4.1.3.

4.4 Limb ozone profile

The validation of gridded limb products (level-3) differs considerably from that of single profiles (level-2). In a first step, the ozonesonde (and satellite) data are reshaped to comparable formats (Section 4.4.2). Then, quality indicators are derived from the comparison of the different incarnations of these data sets (Section 4.4.3). We start this section by explaining the need for such an adapted validation approach.

4.4.1 Challenges for validation of gridded products

The C3S limb profile data products described in Table 3 are averages of single profile retrievals from measurements by limb and occultation instruments. The LMZ products represent monthly mean data in 10° latitude bands, the LP product provides monthly means in smaller cells of 10° latitude by 20° longitude. In addition, some products contain profile data from multiple instruments (_MERGED). Single profile satellite data sets (so-called level-2 data) are usually validated using space and time co-located reference data (Hubert et al., 2016, RD8), but such an approach cannot simply be translated for aggregates of limb data (so-called level-3 data). In the following sections, we describe an approach to evaluate the quality of these data records using ozonesonde data as a reference.

The main challenge in evaluating level-3 satellite data is the launch frequency and the spatial density of the ground-based network which introduces considerable (depending on the bin size of the satellite product) spatial and temporal sampling errors. Most stations launch one balloon per week or twice a month. There are only a handful of sites that perform more frequent soundings, all of which are located in Europe. It is therefore not expected that the small monthly sample of sonde observations is representative of the monthly mean state of the ozone field around the station, especially in winter months which exhibit larger geophysical variability. In addition, there are many latitude bands and latitude-longitude cells without or with just a few stations (Figure 2). Similarly, it is therefore not expected that the data from a handful of stations is representative of the mean state of 10°x20° grid cells and especially not for 10° zonal bands.

Nonetheless, the stratospheric ozone field correlates over several thousand km over several days (Liu et al., 2009). A few 10°x20° grid cells in Europe and North America contain more than two stations with weekly soundings, which makes these prime locations to evaluate data quality. The investigation of larger-scale spatial structure of quality indicators in the stratosphere is more ambitious, especially for the zonally averaged LMZ products. The variability in the troposphere is larger than in the stratosphere, resulting in shorter correlation lengths and timescales (Liu et al., 2009). Combined with increased measurement noise by limb sounders it is particularly challenging to assess satellite data quality in the lower part of the atmosphere.

4.4.2 Data preparation

4.4.2.1 Ozonesonde

Raw ozonesonde profiles retrieved from the public NDACC, WOUDC and SHADOZ data archives are screened according to the criteria outlined in Hubert et al. (2016). At several ozonesonde stations an unexplained drop-off was observed, which affects estimates of satellite stability in the middle stratosphere. We therefore discard, for these sites, all profiles after the drop-off dates reported by Stauffer et al. (2020, Table 1). If there is no GPS altitude information in the ozonesonde data file, the altitude scale is reconstructed via the hydrostatic equation from pressure and temperature readings of the radiosonde attached to the ozonesonde. Ozone concentration and volume mixing ratio (VMR) are computed from partial pressure using the same auxiliary data. Then, VMR and number density profiles are interpolated to the vertical grid of the C3S products using a pseudo-inverse interpolation method described in Calisesi et al. (2005). In a next step, the vertically gridded data are averaged by month over the entire time series. Uncertainties in the derived monthly mean value are reduced by rejecting months and grid levels with <2 (tropics) or <3 (higher latitudes) profiles. The resulting vertically and temporally gridded ozonesonde data set is referred to as the Station Monthly Mean data, one per site: **SMM**(site, grid level, time).

The site-specific seasonal cycle is then computed as an average, for each calendar month (Jan, Feb, …, Dec), of all SMM data in the reference period. The latter was chosen to coincide with that used for the LMZ_MERGED product: Jan 2004 to Dec 2011 (C3S ATBD, RD3). Seasonal cycle entries are discarded for months and grid levels that contain <4 years (tropics) or <5 years (higher latitudes) of SMM data. This requirement ensures a more accurate determination of the observed seasonal cycle, but is only satisfied for a select number of sites. This sonde data set will be referred to as the Station Seasonal Cycle data set, one per site: **SSC**(site, grid level, month).

We then calculate the relative deviation of the SMM data from the local observed SSC data. This deseasonalised relative anomaly data set (expressed in %) will be called the Station Monthly Mean Anomaly data set (**SMMa**) and is defined as

$$\mathrm{SMMa(site, z, t)} = 100 \times \frac{SMM(site, z, t) - SSC\big(site, z, m(t)\big)}{SSC\big(site, z, m(t)\big)},$$

where z stands for vertical grid level (either pressure or altitude), t represents time (i.e. month) and $m(t)$ the corresponding calendar month (i.e. Jan, Feb, …). Hence, by construction, the (dominant part of the) seasonal cycle is removed from the SMMa data and the absolute level averages to zero over the reference period (as for some C3S limb products). In addition, instrument-related multiplicative offsets (i.e. bias) are hereby removed as well.

The previous step is motivated by the need to combine data from different sonde sites over latitude belts and grid cells. Site-dependent instrument biases will generate, in a multiple station averaged SMM data set, not only random uncertainty but also jumps (due to differences in time coverage). However, under the assumption that errors are multiplicative, these sources of errors are suppressed for such a multiple station averaged SMMa data set. The latter will hereafter be called the Cell Monthly Mean Anomaly data set (**CMMa**), where the cell can either stand for a 10° zonal band or for a 10°x20° grid. All SMMa data are weighed equally, which effectively gives more weight to regions (Europe, North America) with more stations. The latter introduces a sampling bias that may be important for the zonal mean CMMa in most bands, but this is less of an issue for the smaller, gridded CMMa data given the large correlation lengths in the stratosphere. In addition, we compute the Cell Monthly Mean data set (**CMM**) in a similar way as the CMMa data but directly from the SMM records. However, unlike for CMMa data, systematic uncertainties from the measurements are not reduced for the CMM record.

4.4.2.2 Satellite data

LMZ products of single instruments contain either monthly mean ozone concentration on an altitude grid or monthly mean volume mixing ratio data on a pressure grid (Table 3). The merged products (LMZ and LP) on the other hand consist also of deseasonalized relative anomalies on an altitude grid, and are hereafter referred to as *LMZa* and *LPa*, respectively. However, the estimation of some quality indicators requires deseasonalized anomaly data for the single-instrument LMZ data as well.

For some instruments (GOMOS, MIPAS, SCIAMACHY, OSIRIS, ACE-FTS, SAGE II, OMPS) the deseasonalized data are taken from the LMZ_MERGED product. For the remaining instruments (SMR, HALOE, Aura MLS, SABER) such data are delivered to us directly by the data providers. Below we denote the deseasonalized anomaly data by LMZa, to clarify it is a product derived from the official C3S limb product.

4.4.3 Estimation of data quality indicators

The data quality assessment includes estimates of bias, short-term variability and long-term stability of the C3S limb products relative to the gridded ozonesonde data.

These indicators are derived from two kinds of comparison time series, each representing percentage differences. The first time series consist of relative differences of the original LMZ and CMM data as ozone concentrations (or volume mixing ratios),

$$\Delta(lat/lon\,,z,t) = 100 \times \left(\frac{LMZ(lat/lon,z,t) - CMM(lat/lon,z,t)}{CMM(lat/lon,z,t)}\right).$$

The second consists of absolute differences of deseasonalized data,

$$\Delta_a(lat/lon\,,z,t) = LMZa(lat/lon\,,z,t) - CMMa(lat/lon\,,z,t),$$

where LMZa and CMMa express a percentage deviation from the seasonal cycle in the respective data record.

The bias B in a C3S limb product will be estimated as the median value of the Δ (single-sensor data) or of the Δ_a (merged data) difference time series. This bias originates from systematic uncertainty in the satellite and ground-based products and from systematics in differences in sampling, as discussed in Section 4.4.1. Further studies are needed, which fall outside the scope of this work, to disentangle these components. For the moment, the reported satellite bias B can not be interpreted as a strict measure of satellite systematic uncertainty. In the analysis of merged satellite products, all data sets are deseasonalized over the same reference period (Jan 2004 – Dec 2011). So, by definition, the bias B for the merged C3S products will be insensitive to differences in this period.

The spread S in the Δ_a difference time series is derived from the 16-84% interpercentile (divided by 2 this corresponds to the standard deviation of a normal distribution). It is generated by random errors in the satellite and the ozonesonde products and by random errors from differences in spatial and temporal sampling. It is out of the scope of this work to quantify the sampling uncertainty, so here we report S as a (conservative) upper limit to the random uncertainty (i.e. precision) in the C3S limb products. Note that we use the anomaly time series to compute S for not only the merged but also the single-sensor C3S products.

The stability α of C3S limb data is estimated from a robust linear regression of the Δ_a difference time series

$$\Delta_a(lat/lon,z,t) = \alpha(lat/lon,z) \times (t - t_0) + \beta(lat/lon\,,z) + \varepsilon(lat/lon\,,z,t).$$

The coefficient α is interpreted as the linear drift between the C3S and ozonesonde data. Drift uncertainty S_α is determined from the fit residuals ε, which receive contributions from random uncertainty as well as from non-modelled temporal structure in the difference time series. In a second step, the zonal/cell estimates of drift are averaged over the globe, to reduce the impact of temporal inhomogeneities in the reference data

$$\alpha(z) = average\big(\alpha(lat/lon\,,z)\big).$$

This basically follows the approach outlined in Hubert et al. (2016), though applied to gridded deseasonalized anomaly data (level-3) instead of retrieved ozone concentration data (level-2). The structure of drift in the horizontal domain will not be quantified.

5. Compliance with user requirements

The quality indicators derived from the quantitative assessment (computation of estimators for bias, precision and stability) described above can be used to assess the compliance of C3S_312b_Lot2 ozone CDRs with target user requirements, and also to quantify Key Performance Indicators (KPI) for project management purposes. In the sections below, we list the user requirements applicable here, which have been taken from the "Target Requirements and Gap Analysis Document" (RD2). They are derived from the GCOS Implementation Plan requirements on ozone (RD10), and from the Ozone_cci User Requirements Document (RD7) describing additional requirements collected from ESA CCI's Climate Modelling and User Group (CMUG). Additional user requirements are derived from the ESA OPEROZ study on operational ozone profile requirements (RD9). The target requirements summarized in the following tables encompass (target/threshold) spatiotemporal resolution and coverage, uncertainty and stability requirements.

The stability requirements correspond directly with the stability metrics described in Section 4, but the "total uncertainty" requirement needs to be assessed against a quantified bias and precision metrics. To that end, the estimated satellite bias and spread are combined quadratically and this number is considered to be a conservative (as in "pessimistic") estimate of the total uncertainty. Please note that in some cases, e.g. for the single-instrument LMZ limb products, the quoted total uncertainty should be interpreted with great caution since the bias estimates are not necessarily accurate and their significance is poorly known.

Note that there remains some ambiguity in these requirements as to the exact quantitative meaning of "total uncertainty" and "accuracy". We take it here to be the measurement uncertainty due to both random errors and systematic effects.

5.1 Total and tropospheric ozone column products

Table 4. (Target) user requirements for Total and Tropospheric ozone denoted as (goal/threshold). GCOS requirements (in brackets) are target user requirements and only denoted if more demanding than the goal. Copied from RD2.

Total and Tropospheric Ozone			**Stability: 1% / 3% per decade**	
Observable	Horizontal resolution	Horizontal coverage	Update frequency	Uncertainty
Total ozone	20 / 100	Global (incl. polar night)	Daily/ weekly (4h)	2% / 3%
Tropospheric ozone	20 / 200	Global (incl. polar night)	Daily/ weekly (4h)	8% / 16%

5.2 Nadir ozone profile products

Table 5. (Target) user requirements for ozone profile (nadir) denoted as (goal/threshold). GCOS requirements (in brackets) are target user requirements and only denoted if more demanding than the goal. Copied from RD2.

Ozone profile (nadir)				**Stability: 1% / 3% per decade**		
Observable	Horizontal resolution	Horizontal coverage	Vertical resolution (km)	Vertical coverage	Update frequency	Uncertainty
Ozone profile (nadir)	20 / 200	Global incl. polar night	6 / partial column (5)	Troposphere	Daily / weekly (4h)	8% / 16%
Ozone profile (nadir)	100 / 200 (20 / 50)	Global incl. polar night	6 / partial column (1 /2)	Lower stratosphere	Daily / weekly (4h)	8% / 16% or 50 / 100 ppbv
Ozone profile (nadir)	100 / 200 (20 / 50)	Global incl. polar night	6 / partial column (1 / 2)	Middle stratosphere	Daily / weekly (4h)	4% / 8% or 50 / 100 ppbv
Ozone profile (nadir)	200 / 400 (20 / 50)	Global incl. polar night	6 / partial column (3)	Upper stratosphere / mesosphere	Daily / weekly	4% / 8%

5.3 Limb ozone profile products

Table 6. (Target) user requirements for ozone profile (limb) denoted as (goal/threshold). GCOS requirements (in brackets) are target user requirements and only denoted if more demanding than the goal. Copied from RD2.

Ozone profile (limb)				**Stability: 1% / 3% per decade**		
Observable	Along-track sampling (km)	Horizontal coverage	Vertical resolution (km)	Vertical coverage	Update frequency	Uncertainty
Ozone profile (limb)	100 / 200	Global incl. polar night	1 / 2	Lower stratosphere	Daily / weekly (4h)	8% / 16% or 50 / 100 ppbv
Ozone profile (limb)	100 / 200	Global incl. polar night	1 / 2	Middle stratosphere	Daily / weekly (4h)	4% / 8% or 50 / 100 ppbv
Ozone profile (limb)	200 / 400 (100 / 200)	Global incl. polar night	2 / 4	Upper Stratosphere / mesosphere	Daily / weekly	4% / 8%

6. Summary of validation results

Validation results of the ozone climate data records procured by the C3S_312b_Lot2 service by February 2020 are reported in the Product Quality Assessment Report (RD5). This version gives the following conclusions:

- The long-term stability of the **total ozone column data products** w.r.t. the ground-based network, and in particular the stability of the long time series of TC_GTO-ECV (level-3) and TC_MSR (level-4), meets user requirements, with drifts typically below the 1%/decade level. The mean difference and spread on the differences indicate systematic and random errors below 2% and 3-4% respectively, also satisfying the user requirements. The earliest decade (1970s) of the TC_MSR product unfortunately contains some data gaps, resulting from the insufficient amount of satellite measurements constraining the data merging system.
- The C3S level-3 **nadir ozone profile products** show, between about 10 hPa and the tropopause (100-200 hPa), relative differences and spreads of the order of 5-10% and 10%, respectively, for all instruments, while the troposphere shows a 10-20% bias (both positive and negative) and 40% spread. Strong outliers occur, however, typically in the Antarctic local winter (JJA) and spring (SON) due to strong ozone variability around the polar vortex. GOME-2B moreover shows a clear meridian bias dependence in the troposphere, going from 30% positive in the Arctic to up to 50% negative in the southern hemisphere. The decadal drift is order of 5-10% per decade and insignificant for GOME and OMI at all altitudes under consideration, while a significant positive drift of the order of 20 and 30% per decade is observed for, respectively, SCIAMACHY and GOME-2A below the tropopause. The GOME-2B drift assessment requires a longer time series to draw significant conclusions.
- The two IASI level-3 monthly gridded **tropospheric ozone column products** show a strong seasonal variation in their comparison with 0-6 km integrated ozonesonde data, ranging up to 100% at the southern pole. Median relative differences range between 30% negative in the northern mid-latitudes and 30% positive in Antarctica, with a nearly zero overall bias around the equator. Globally averaged biases are of the order of 25% negative. The spread decreases from about 30% in the tropics to about 10% towards the poles. On the global scale, its seasonality decreases from 2016 onwards for both IASI-A and IASI-B. Decadal drift results are also similar for IASI-A and IASI-B. An order of 10 %/decade significant negative drift is detected for both instruments on the global scale.
- Estimates of bias, comparison spread and drift for the C3S level-3 **limb ozone profile products** are consistent with those obtained for level-2 data. The considerably looser level-3 co-location criteria allow for more precise drift estimates than for level-2 data. Insignificant drifts of less than 1-3% per decade are found for most data products. A few products exhibit larger and significant drift of 3-5% per decade in part of the atmosphere. The significant positive drift in OMPS-LP data in the Feb 2019 data release is reduced by 1-2% per decade for the Feb 2020 data release. These become insignificant in the lower stratosphere. Drift estimates for the merged monthly zonal mean data record are also slightly reduced, by ~0.5% per decade. Comparison spreads above 20 km are slightly larger than at level-2 for all sensors except SCIAMACHY and SMR. The spread increases from tropics (~5%) to poles (~12%), which is believed to be driven by differences in sampling in the presence of higher natural variability

rather than poorer satellite random uncertainty. At lower altitudes spreads rapidly increase to 30-40% or more, again due to natural variability though the lower ozone concentrations and limited signal-to-noise ratio of satellite measurements play a role too. Inhomogeneities in the ozonesonde record make it inherently more challenging to estimate biases down to the few percent level. As a result, the satellite bias field has notable variance in the horizontal domain. Nonetheless, overall the satellite and ground-based data agree within 5% or better in the lower stratosphere and only three data records exhibit larger biases of 5-10%. Larger mean differences are found in the UTLS and below, where ozone concentrations are low.

7. References

Basher, R. E., Units For Column Amounts of Ozone and Other Atmospheric Gases. Q.J.R. Meteorol. Soc., 108: 460–462. doi:10.1002/qj.49710845611, 1982.

Balis, D., Lambert, J.-C., Van Roozendael, M., Loyola, D., Spurr, R., et al.: Ten years of GOME/ERS-2 total ozone data – the new GOME Data Processor (GDP) Version 4: II Ground-based validation and comparisons with TOMS V7/V8, J. Geophys. Res.-Atmos., 112, D07307, doi:10.1029/2005JD006376, 2007.

Calisesi, Y., V. T. Soebijanta, and R. van Oss, Regridding of remote soundings: Formulation and application to ozone profile comparison, J. Geophys. Res., 110, D23306, doi:10.1029/2005JD006122, 2005.

Coldewey-Egbers, M., Loyola, D. G., Koukouli, M., Balis, D., Lambert, J.-C., et al.: 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.

Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., et al.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records, Atmos. Meas. Tech., 9, 2497-2534, https://doi.org/10.5194/amt-9-2497-2016, 2016.

Keppens, A., Lambert, J.-C., Granville, J., Miles, G., Siddans, R., et al.: Round-robin evaluation of nadir ozone profile retrievals: methodology and application to MetOp-A GOME-2, Atmos. Meas. Tech., 8, 2093-2120, https://doi.org/10.5194/amt-8-2093-2015, 2015.

Keppens, A., Compernolle, S., Verhoelst, T., Hubert, D., and Lambert, J.-C.: Harmonization and comparison of vertically resolved atmospheric state observations: methods, effects, and uncertainty budget, Atmos. Meas. Tech., 12, 4379–4391, https://doi.org/10.5194/amt-12-4379-2019, 2019.

Kerr, J. B., I. A. Asbridge, and W. F. J. Evans, Intercomparison of total ozone measured by the Brewer and Dobson spectrophotometers at Toronto, J. Geophys. Res., 93 (D9), 11,129– 11, 140, 1988.

Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, Geosci. Model Dev., 8, 911–921, https://doi.org/10.5194/gmd-8-911-2015, 2015.

Liu, G., D. W. Tarasick, V. E. Fioletov, C. E. Sioris, and Y. J. Rochon, Ozone correlation lengths and measurement uncertainties from analysis of historical ozonesonde data in North America and Europe, J. Geophys. Res., 114, D04112, doi:10.1029/2008JD010576, 2009.

Smit, H., and the Panel for the Assessment of Standard Operating Procedures for Ozonesondes (ASOPOS), Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, GAW Report No. 201, September 2014.

Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Witte, J. C., Tarasick, D. W., Davies, J., et al. (2020). A post-2013 dropoff in total ozone at a third of global ozonesonde stations: Electrochemical concentration cell instrument artifacts? Geophys. Res. Lett., 47, e2019GL086791. https://doi.org/10.1029/2019GL086791.

Staehelin J., J. Kerr, R. Evans and K. Vanicek, Comparison of total ozone measurements of Dobson and Brewer spectrophotometers and recommended transfer functions, WMO TD N. 1147, No 149, 2003.

Van Roozendael, M.,P. Peters, H. K. Roscoe, et al., Validation of Ground-Based Visible measurements of Total Ozone by Comparison with Dobson and Brewer Spectrophotometers, J. Atmos. Chem, 29, 5–83, 1998.

Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J.-P., Redondas, A., Van Roozendael, M., and Lambert, J.-C.: Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons, Atmos. Meas. Tech., 8, 5039-5062, https://doi.org/10.5194/amt-8-5039-2015, 2015.

World Meteorological Organization, Scientific assessment of ozone depletion: 2006, Global Ozone Res. Monit. Proj., Rep. 50, Geneva, Switzerland, 2006.

Appendix A: Details of validation data sets

Table 7. Details of Dobson and Brewer hosting stations (listed from South to North) considered for the validation of C3S_312b_Lot2 total ozone column data products (see also Figure 1).

Station (Country)	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible Institute	Instrument
Amundsen-Scott (Antarctica)	-89.98	-24.80	2810	NOAA-CMDL	Brewer
Doctor Sobral (Antarctica)	-81.07	-40.50	100	AAS-AAAC	Brewer
Arrival Heights (Antarctica)	-77.83	166.67	184	NIWA	Dobson
Princess Elisabeth (Antarctica)	-71.95	23.35	1350	RMIB	Brewer
Maitri (Antarctica)	-70.45	11.45	330	IMD	Brewer
San Martin (Antarctica)	-68.13	-67.11	30	AAS-AAAC	Brewer
Marambio (Antarctica)	-64.23	-56.62	198	IAA-CHMI	Brewer
Marambio (Antarctica)	-64.23	-56.62	198	SMNA	Dobson
Ushuaia (Argentina)	-54.85	-68.31	7	SMNA	Dobson
Comodoro Rivadavia (Argentina)	-45.78	-67.50	46	SMNA	Dobson
Broadmeadows (Australia)	-37.69	144.95	108	ABM	Dobson
Buenos Aires (Argentina)	-34.58	-58.48	25	SMNA	Dobson
Perth (Australia)	-31.92	115.96	2	NOAA-CMDL	Dobson
Salto (Uruguay)	-31.43	-57.97	41	DNMUY	Dobson
Springbok (South Africa)	-29.67	17.90	1006	SAWS	Dobson
Irene (South Africa)	-25.92	28.22	1523	SAWS	Dobson
Cachoeira Paulista (Brazil)	-22.68	-45.00	573	INPE	Dobson
Maun (Botswana)	-19.98	23.43	950	DMSB	Dobson
Darwin (Australia)	-12.42	130.89	30	ABM	Dobson
Marcapomacocha (Peru)	-11.40	-76.32	4479	NOAA-CMDL	Dobson
Natal (Brazil)	-5.83	-35.20	32	INPE	Dobson
Mahe (Seychelles)	-4.68	55.53	6	SNMS	Dobson
Nairobi (Kenya)	-1.27	36.80	1795	KMD	Dobson
Singapore (Singapore)	1.33	103.88	14	MSS	Dobson
Petaling Jaya (Malaysia)	3.10	101.64	86	MMS	Brewer
Paramaribo (Suriname)	5.81	-55.22	7	KNMI	Brewer
Lagos (Nigeria)	6.60	3.33	10	NMS	Dobson
Songkhla (Thailand)	7.20	100.60	13	TMD	Brewer
Kodaikanal (India)	10.23	77.47	2343	IMD	Brewer
Bangkok (Thailand)	13.67	100.61	53	TMD	Brewer
Bangkok (Thailand)	13.67	100.61	53	TMD	Dobson
Manila (Philippines)	14.65	121.05	61	MSP	Dobson
Poona (India)	18.53	73.85	559	IMD	Brewer
Poona (India)	18.53	73.85	559	IMD	Dobson
Mexico City (Mexico)	19.33	-99.18	2268	MIG	Dobson
Mauna Loa (United States)	19.53	-155.58	3397	MSC-MLO	Brewer

Station (Country)	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible Institute	Instrument
Mauna Loa (United States)	19.53	-155.58	3397	NOAA-CMDL	Dobson
Hanoi (Viet Nam)	21.03	105.85	5	HSSRV	Brewer
Cape D'Aguilar (Hong Kong)	22.21	114.26	60	HKPU	Brewer
Tamanrasset (Algeria)	22.80	5.52	1382	RMDA	Brewer
Tamanrasset (Algeria)	22.80	5.52	1382	RMDA	Dobson
Chengkung (Taiwan, Province Of China)	23.10	121.36	39	CWBT	Brewer
Havana (Cuba)	23.28	-82.55	50	IMC	Dobson
Aswan (Egypt)	23.97	32.78	190	EMA	Dobson
Kunming (China)	25.03	102.68	1917	CAS-IAP	Dobson
Taipei (Taiwan, Province Of China)	25.04	121.51	22	CWBT	Brewer
Varanasi (India)	25.32	83.03	76	IMD	Dobson
Naha (Japan)	26.20	127.68	27	JMA	Dobson
Hurghada (Egypt)	27.28	33.75	7	EMA	Dobson
Izaña (Spain)	28.30	-16.50	2367	AEMET	Brewer
Santa Cruz (Spain)	28.46	-16.26	36	AEMET	Brewer
New Delhi (India)	28.65	77.22	220	IMD	Dobson
Lhasa (China)	29.67	91.13	3650	CAMS-IAC	Brewer
Cairo (Egypt)	30.08	31.28	37	EMA	Dobson
Linan (China)	30.30	119.73	132	CAMS-IAC	Brewer
Tallahassee (United States)	30.40	-84.35	21	NOAA-CMDL	Dobson
Mrsa Matrouh (Egypt)	31.33	27.22	35	EMA	Brewer
Kagoshima (Japan)	31.55	130.55	31	JMA	Dobson
Isfahan (Iran, Islamic Republic Of)	32.48	51.42	1550	MDI	Brewer
Isfahan (Iran, Islamic Republic Of)	32.48	51.42	1550	IGUT	Dobson
Funchal (Portugal)	32.64	-16.89	49	PIM	Brewer
Tehran University (Iran, Islamic Republic Of)	35.73	51.38	1419	IGUT	Dobson
Tsukuba (Japan)	36.05	140.13	31	JMA	Dobson
Nashville (United States)	36.25	-86.57	182	NOAA-CMDL	Dobson
Mt Waliguan (China)	36.29	100.90	3810	CAMS-IAC	Brewer
Hanford (United States)	36.32	-119.63	73	NOAA-CMDL	Dobson
Anmyeon-do (Korea, Republic Of)	36.54	126.33	57	KMA	Brewer
El Arenosillo (Spain)	37.10	-6.73	41	INTA	Brewer
El Arenosillo (Spain)	37.10	-6.73	41	INTA	Dobson
Seoul (Korea, Republic Of)	37.57	126.98	84	Yonsei_U	Brewer
Seoul (Korea, Republic Of)	37.57	126.98	84	Yonsei_U	Dobson
Wallops Island (United States)	37.93	-75.48	13	NOAA-CMDL	Dobson
Athens (Greece)	37.98	23.73	280	U_ATHENS	Dobson
Murcia (Spain)	38.00	-1.16	69	AEMET	Brewer
Lisbon (Portugal)	38.77	-9.15	105	PIM	Brewer

Station (Country)	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible Institute	Instrument
Lisbon (Portugal)	38.77	-9.15	105	PIM	Dobson
Greenbelt (United States)	38.99	-76.83	100	NASA-TOMS	Brewer
Ankara (Turkey)	39.97	32.86	913	TSMS	Brewer
Xianghe (China)	39.98	116.37	80	CAS-IAP	Dobson
Boulder (United States)	40.03	-105.25	1689	NOAA-CMDL	Dobson
Amberd (Armenia)	40.38	44.25	2070	AHMS	Dobson
Thessaloniki (Greece)	40.52	22.97	50	AUTH	Brewer
Zaragoza (Spain)	41.63	-0.88	258	AEMET	Brewer
Rome University (Italy)	41.90	12.50	75	U_ROME	Brewer
Lannemezan (France)	43.12	0.37	590	LA-OMP	Dobson
Toronto University (Canada)	43.66	-79.40	174	U_Toronto	Brewer
Kislovodsk (Russian Federation)	43.73	42.66	2070	RAS-IAP	Brewer
Toronto (Canada)	43.78	-79.47	198	MSC	Dobson
Observatoire de Haute-Provence (France)	43.94	5.71	650	NOAA-CMDL	Dobson
Sestola (Italy)	44.22	10.77	1030	AM-IMS	Brewer
Sestola (Italy)	44.22	10.77	1030	AM-IMS	Dobson
Bucharest (Romania)	44.48	26.13	100	RNIMH	Dobson
Halifax (Canada)	44.67	-63.57	31	MSC	Brewer
Longfengshan (China)	44.73	127.58	334	CAMS-IAC	Brewer
Bordeaux (France)	44.84	-0.53	73	OBX	Dobson
Montreal (Canada)	45.48	-73.75	31	MSC	Brewer
Aosta (Italy)	45.74	7.36	570	ARPA-VDA	Brewer
Ispra (Italy)	45.80	8.63	240	JRC_EU	Brewer
Bismarck (United States)	46.77	-100.75	511	NOAA-CMDL	Dobson
Arosa (Switzerland)	46.78	9.68	1840	MCH	Brewer
Arosa (Switzerland)	46.78	9.68	1840	MCH	Dobson
Caribou (United States)	46.87	-68.03	192	NOAA-CMDL	Dobson
Budapest-Lorinc (Hungary)	47.43	19.18	139	HMS	Brewer
Hohenpeißenberg (Germany)	47.81	11.01	975	DWD-MOHp	Brewer
Hohenpeißenberg (Germany)	47.81	11.01	975	DWD-MOHp	Dobson
Poprad Ganovce (Slovakia)	49.03	20.32	706	SHMI	Brewer
Winnipeg (Canada)	49.90	-97.24	239	MSC	Brewer
Hradec Kralove (Czech Republic)	50.18	15.83	285	CHMI-HK	Dobson
Bratt's Lake (Canada)	50.21	-104.71	592	MSC	Brewer
Kyiv-Goloseyev (Ukraine)	50.36	30.50	206	KNU	Dobson
Uccle (Belgium)	50.80	4.35	100	RMIB	Dobson
Belsk (Poland)	51.84	20.79	180	PAS	Brewer
Belsk (Poland)	51.84	20.79	180	PAS	Dobson
Valentia (Ireland)	51.94	-10.25	14	ME	Brewer
Saskatoon (Canada)	52.11	-106.71	550	SCI-TEC	Brewer

Station (Country)	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible Institute	Instrument
Lindenberg (Germany)	52.21	14.12	127	DWD-MOL	Dobson
Potsdam (Germany)	52.22	13.05	89	DWD-MOP	Brewer
Potsdam (Germany)	52.22	13.05	89	DWD-MOP	Dobson
Goose Bay (Canada)	53.31	-60.36	40	MSC	Brewer
Manchester (United Kingdom)	53.47	-2.23	76	U_Manchester	Brewer
Obninsk (Russian Federation)	55.10	36.61	100	IEM-SPA	Brewer
Moscow (Russian Federation)	55.75	37.57	187	CAO	Dobson
Tomsk (Russian Federation)	56.47	84.95	280	IOA	Brewer
Norrkoeping (Sweden)	58.58	16.15	43	SMHI	Brewer
Oslo (Norway)	59.94	10.72	90	U_Oslo	Brewer
Jokioinen (Finland)	60.81	23.50	103	FMI	Brewer
Yakutsk (Russian Federation)	62.02	129.72	100	CAO	Brewer
Reykjavik (Iceland)	64.13	-21.90	64	IMO	Dobson
Vindeln (Sweden)	64.24	19.77	225	SMHI	Brewer
Vindeln (Sweden)	64.24	19.77	225	SMHI	Dobson
Fairbanks (United States)	64.82	-147.87	138	NOAA-CMDL	Dobson
Søndre Strømfjord (Greenland)	67.00	-50.62	300	DMI	Brewer
Sodankylä (Finland)	67.37	26.63	179	FMI	Brewer
Andoya (Norway)	69.28	16.00	380	NILU	Brewer
Barrow (United States)	71.32	-156.60	11	NOAA-CMDL	Dobson
Resolute (Canada)	74.72	-94.98	64	MSC	Brewer
Ny-Ålesund (Norway)	78.93	11.93	10	CNR	Brewer
Eureka (Canada)	79.99	-85.93	610	MSC	Brewer

Table 8. Details of the 64 ozonesonde launching stations (listed from North to South) considered for the validation of C3S_312b_Lot2 ozone profile and tropospheric ozone data products (see also Figure 2).

Ozonesonde station	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible institute	Data archive
Alert	82.5	-62.5	62	ECCC	WOUDC
Eureka	80.0	-85.9	610	ECCC	WOUDC
Ny-Ålesund	78.9	11.9	10	AWI-NA	WOUDC
Thule	76.5	-68.7	57	DMI	NDACC
Resolute	74.7	-95.0	64	ECCC	WOUDC
Summit	72.3	-38.3	3200	NOAA	NDACC
Scoresbysund	70.5	-22.0	67	DMI	NDACC
Sodankylä	67.4	26.6	179	FMI	NDACC
Lerwick	60.1	-1.2	82	UKMO	WOUDC
Churchill	58.7	-94.1	35	ECCC	WOUDC
Edmonton	53.6	-114.1	766	ECCC	WOUDC
Goose Bay	53.3	-60.4	40	ECCC	WOUDC
Legionowo	52.4	21.0	96	PIMWM	WOUDC
Lindenberg	52.2	14.1	127	DWD-MOL	WOUDC
De Bilt	52.1	5.2	15	KNMI	NDACC
Valentia	51.9	-10.3	14	ME	WOUDC
Uccle	50.8	4.4	100	RMIB	WOUDC
Bratts Lake	50.2	-104.7	592	ECCC	WOUDC
Praha	50.0	14.4	304	CHMI-PR	WOUDC
Kelowna	49.9	-119.4	456	ECCC	WOUDC
Hohenpeißenberg	47.8	11.0	975	DWD-MOHp	WOUDC
Payerne	46.8	7.0	491	MeteoSwiss	WOUDC
Egbert	44.2	-79.8	252	ECCC	WOUDC
Observatoire de Haute Provence	43.9	5.7	650	LATMOS-CNRS	NDACC
Yarmouth	43.9	-66.1	9	ECCC	WOUDC
Sapporo	43.1	141.3	26	JMA	WOUDC
Madrid	40.5	-3.7	680	AEMET	WOUDC
Boulder	40.0	-105.3	1689	NOAA	NDACC
Wallops Island	37.9	-75.5	13	NASA	WOUDC
Tsukuba	36.1	140.1	31	JMA	WOUDC
Huntsville	34.7	-86.6	196	UAH	WOUDC
Isfahan	32.5	51.4	1550	MDI	WOUDC
Kagoshima	31.6	130.6	31	JMA	WOUDC
New Delhi	28.7	77.2	220	IMD	WOUDC
Izaña	28.3	-16.5	2367	AEMET	NDACC
Naha	26.2	127.7	27	JMA	WOUDC

Ozonesonde station	Latitude [deg N]	Longitude [deg E]	Altitude [m a.s.l.]	Responsible institute	Data archive
Hong Kong Observatory	22.3	114.2	66	HKO	WOUDC
Hanoi	21.0	105.9	5	HSSRV	SHADOZ
Hilo	19.7	-155.1	11	NOAA	SHADOZ
Heredia	10.0	-84.1	1176	DWD-GRUAN	SHADOZ
Paramaribo	5.8	-55.2	7	KNMI	SHADOZ
Kaashidhoo	5.0	73.5	1	NOAA	SHADOZ
Sepang Airport	2.7	101.7	17	MMD	SHADOZ
San Cristobal	-0.9	-89.6	8	NOAA	SHADOZ
Nairobi	-1.3	36.8	1795	MeteoSwiss	SHADOZ
Malindi	-3.0	40.2	-6	U Rome-CRPSM	WOUDC
Natal	-5.8	-35.2	32	NASA	SHADOZ
Watukosek	-7.5	112.6	50	Hokkaido U	SHADOZ
Ascension Island	-8.0	-14.4	79	NASA	SHADOZ
Samoa	-14.3	-170.6	82	NOAA	SHADOZ
Papeete	-18.0	-149.0	2	NOAA	SHADOZ
Suva	-18.1	178.4	6	NOAA	SHADOZ
Saint Denis	-20.9	55.5	110	U LaReunion-CNRS	SHADOZ
Irene	-25.9	28.2	1523	SAWS	SHADOZ
Broadmeadows	-37.7	145.0	108	ABM	WOUDC
Lauder	-45.0	169.7	370	NIWA	NDACC
Macquarie	-54.5	159.0	6	ABM	WOUDC
Marambio	-64.2	-56.6	198	FMI, SMNA	WOUDC
Dumont d'Urville	-66.7	140.0	45	LATMOS-CNRS	NDACC
Davis	-68.6	78.0	18	ABM	WOUDC
Syowa	-69.0	39.6	22	JMA	WOUDC
Neumayer	-70.7	-8.3	43	AWI-NM	WOUDC
Mac Murdo	-77.9	166.6	10	U Wyoming	NDACC
Amundsen Scott	-90.0	-24.8	2810	NOAA	NDACC

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu