



The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative

M. Coldewey-Egbers¹, D. G. Loyola¹, M. Koukouli², D. Balis², J.-C. Lambert³, T. Verhoelst³, J. Granville³, M. van Roozendaal³, C. Lerot³, R. Spurr⁴, S. M. Frith⁵, and C. Zehner⁶

¹German Aerospace Center (DLR), Wessling, Germany

²Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece

³Belgian Institute for Space Aeronomy (BIRA), Brussels, Belgium

⁴RT Solutions Inc., Cambridge, Massachusetts, USA

⁵Science Systems and Applications Inc., National Aeronautics and Space Administration/Goddard Space Flight Center (NASA/GSFC), Greenbelt, Maryland, USA

⁶European Space Agency/European Space Research Institute (ESA/ESRIN), Frascati, Italy

Correspondence to: M. Coldewey-Egbers (melanie.coldewey-egbers@dlr.de)

Received: 13 March 2015 – Published in Atmos. Meas. Tech. Discuss.: 6 May 2015

Revised: 1 September 2015 – Accepted: 2 September 2015 – Published: 24 September 2015

Abstract. We present the new GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record which has been created within the framework of the European Space Agency's Climate Change Initiative (ESA-CCI). Total ozone column observations – based on the GOME-type Direct Fitting version 3 algorithm – from GOME (Global Ozone Monitoring Experiment), SCIAMACHY (SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY), and GOME-2 have been combined into one homogeneous time series, thereby taking advantage of the high inter-sensor consistency. The data record spans the 15-year period from March 1996 to June 2011 and it contains global monthly mean total ozone columns on a $1^\circ \times 1^\circ$ grid. Geophysical ground-based validation using Brewer, Dobson, and UV-visible instruments has shown that the GTO-ECV level 3 data record is of the same high quality as the equivalent individual level 2 data products that constitute it. Both absolute agreement and long-term stability are excellent with respect to the ground-based data, for almost all latitudes apart from a few outliers which are mostly due to sampling differences between the level 2 and level 3 data. We conclude that the GTO-ECV data record is valuable for a variety of climate applications such as the long-term monitoring of the past evolution of the ozone layer, trend analysis and the evaluation of chemistry–climate model simulations.

1 Introduction

In 2010 the European Space Agency (ESA) set up the Climate Change Initiative (CCI) program, which aims to realize the full potential of long-term Earth observation data records for a number of Essential Climate Variables (ECVs) from the atmospheric, oceanic and terrestrial domains (Hollmann et al., 2013). These data records are essential to assess the state and future evolution of climate, as observations from space provide unique information and global coverage. However they are often limited by a lack of homogeneity and continuity. Therefore the aim of the ESA-CCI program is to provide stable and long-term Climate Data Records (CDRs) derived from multiple satellite data sets which are then suitable for both monitoring and modelling of climate and which meet the target requirements defined within the Global Climate Observing System (GCOS, 2011).

In this paper we focus on measurements of the ozone layer which protects life on Earth from harmful ultraviolet solar radiation and which plays an important role in the radiation budget of the atmosphere. As a consequence of the 1987 Montreal Protocol (UNEP, 1986) and subsequent phasing-out of the emissions of the ozone-depleting substances (ODSs) the stratospheric ozone layer is expected to recover within the next decades (WMO, 2011, 2014). However, significant uncertainty remains as to the timing of

Table 1. European total ozone satellite instrument characteristics.

Parameter	GOME	SCIAMACHY	GOME-2A
Data availability	06/1995–07/2011 ^a	08/2002–04/2012	01/2007–today
Organization	ESA	ESA	EUMETSAT ^b
Satellite	ERS-2	ENVISAT	MetOp-A
Spectral coverage	240–790 nm	240–2380 nm	240–790 nm
Spectral resolution	0.2–0.4 nm	0.2–1.5 nm	0.2–0.4 nm
PMD ^c coverage	3 p-PMD 300–800 nm	6 p-PMD 320–2380 nm	15 p-PMD and 15 s-PMD 310–790 nm
Viewing geometries	nadir	nadir, limb, occultation	nadir
Ground pixel size	320 × 40 km ²	60 × 30 km ²	40 × 80 km ²
Swath width	960 km	960 km	1920 km
Altitude	785 km	800 km	817 km
Equator crossing	10:30 a.m. LT ^d	10:00 a.m. LT	09:30 a.m. LT
Global coverage	3 days	6 days	almost daily ^e
Reference	Burrows et al. (1999)	Bovensmann et al. (1999)	Callies et al. (2000)

^a No global coverage since June 2003; ^b European Organisation for the Exploitation of METeorological SATellites; ^c PMD: polarization measurement device detecting polarized light perpendicular (p-) or parallel (s-) to the optical plane; ^d LT: local time; ^e until June 2013.

this recovery, because of complex interaction with climate change and continuously increasing emissions of greenhouse gases.

Within the phase I of the ESA's Ozone CCI (Ozone_cci) project, total ozone and ozone profile data records from nadir ultra-violet (UV) backscatter sensors, as well as ozone profiles from limb and occultation sensors (Sofieva et al., 2013), have been created. In this paper we introduce the multi-sensor total ozone data record which covers the period 1996–2011. The record is based on observations from three European instruments – all mounted on sun-synchronous low earth orbit platforms – namely the Global Ozone Monitoring Experiment (GOME) onboard the second European Remote Sensing satellite (ERS-2), the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartography (SCIAMACHY) onboard the ENVironmental SATellite (ENVISAT), and GOME-2 (referred to as GOME-2A in the following) onboard the first of a series of three Meteorological Operational satellites (MetOp-A). Detailed descriptions of the instruments are given in Burrows et al. (1999), Bovensmann et al. (1999), and Callies et al. (2000), respectively. A brief overview of the main platform and sensor characteristics is presented in Table 1. GOME data are available for July 1995 to June 2011, but their global coverage ended in June 2003 due to the permanent loss of the ERS-2 onboard data storage capability. As a consequence, the data coverage has been initially limited to the European and North Atlantic sector since only data within reach of an ERS-2 receiving station were transmitted to ground. Subsequently additional ground stations have been brought online and the data coverage has been incrementally increased. On 4 July 2011 the ERS-2 science mission ended. SCIAMACHY/ENVISAT

was launched in March 2002 and provided data from August 2002 to April 2012, which marks the end of the ENVISAT mission due to the unexpected loss of contact with the satellite.

As part of the Ozone_cci project, the total ozone data sets have been recently reprocessed with the retrieval algorithm GOME-type Direct FITting version 3 (GODFIT_V3) for the entire time series of the GOME, SCIAMACHY, and GOME-2A observations (Lerot et al., 2014). The GODFIT_V3 algorithm leads to high-quality retrievals in all conditions, including high solar zenith angles and large optical depths. Since the deployment of the GODFIT version used in the GOME Data Processor (GDP) operational ground segment version 5 (Van Roozendaal et al., 2012), a number of new developments in GODFIT have made the algorithm even more robust. These improvements are related to the tropospheric ozone content, a semi-empirical Ring correction and a correction for atmospheric polarization, as well as enhanced computational performance.

Although common retrieval settings are used for all three sensors, significant differences may appear when individual data sets are compared. These differences are largely due to calibration issues in the level 1 data. To improve inter-sensor consistency, a soft-calibration scheme for measured reflectances has been developed by Lerot et al. (2014), which relies on a statistical comparison of the level 1 sun-normalized radiances with simulated spectra at a few reference sites (viz., European stations equipped with Brewer spectrophotometers). The identification and correction of any artificial offset or spectral structures in the measured reflectances greatly improves the agreement between individual level 2 total ozone data sets. On the other hand, this pro-

cedure introduces a dependency of the satellite data on the observations from the Brewer instruments themselves, and this has to be kept in mind for the assessment of the geophysical validation results.

Together, these level 2 data sets based on the GODFIT_V3 retrieval algorithm span the time period 1996–2012. They have been recently validated, using ground-based measurements with Brewer and Dobson spectrophotometers as well as UV–visible DOAS/SAOZ (Differential Optical Absorption Spectroscopy/Système d'Analyse par Observation Zénithale) instruments (Koukouli et al., 2015) as a reference. The main findings were that the three sensors are characterized by similar patterns (such as seasonality and solar zenith angle dependence) against the reference data sets. No trends or unexplained jumps were detected. Furthermore, a marked improvement in quality with respect to the operational products was identified, along with an enhanced inter-sensor consistency.

Following the papers by Lerot et al. (2014) describing the retrieval algorithm itself, and by Koukouli et al. (2015) presenting the geophysical validation of the level 2 data, this paper is the third article on the ESA-CCI total ozone ECV. It describes the construction and validation of a cohesive merged level 3 data product. The aim is to show that the combination of the three individual homogenized total ozone data sets forms a consistent long-term time series, which meets the GCOS requirements and is therefore suitable for climate applications.

The paper is organized as follows: Sect. 2 contains a detailed description of the generation of the GTO-ECV CCI total ozone data record. Section 3 is dedicated to the validation of the level 3 merged product using ground-based measurement systems, and Sect. 4 shows the results of comparisons with two comparable satellite-based data records. Section 5 contains the summary and outlook.

2 Construction of GTO-ECV data record

In this section we describe the construction of the level 3 data set and the inter-satellite calibration approach, which has been developed and applied to combine the individual observations into a homogeneous long-term product. An analysis of issues related to spatial and temporal sampling is presented in Sect. 2.2. Section 2.3 contains a short description of the final output NetCDF (Network Common Data Form) files.

2.1 Level 3 algorithm description and merging approach

The level 3 algorithm is designed to map the level 2 measurements, processed with the GODFIT_V3 retrieval algorithm, onto a daily fixed global grid of $1^\circ \times 1^\circ$ in longitude and latitude. This spatial resolution has been selected according to

the user requirements defined for the ESA-CCI total ozone ECV product (van der A, 2011) which specify a horizontal resolution of 20–100 km. These requirements are based on the ozone requirements of GCOS, CMUG (Climate Modelling User Group), IGACO (Integrated Global Atmospheric Chemistry Observations), and the World Meteorological Organization (WMO). Each grid cell contains an average of all level 2 data from the same GMT (Greenwich Mean Time) day, that overlap with the level 3 cell. Cell values are computed as weighted averages in which the fractional area of overlap of the satellite ground pixel with the given grid cell is used as the weight. Level 2 data can be mapped onto more than one grid cell. The gridding algorithm is applied separately to GOME, SCIAMACHY, and GOME-2A measurements.

The next step is to merge the individual level 3 data sets from the three sensors into one homogeneous record using an inter-instrument calibration approach. Predecessors of this algorithm are described in Loyola et al. (2009a) and Loyola and Coldewey-Egbers (2012). We apply an external adjustment to SCIAMACHY and GOME-2A results with respect to the GOME results in order to account for inter-sensor differences, which possibly remain from the GODFIT_V3 level 2 algorithm, albeit these differences are small and the inter-sensor consistency is high (Lerot et al., 2014). Furthermore, all three individual data records exhibit good temporal stability, well within the GCOS target requirement of 1–3 % per decade (Koukouli et al., 2015). We selected the GOME data record to serve as the reference data base because it has the longest overlap periods with the other two sensors and, furthermore, it was found to be the most stable instrument over its lifetime before the application of the soft-calibration correction (Lerot et al., 2014).

The calculation of the correction factors is based on a comparison of 1° zonal monthly means, which are computed at first for GOME and SCIAMACHY. These zonal monthly means are based on common daily gridded data only in order to minimize the differences in spatial and temporal sampling. In particular this becomes important after June 2003, when GOME lost its global coverage. We did not consider diurnal changes of ozone in the merging approach, since all three instruments provide measurements within 1 h of each other (see Table 1). However, the peak-to-peak difference in total ozone may reach 1 % over the course of a day (Sakazaki et al., 2013).

The correction factors for SCIAMACHY with respect to GOME are derived using the ratios of these zonal monthly means. The correction factors comprise two parts: (1) a “basic” correction for each month of the year (averaged over all years from 2002 to 2011) in terms of third-order polynomials as a function of latitude, and (2) an offset for each individual month, which is added to the “basic” correction. This offset does not depend on latitude, but it accounts for the time-dependence (i.e. short-term fluctuations) in the differences between SCIAMACHY and GOME from 2002 to

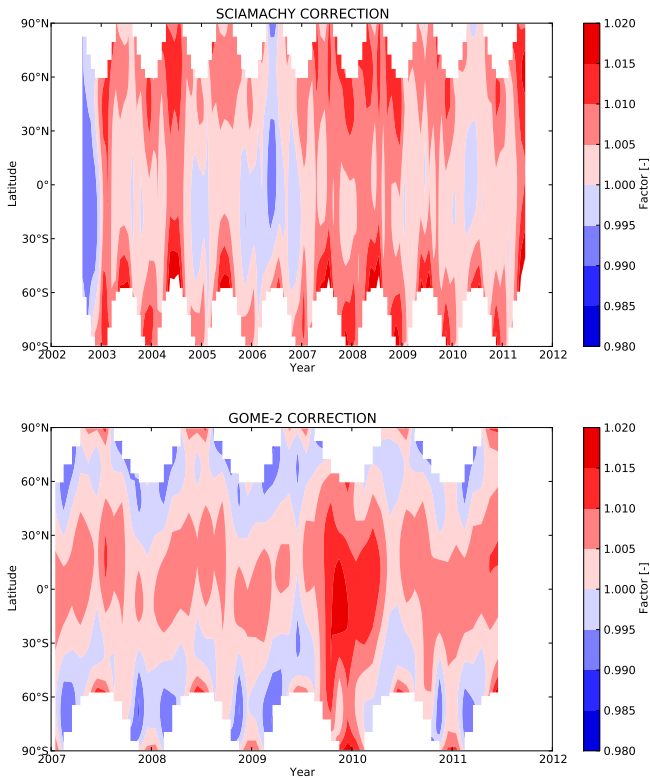


Figure 1. Correction factors for SCIAMACHY 2002–2011 (top panel) and GOME-2A 2007–2011 (bottom panel) as a function of latitude.

2011. The correction factors are then applied to the SCIAMACHY daily gridded data by linear interpolation in time. They are shown in the top panel of Fig. 1 as a function of latitude and time. The correction is well below 2% without obvious trends. It is between -0.5 and 1.0% in the tropical region and increases slightly toward higher latitudes.

In preparation for the GOME-2A adjustment, an intermediate product of averaged GOME and corrected SCIAMACHY daily gridded data is generated for the overlap period with GOME-2A from January 2007 to June 2011. This is referred to as GS_MERGED in the following. 1° zonal monthly means are computed for GS_MERGED and GOME-2A based on common daily gridded data only. The correction factors for GOME-2A with respect to GS_MERGED are derived similarly to those for SCIAMACHY: fourth-order polynomials as a function of latitude and month plus a time-dependent offset. They are applied to GOME-2A daily gridded data by linear interpolation in time. The bottom panel of Fig. 1 shows the GOME-2A correction, which is also well below 2%, as for SCIAMACHY, and without trends. It is between 0.0 and 1.0% in the tropics and decreases towards higher latitudes. In September 2009 the behaviour of the GOME-2A instrument changed owing to the second throughput test (Lacan and Lang, 2011). The soft-calibration scheme applied within the GODFIT_V3 re-

Table 2. Latitudinal and monthly constraints imposed on the GTO-ECV CCI monthly mean calculation.

Month	Latitude range
January	60.0° N– 90.0° S
February	70.0° N– 90.0° S
March	80.0° N– 80.0° S
April	90.0° N– 65.0° S
May	90.0° N– 60.0° S
June	90.0° N– 57.5° S
July	90.0° N– 57.5° S
August	90.0° N– 62.5° S
September	82.5° N– 72.5° S
October	72.5° N– 85.0° S
November	65.0° N– 90.0° S
December	60.0° N– 90.0° S

trieval algorithm mitigates the long-term impact of this test, so that only an insignificant increase in the correction factors can be identified. Furthermore, no unexpected jumps with respect to ground-based data were found during the geophysical validation exercise of the level 2 data (Koukouli et al., 2015). Only a small increase in the correction factors for a limited period in time is visible, which is caused by the low time resolution of 1 year for the soft-calibration scheme.

Once SCIAMACHY and GOME-2A data have been adjusted, $1^\circ \times 1^\circ$ monthly mean gridded data are computed for each instrument. In order to provide representative monthly means that contain a sufficient number of measurements equally distributed over time, cut-off values for latitude as a function of the month have been defined (see Table 2). Thereby we avoid calculating monthly averages based on a small number of measurements at the beginning or end of a month which appear close to the polar night. Nonetheless, differences in monthly means among the instruments may occur due to regular differences in spatial and temporal sampling (see Table 1). This will be discussed in the next subsection.

Subsequently, the three data sets are combined into one single record as follows: only one instrument is used at a time, i.e. the merged GTO-ECV total ozone time series contains GOME measurements from March 1996 to March 2003, adjusted SCIAMACHY measurements from April 2003 to March 2007, and adjusted GOME-2A measurements from April 2007 to June 2011. We decided not to include GOME data after the onboard tape recorder failure because of the very limited spatial coverage. Furthermore we omit SCIAMACHY data after the start of the GOME-2A record since a significant increase in data coverage and, hence, a reduction in sampling uncertainty is not expected. The whole procedure is summarized in Fig. 2. Red-shaded boxes denote data records which are part of the official ESA Ozone CCI Climate Research Data Package (CRDP). Blue-shaded boxes denote intermediate data sets needed to create

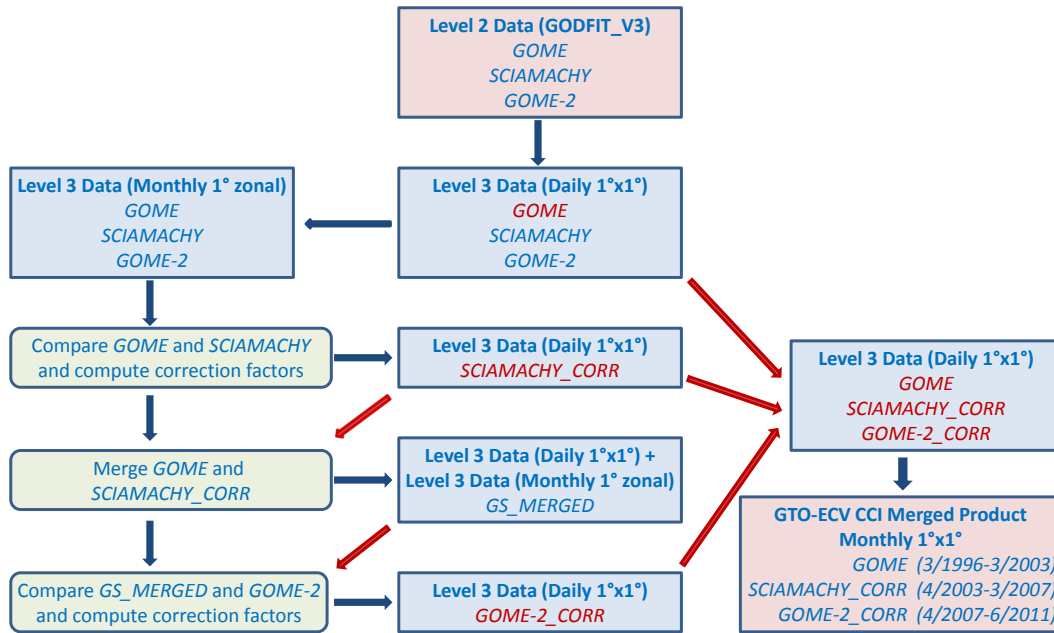


Figure 2. Flow chart of the GTO-ECV CCI level 3 algorithm and merging approach. Red-shaded boxes denote data records which are part of the official ESA Ozone CCI Climate Research Data Package (CRDP). Blue-shaded boxes denote intermediate data sets needed to create the merged final product, and green shading denotes the steps of the merging approach.

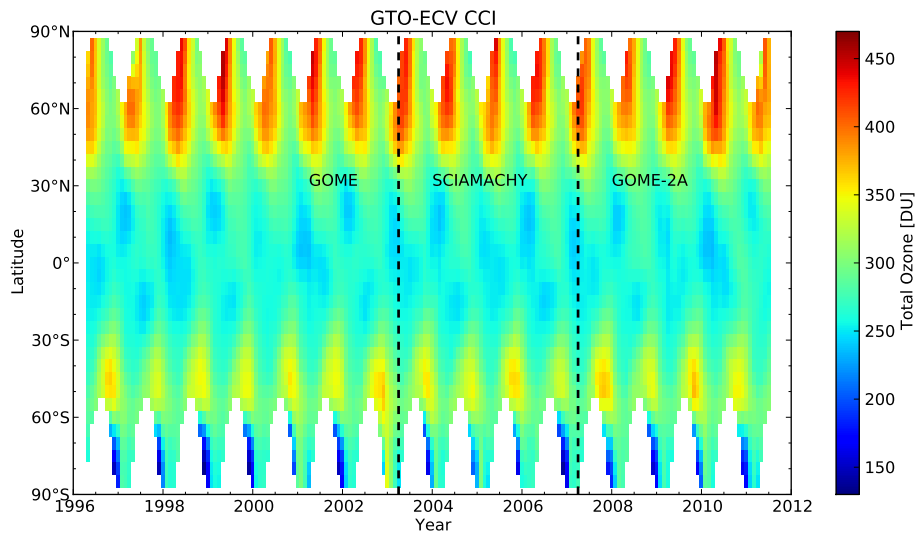


Figure 3. GTO-ECV CCI total ozone column data record 1996–2011 as a function of latitude. Vertical black dashed lines indicate the change-over from GOME to SCIAMACHY in April 2003 and from SCIAMACHY to GOME-2A in April 2007.

the merged final product and green shading denotes the three steps of the merging approach.

The complete data record with typical total ozone features is shown in Fig. 3. Highest ozone values occur in northern hemispheric springtime, whereas monthly mean values are below 200 DU from September to November south of 70° S. Extreme events such as the anomalous Antarctic ozone hole in 2002 and the severe ozone loss in 1997 and 2011 in the Arctic are visible. Instrument switches from GOME

to SCIAMACHY in April 2003 and from SCIAMACHY to GOME-2A in April 2007 are indicated with the black vertical bars.

2.2 Illustration of sampling issues

As already noted in the previous section, inhomogeneous or incomplete sampling – intrinsic to these types of satellite sensors – may have systematic effects and may therefore lead

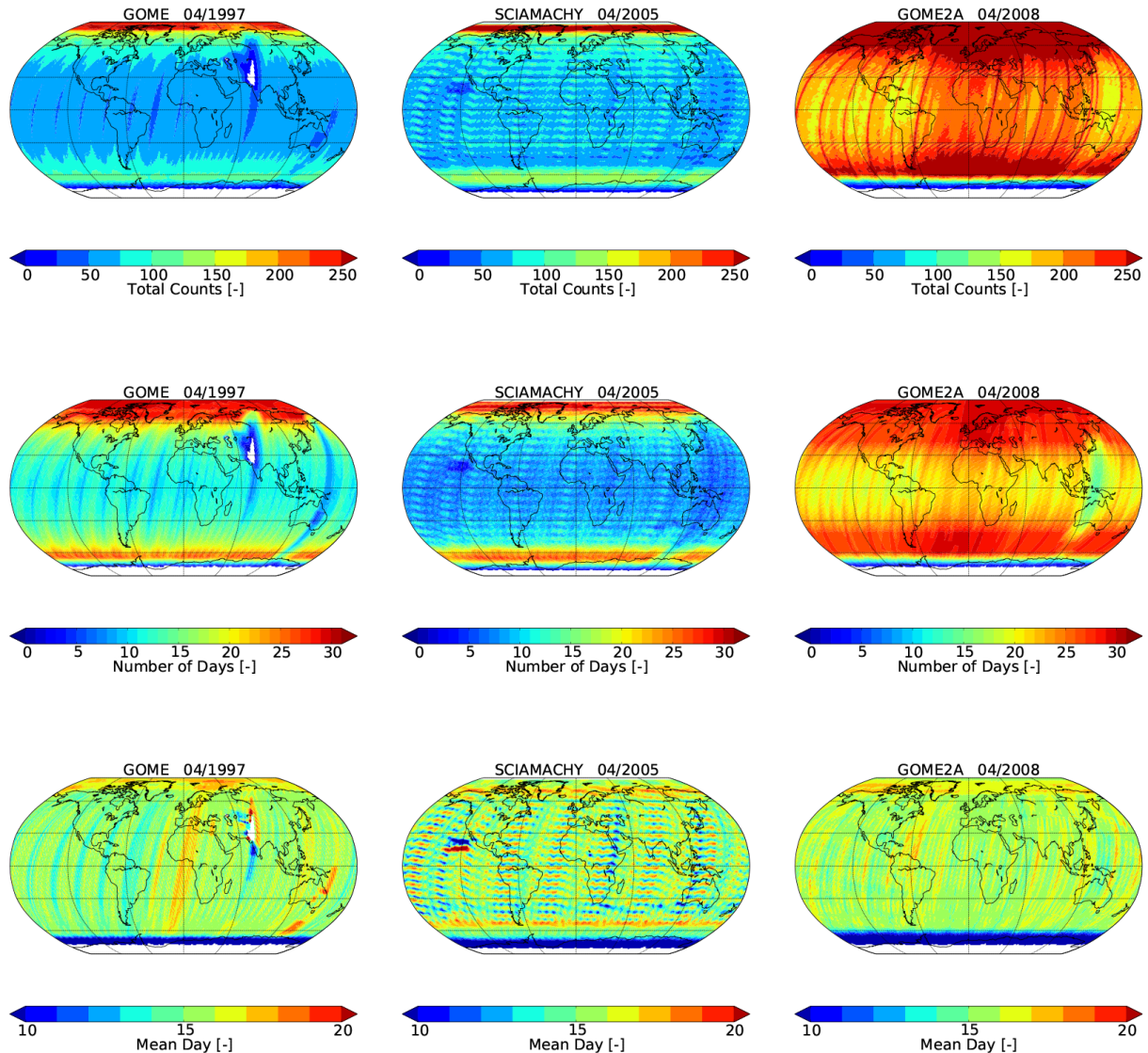


Figure 4. Sampling patterns of GOME (left column), SCIAMACHY (middle column), and GOME-2A (right column) exemplified for April 1997, 2005, and 2008, respectively. Top row: total number of measurements per month and grid cell, middle row: number of days for which measurements are available, and bottom row: effective mean day d_{eff} representing the monthly mean according to Eq. (1).

to erroneous average estimates (e.g. Sofieva et al., 2014). Since the platforms are in polar orbits, for each day there are coverage gaps in the tropics (even for GOME-2A, which has the largest swath width) as well as repeated views of the summertime poles, leading to non-uniform undersampling or oversampling of ozone. This can result in inaccurate monthly average estimates, in particular when natural variability is strongest, i.e. in spring months in the Northern Hemisphere or under ozone hole conditions. The problem is exacerbated when the satellites sample only a few days at the beginning or end of the month owing to the beginning or end of the polar night.

Figure 4 exemplifies the diverse sampling patterns of GOME (left column), SCIAMACHY (middle column), and

GOME-2A (right column) for April 1997, 2005, and 2008, respectively. The total number of measurements per month and grid cell, i.e. mapped level 2 data according to the level 3 algorithm described above, are shown in the top row, the number of days for which measurements are available is indicated in the middle row, and the effective mean day d_{eff} representing the monthly mean is found in the bottom row. The latter has been calculated using

$$d_{\text{eff}} = \frac{\sum_{d=1}^D d \cdot n_d}{\sum_{d=1}^D n_d} \quad (1)$$

D is the maximum number of days in the month, i.e. 31 in January, 30 in April, etc., and n_d is the number of measurements per day and grid cell.

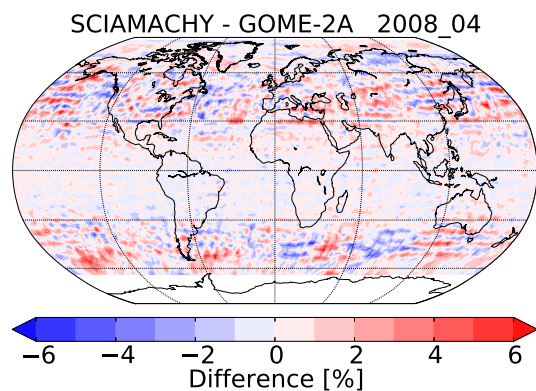


Figure 5. Percentage differences between SCIAMACHY and GOME-2A monthly mean total ozone for April 2008.

GOME-2A has the densest and most uniform sampling, i.e. the highest number of measurements (top right panel). The effective day is close to the middle of the month (between day 14 and 16, bottom right panel), although some longitudinal structures are visible in particular in the tropics. The GOME sampling is less dense and the effective mean day shows a larger spread around the middle of the month as well as pronounced longitudinal structures in low and middle latitudes (bottom left panel). The sampling pattern of SCIAMACHY strongly reflects the alternation of the nadir and limb measurement modes for this instrument, leading to extreme longitudinal as well as latitudinal structures (middle panels).

Toward the north polar regions (in April) the number of measurements increases due to overlapping orbits and hence multiple views per day. Toward the south polar regions the number of measurements increases up to about 65° S and then rapidly decreases due to the beginning polar night. The effective mean day (bottom row) indicates that only the first half of April is sampled. We decided to exclude these regions close to the polar night from the level 3 data record. Therefore, we defined cut-off latitudes (see Table 2) for each month in order to avoid using data covering only a limited part of the month.

Figure 5 illustrates the impact of the diverse sampling patterns on the monthly averages. It shows the percentage differences between SCIAMACHY (sparse sampling) and GOME-2A (dense and most uniform sampling) monthly mean total ozone for April 2008. Biases of $\pm 5\%$ reflect the differences in the sampling patterns, in particular in the middle latitudes, where natural variability is strong in this month. It is less pronounced in the tropics, where variability is low, and in the north polar region, where the SCIAMACHY sampling is enhanced due to overlapping orbits and, thus, multiple views per day. These sampling issues will be addressed in more detail in the second phase of the Ozone_cci project.

Table 3. Description and dimensions of all variables contained in the level 3 monthly mean total ozone NetCDF files. $N_{\text{lat}} = 180$ and $N_{\text{lon}} = 360$.

Description of variable	Unit	Dimension
Latitude of grid center	degree	N_{lat}
Longitude of grid centre	degree	N_{lon}
Mean total ozone column	DU	$N_{\text{lat}} \times N_{\text{lon}}$
Standard deviation of mean total ozone column	DU	$N_{\text{lat}} \times N_{\text{lon}}$
Standard error of mean total ozone column	DU	$N_{\text{lat}} \times N_{\text{lon}}$
Number of measurements used to derive the monthly mean total ozone column	–	$N_{\text{lat}} \times N_{\text{lon}}$

2.3 GTO-ECV data files

The final GTO-ECV CCI total ozone monthly mean output data are stored in NetCDF files (one file per month), which are publicly available via www.esa-ozone-cci.org. All files follow the NetCDF Climate and Forecast (CF) metadata convention version 1.5. Table 3 gives an overview of the content of the individual files. The reported grid of the data record is $1^{\circ} \times 1^{\circ}$ in longitude and latitude, i.e. the dimensions are 360×180 and the centre of the first grid cell is located at latitude 89.5° N and longitude 0.5° E. Besides the mean total ozone column, the corresponding standard deviation (SD), the standard error, and the number of measurements per month are provided. The sample standard deviation is the standard deviation of the monthly mean obtained from the daily gridded values. It characterizes the scatter of the measured data encompassing the natural variability, the measurement error as well as the sampling uncertainty. The standard deviations of the GTO-ECV product are compared with those from another satellite-based data record in Sect. 4.2. The standard error (SE), however, quantifies the spatial-temporal sampling errors inherent to the satellite measurements. These errors have been estimated using the aforementioned standard deviation (SD) and the number of available measurements per grid cell (N_{meas}) according to

$$\text{SE} = \frac{\text{SD}}{\sqrt{N_{\text{meas}}}} r. \quad (2)$$

The factor r has been obtained using an Observing System Simulation Experiment (OSSE) for which high-resolution ECMWF (European Centre for Medium-Range Weather Forecasts) data were taken as the reference data set. Then, three sets of daily observations were simulated from the reference using the sampling patterns appropriate to GOME, SCIAMACHY, and GOME-2A, respectively. Finally, the average monthly simulations are compared with the corresponding monthly reference in order to estimate the sampling errors corresponding to the total ozone monthly averages. The standard error is shown in Fig. 6 for April 1997

(GOME, top panel), 2005 (SCIAMACHY, middle panel), and 2008 (GOME-2A, bottom panel). The errors increase from the tropics to higher latitudes following the increasing ozone variability. GOME errors are larger than those for SCIAMACHY and GOME-2A due to the much larger ground-pixel size (see Table 1). The SCIAMACHY errors reflect the sampling pattern seen in Fig. 4, middle column, with latitudinal and longitudinal variance. GOME-2A errors are quite small and do not have noticeable structures.

3 Ground-based validation

The validation of level 2 satellite total ozone columns using independent ground-based observations has been a substantial part of retrieval algorithm development for many decades. A well-established procedure exists in assessing the level 2 total ozone products using global ground-based Brewer, Dobson, and UV-visible SAOZ spectrophotometer measurements (e.g. Balis et al., 2007b; Loyola et al., 2011; Koukouli et al., 2012; Labow et al., 2013, and references therein). Taking into account that the long-term climate study of the total ozone atmospheric content is based on using level 3 gridded products, one must ensure that the transition from level 2 to level 3 does not introduce artifacts. These might be induced by the level 3 algorithm itself, mainly through sampling issues which could lead to inaccurate average estimates, or by the merging approach through improper inter-sensor calibration. The aim of the following section is to compare the current level 2 validation of the individual satellite GOME, SCIAMACHY, and GOME-2A GODFIT_V3 products with the new level 3 GTO-ECV CCI integrated long-term record of total ozone on a global scale.

3.1 Representativeness of the ground-based network

The representativeness of the ground-based reference network used to validate a product with global coverage determines both the validation approach, and the representativeness of the validation results. While validation results (and the level 3 data themselves) are often shown and used as zonal averages, e.g. plotted against time and latitude as in Fig. 3, Fig. 7 illustrates the significant spatial representativeness error when comparing zonal means of global gridded data with zonal means based on the limited geospatial coverage of the ground-based network.

For this figure, IFS-MOZART (Integrated Forecasting System – Model for OZone And Related chemical Tracers) modelled fields (Inness et al., 2013) were averaged to zonal monthly means, either using all data or using only data coincident (in geolocation) with the Dobson, Brewer, and SAOZ instruments. The relative difference between these two simulated zonal means yields estimated spatial representativeness errors. As these errors exceed the expected performance of the level 3 product, the validation work presented

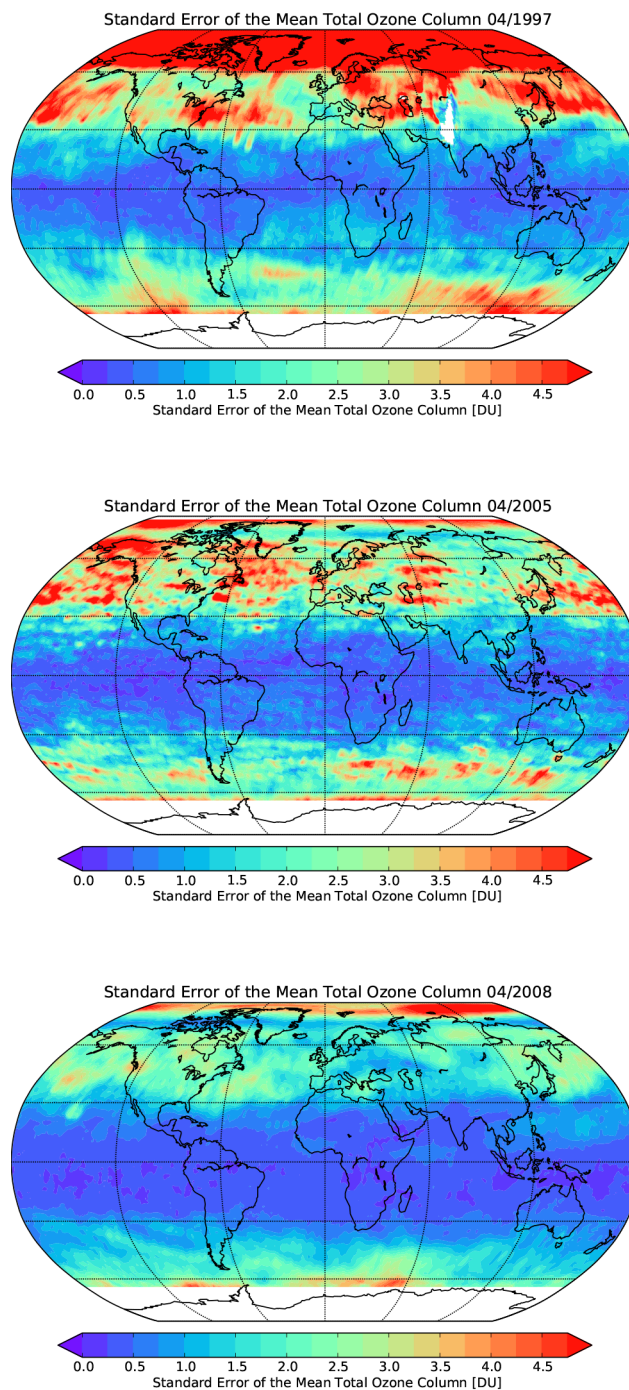


Figure 6. Standard error associated with the product for GOME (top), SCIAMACHY (middle), and GOME-2A (bottom) for April 1997, 2005, and 2008, respectively.

here is based solely on level 3 grid-cells co-located with the ground stations, and on zonal statistics derived from those co-locations. Besides avoiding the spatial representativeness error, this approach allows for a more direct comparison with the validation results of the level 2 data sets. However, it must be kept in mind that this validation strategy is blind

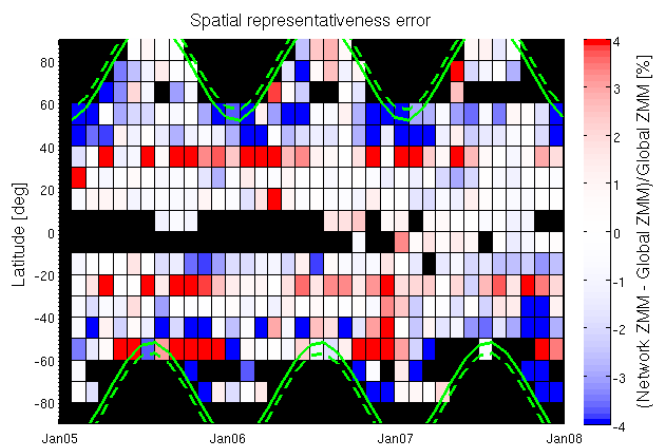


Figure 7. Simulated differences between zonal means based either on data coincident (in geolocation) with ground-based reference instruments, or on full global gridded data. These differences constitute the so-called spatial representativeness error. Data used for this graph are 6-hourly modelled fields calculated with IFS-MOZART for MACC (Inness et al., 2013). The green solid and dashed lines correspond to 75 and 80° solar zenith angles at noon, respectively.

to the product quality outside of the ground network. This issue is tackled by comparing the product with other satellite data sets in Sect. 4. Temporal representativeness errors, due to limited numbers of measurements within each month at a given station, are minimized in the following by requiring at least 10 measurements per month for an accepted collocation. In view of the temporal sampling issues known to be present in the level 3 data set (see Sect. 2.2), no attempt was made here to further characterize the errors due to limitations in temporal sampling of the reference measurements.

3.2 Comparison with Dobson and Brewer measurements

The Brewer and Dobson spectrophotometer measurements, as extracted from the World Ozone and UV radiation Data Center (WOUDC, <http://www.woudc.org>) have already been used numerous times in the last 2 decades for the validation of various satellite-based global total ozone records (e.g. Loyola et al., 2011; Koukouli et al., 2012; Labow et al., 2013, and references therein). A comprehensive description of the individual station selection criteria has been presented in Balis et al. (2007a, b). Station selection updates may be found in more recent papers listed above. The measurements involved in this current study are the same as those used and discussed in the companion level 2 validation paper by Koukouli et al. (2015) in which all level 2 comparisons shown in the following are discussed.

For comparison consistency, the Dobson–Brewer WOUDC ground-based data set was transformed into a monthly level 3 field in order to match the $1^\circ \times 1^\circ$ grid of the GTO-ECV CCI data. Measurements from all stations

were gridded in the same latitude–longitude boxes with some specific considerations. First, only the direct sun observations were used. Even though in some cases, as is shown in the subsequent figures, this severely decreases the number of measurements, after rigorous testing it was found that the usage of direct sun ground-based observations ensures an optimal level 3 ground-based product. Secondly the threshold on the number of measurements available before the computation of the associated monthly mean was investigated. As a compromise between obtaining the highest global coverage possible and the most representative monthly means, especially at high latitudes, a lower limit of 10 measurements per month and grid box was imposed.

The validation of the GTO-ECV CCI level 3 product against the Dobson and Brewer network is presented here as a series of comparative figures: in each plot, four lines are presented, namely the level 3 comparison (in dark blue) and three level 2 comparisons for GOME (in light blue), SCIAMACHY (in green), and GOME-2A (in red), respectively. In order to compare as closely as possible the same validation results for level 3 and level 2, a time constraint was imposed on the level 2 comparisons according to the time periods for each instrument in the merged data record (see Sect. 2.1). Furthermore, the same latitudinal constraints for the monthly means were imposed (see Table 2).

Figure 8 shows the latitudinal dependency of the percentage differences for both Brewer (left) and Dobson (right) instrument types. The three satellite instruments reveal a remarkable inter-sensor consistency for all latitudes and an excellent agreement with the ground data. The level 3 comparison (blue) closely follows that for level 2. The slight positive deviation of about 0.5 % of level 3 data (compared to level 2) for the 40–60° N belt (right panel) will be discussed in the next section.

3.2.1 Northern Hemisphere statistics

The Northern Hemisphere time series comparisons are shown separately for the Brewer and Dobson instrument types in Fig. 9. The Brewer comparisons (left panel) show very good agreement between level 3 and individual level 2 lines, well within the $\pm 1\%$ difference level for most of the 15-year data record and with negligible bias. The two outliers during the GOME period and the two during the SCIAMACHY period are discussed below. The Dobson analysis (right panel) shows equally good comparisons, falling within the 1.5 % difference level with a bias of $\sim 1\%$, due to the known differences in the treatment of the stratospheric temperature dependence of the ozone absorption cross sections and how this issue is dealt with by the ground-based algorithm (Van Roozendaal et al., 1998; Scarnato et al., 2009, and references therein). Koukouli et al. (2015, their Table IV) have shown that no long-term drift in the individual level 2 data sets was found for both Dobson and Brewer comparisons. For the corresponding level 3 comparisons in the

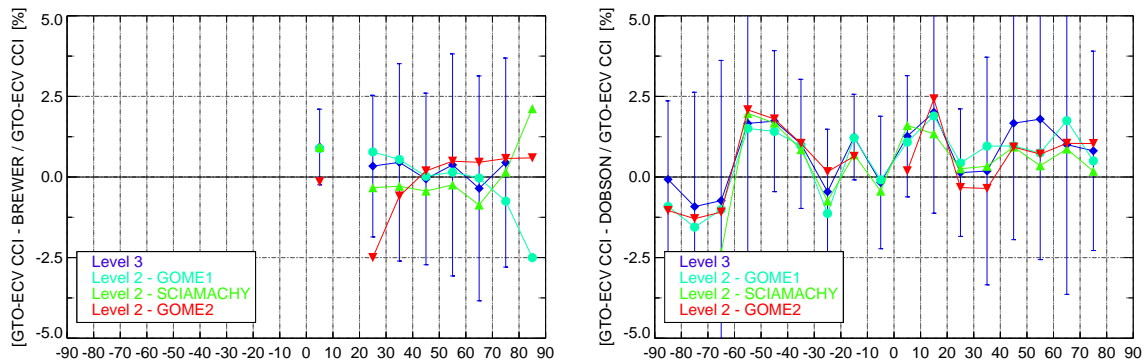


Figure 8. Percentage difference between satellite data records and ground-based data as a function of latitude. Left: Brewer comparisons and right: Dobson comparisons. Level 3 comparison in dark blue, GOME level 2 comparison in light blue, SCIAMACHY in green, and GOME-2A in red. The $1\text{-}\sigma$ standard deviation of the average is only given for the level 3 lines.

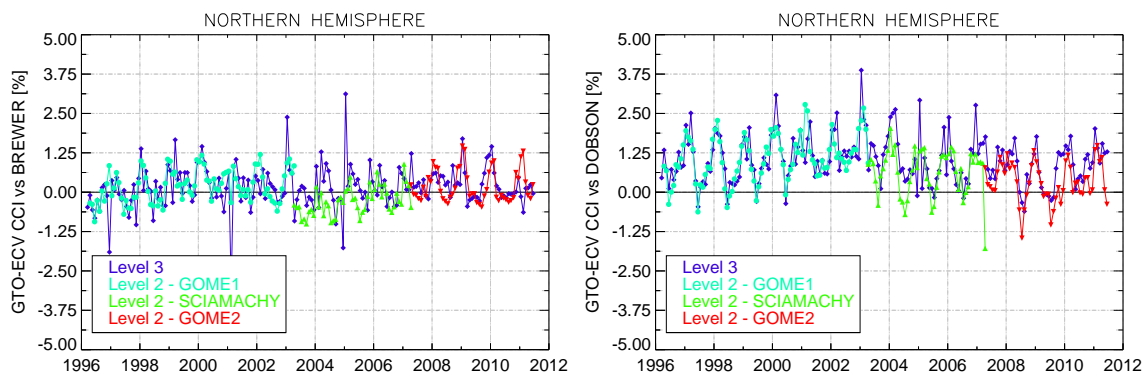


Figure 9. Percentage difference between satellite data records and ground-based data as a function of time for the Northern Hemisphere. Left: Brewer comparisons and right: Dobson comparisons. Level 3 comparison in dark blue, GOME level 2 comparison in light blue, SCIAMACHY in green, and GOME-2A in red.

Northern Hemisphere, the drift (per decade) of the differences with respect to ground-based data is also negligible, i.e. -0.12 ± 0.12 and 0.17 ± 0.11 %, respectively. This makes the GTO-ECV CCI data record exceedingly useful for longer-term analysis of the ozone layer such as decadal trend studies (e.g. Coldewey-Egbers et al., 2014).

Figure 10 shows the percentage differences as a function of time for six zonal belts $0\text{--}10$, $30\text{--}40$, $40\text{--}50$, $50\text{--}60$, $60\text{--}70$, and $70\text{--}90^\circ$ N (from top to bottom) for the Brewer comparisons (left) and the Dobson comparisons (right). The patterns of the level 3 comparison are nearly identical to those from the individual level 2 comparisons. The agreement for both types of instruments is excellent up to high latitudes, except for a small number of outliers in the $60\text{--}70^\circ$ N belt for the Brewer comparisons and in the $50\text{--}60^\circ$ N belt for the Dobson comparisons. For this latter belt, some strong disagreement up to 5–10% between level 3 and level 2 coincidences is shown for the SCIAMACHY period. These outliers, during years 2004 and 2005, were basically due to sampling issues. A different set of days was considered for creating the monthly mean differences for the level 2 data set and the level

3 data set, due to the 6-day SCIAMACHY global coverage and the scarcity of ground-based stations in those latitudes. We have to keep in mind that the level 2 comparisons are based on coincident measurements with respect to geolocation (150 km radius) and time (same day), whereas the level 3 comparisons are based on coincident measurements with respect to geolocation (same $1^\circ \times 1^\circ$ grid box) only. Therefore, a different set of days might form the basis for the level 3 monthly averages from ground-based and satellite-based data, respectively. Consequently, these larger differences do not necessarily indicate poorer quality of the level 3 data record. A similar reason explains the outliers noted in the $60\text{--}70^\circ$ N belt for the Brewer comparisons. Furthermore, for the high latitude belts it is possible that we include comparisons with one ground-based station alone. Overall, considering the excellent agreement for the remainder of the belts, the consistency between the level 2 and level 3 validation results is very satisfactory. As for the entire Northern Hemisphere statistics (see Fig. 9) no long-term drift in the differences is found for the individual latitude belt statistics.

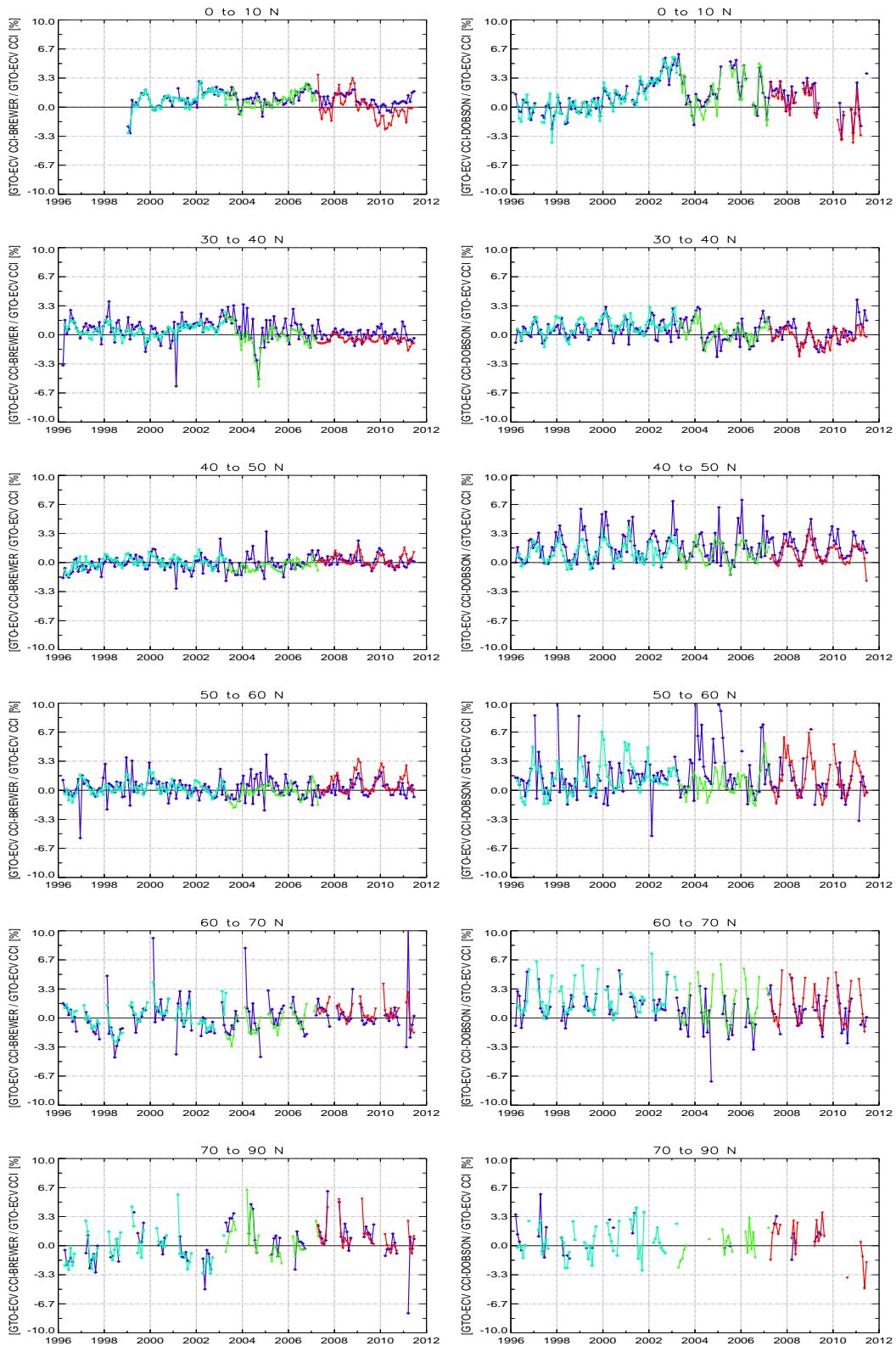


Figure 10. Percentage difference between satellite data records and ground-based data as a function of time for the Northern Hemisphere for six zonal belts 0–10, 30–40, 40–50, 50–60, 60–70, and 70–90° N from top to bottom. Left column: Brewer comparisons and right column: Dobson comparisons. Level 3 comparison in dark blue, GOME level 2 comparison in light blue, SCIAMACHY in green, and GOME-2A in red.

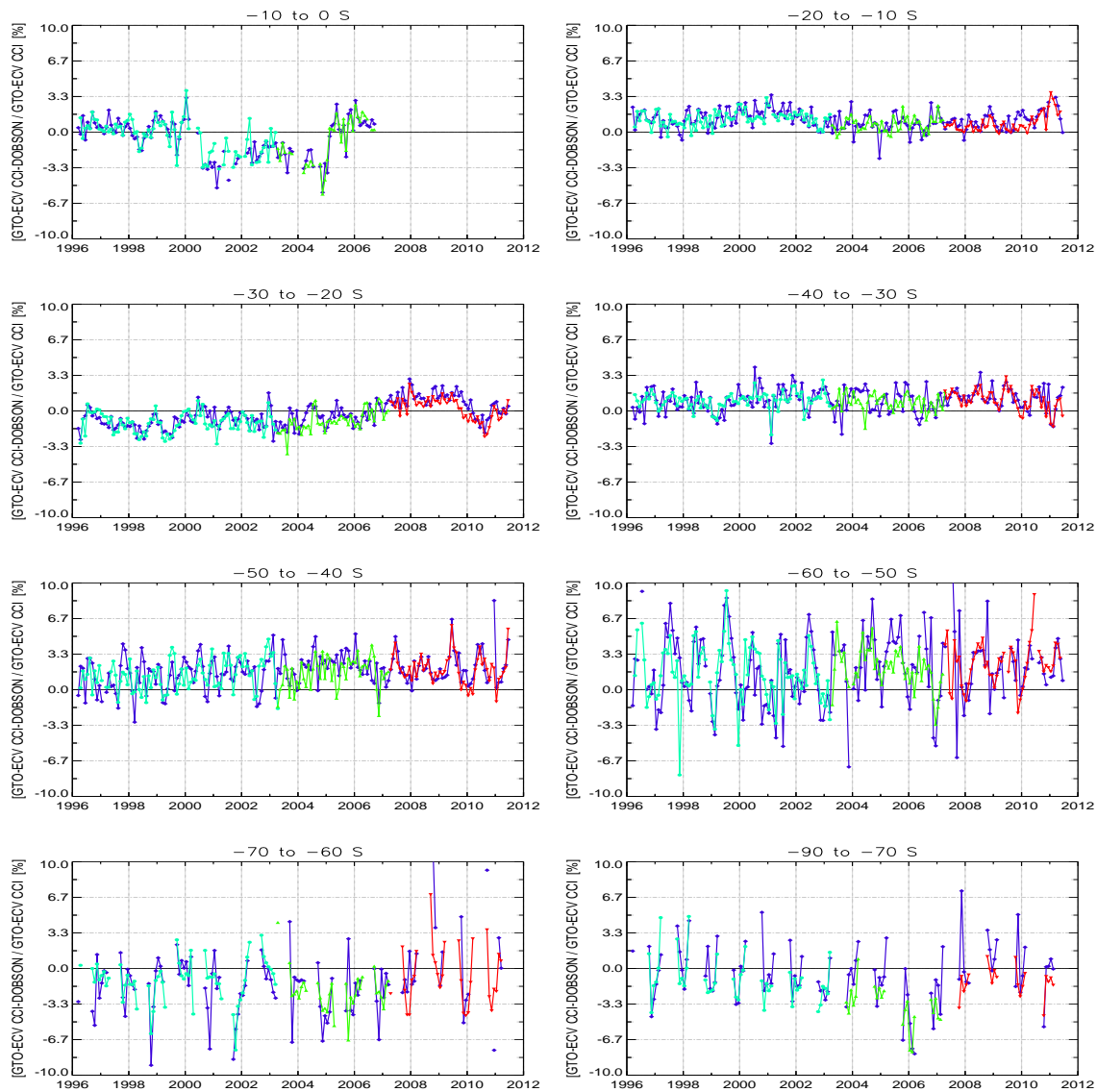


Figure 11. Percentage difference between satellite data records and Dobson ground-based data as a function of time for the Southern Hemisphere for seven 10° zonal belts from $0\text{--}70^\circ\text{S}$ and one belt from $70\text{--}90^\circ\text{S}$. Level 3 comparison in dark blue, GOME level 2 comparison in light blue, SCIAMACHY in green, and GOME-2A in red.

3.2.2 Southern Hemisphere statistics

In the Southern Hemisphere, the validation is restricted to Dobson measurements. Figure 11 shows the percentage differences between satellite and ground-based data as a function of time for seven 10° belts from $0\text{--}70^\circ\text{S}$ and one belt from $70\text{--}90^\circ\text{S}$ (top to bottom). As for the Northern Hemisphere, the level 3 comparisons show a near-perfect agreement with the level 2 comparisons up to 50°S . The outliers in higher latitudes are mostly due to differences in sampling as explained in the previous section. The mean bias between GTO-ECV CCI level 3 data and the Dobson ground-based network in the Southern Hemisphere is $0.66 \pm 1.63\%$ and the drift per decade is $0.77 \pm 0.12\%$.

3.2.3 Seasonal and latitudinal dependence

The seasonal variability of the GTO-ECV CCI data compared to the Dobson network is shown in Fig. 12 as a contour plot of latitude vs. month of year. Very small seasonal features are observed with a slight oscillation of $\pm 1\%$. For the very high southern latitudes some underestimations are seen for the summer months (around -2 to -2.5%) and overestimations for the winter months (around $+3$ to $+4\%$). This seasonality probably originates from the Dobson sensitivity to atmospheric effective temperature, which leads to positive differences between Dobson and satellite observations for high effective temperatures in local summer (negative differences in winter). For the Brewer stations no significant fea-

Table 4. The statistics following the Figures presented in Sect. 3.2 for the Northern Hemisphere.

GTO-ECV CCI	Dobson	Brewer
Mean bias	$1.00 \pm 0.75 \%$	$0.16 \pm 0.66 \%$
Monthly mean variability	$\pm 2.35 \%$	$\pm 2.20 \%$
Drift per decade	$-0.12 \pm 0.12 \%$	$0.17 \pm 0.11 \%$
Seasonal variation of biases	$0.80 \pm 0.21 \%$	$0.16 \pm 0.30 \%$
Latitudinal variation of biases	$0.72 \pm 0.96 \%$	$0.30 \pm 0.41 \%$

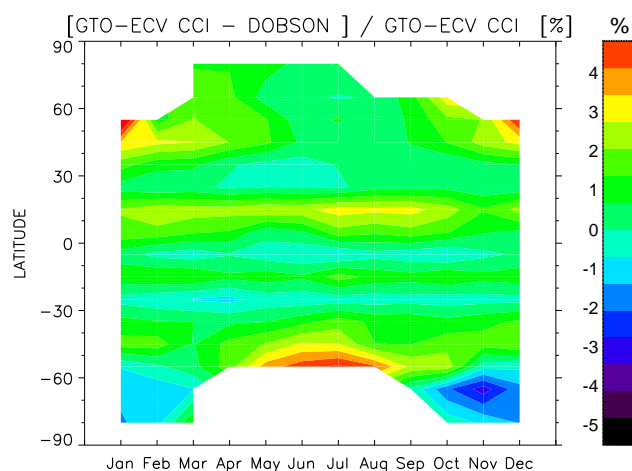
tures are observed in the contour comparison; the seasonal variation of biases is $0.16 \pm 0.30 \%$ and the latitudinal variation of biases is $0.30 \pm 0.41 \%$.

3.2.4 Summary of the Brewer and Dobson comparisons

In conclusion, the GTO-ECV CCI level 3 validation results were found to be very consistent with the separate GOME, SCIAMACHY, and GOME-2A level 2 validation comparisons. In particular, on a monthly mean basis, for the Dobson comparisons, both the Northern and Southern Hemisphere time series are in very close agreement. Similarly, for the Brewer comparisons (Northern Hemisphere), an excellent agreement is found apart from a handful of outliers. On a seasonal basis, both the Brewer and the Dobson level 3 comparisons show close agreement with the level 2 comparisons.

According to Table 5 of the Ozone_cci User Requirement Document (van der A, 2011) it is stated that the decadal stability of the total ozone column provided by the three instruments must fall within 1–3 %, the long-term accuracy of each product at 2 % and short term accuracy at 3 %. The seasonal cycle and inter-annual variability must also fall within the 3 % level. In Table 4, the statistics extracted from the Dobson and Brewer comparisons for the Northern Hemisphere are summarized. Under the header “mean bias” we refer to the mean bias and standard deviation ($1-\sigma$) of the time series (see Fig. 9). It is $1 \pm 0.75 \%$ for the Dobson comparisons and $0.16 \pm 0.66 \%$ for the Brewer comparisons, respectively. The header “monthly mean variability” refers to the standard deviation of the standard deviations of the monthly mean values in the Northern Hemisphere time series. The header “drift per decade” refers to the decadal drift and drift error calculated from the Northern Hemisphere time series (Fig. 9); the header “seasonal variation of biases” indicates the mean difference from the seasonal plots (see Fig. 12) and the amplitude of the seasonal variability. The header “latitudinal variation of biases” refers to the mean bias and standard deviation as calculated by the latitudinal variability plots on a global scale.

It is evident that the product easily meets the User requirement levels listed above. Hence, we can conclude that the current GTO-ECV CCI level 3 total ozone product is of the same high quality as the constituent level 2 total ozone products. As the relative drift compared to the ground-based ref-

**Figure 12.** Seasonal variability of the GTO-ECV CCI data compared to the Dobson network as a contour plot of latitude vs. month of the year.

erence is less than 1 % per decade, the GTO-ECV data record will be useful for studies of long-term total ozone trends.

3.3 Comparison with SAOZ UV–visible instruments

The NDACC (Network for the Detection of Atmospheric Composition Change, <http://www.ndacc.org>) UV–visible working group operates about 35 certified SAOZ zenith-sky UV–visible absorption spectrometers (Pommereau and Goutail, 1988) distributed from the Arctic to the Antarctic. Most of the instruments perform twice-daily measurements of the total ozone column during twilight between 86 and 91° solar zenith angle at all latitudes and seasons. The retrieval is based on the DOAS approach in the visible Chappuis band of ozone between 470 and 540 nm.

Figure 13 shows time series of monthly mean differences between GTO-ECV CCI level 3 data and the UV–visible network grouped by latitude zones of 30° . Red dots correspond to comparisons for single stations and the white-faced red circles represent the mean of those differences over all stations within a given latitude zone. For belts $0\text{--}30^\circ$ N (bottom left panel) and $30\text{--}60^\circ$ S (middle panel on the right) only one station contributes data for the better part of the time series, and the zonal mean therefore coincides with the station’s difference.

These comparisons with UV–visible instruments in general confirm the validation results based on Dobson and Brewer comparisons. Large discrepancies are evident in the southernmost bin, in particular during Antarctic ozone hole conditions. These are in large part due to co-location space–time mismatches and differences in horizontal smoothing of the large gradients occurring at the border of the polar vortex (Verhoelst et al., 2015). The positive bias observed in the northernmost bin, which is not seen in the comparisons with Brewer observations, is noteworthy. While the GODFIT_V3

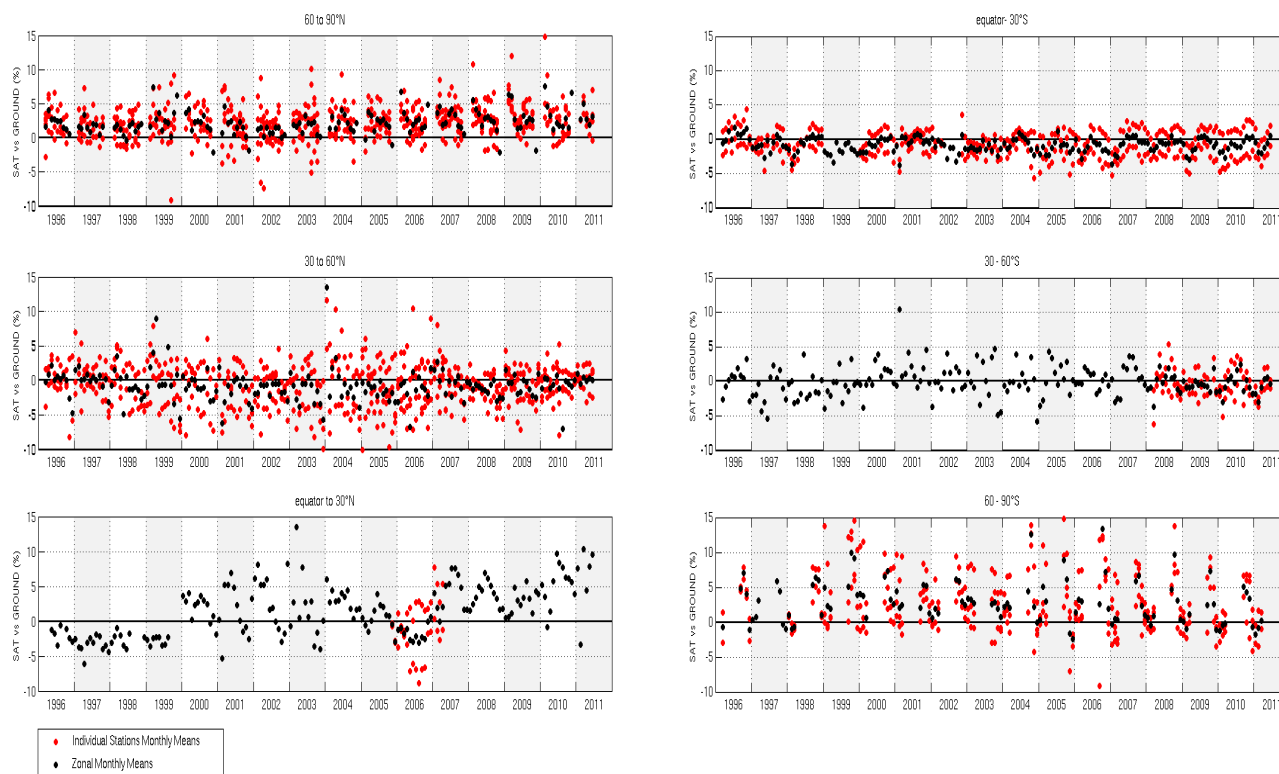


Figure 13. Time series of monthly mean relative differences for NDACC UV-visible instruments for six 30° latitude zones; Northern Hemisphere in the left panels (from top to bottom: high, middle, and low latitudes), and Southern Hemisphere in the right panels (from top to bottom: low, middle, and high latitudes). Red dots correspond to individual stations, black dots correspond to the zonal means. If only one station contributes, the single-station differences are coincident with the zonal mean.

retrieval uses more recent ozone cross sections than those used in the default Brewer data processing, the good agreement between the GTO-ECV CCI total ozone column level 3 product and the Brewer observations should be interpreted with care as the GODFIT_V3 uses a soft-calibration scheme based on total ozone measurements obtained with Brewer measurements at a set of northern mid-latitude reference sites (Lerot et al., 2014). As such, the accuracy of the GTO-ECV CCI level 3 product somehow depends on that of the Brewer network. On the other hand, as this positive bias between the GTO-ECV product and the SAOZ instruments only appears at high latitudes, errors in the SAOZ AMFs cannot be ruled out either.

4 Comparison with other satellite data

In this section the GTO-ECV CCI level 3 monthly mean total ozone product is compared with two other satellite-based data records: (1) its predecessor product GTO-ECV GDP and (2) the SBUV version 8.6 merged ozone data record.

4.1 GTO-ECV GDP

The preceding GTO-ECV GDP data record (Loyola et al., 2009a; Loyola and Coldewey-Egbers, 2012) is based on GOME, SCIAMACHY, and GOME-2A total ozone columns obtained with the GDP 4.X retrieval algorithm (Van Roozendael et al., 2006; Lerot et al., 2009; Loyola et al., 2011; Hao et al., 2014). The first version of GTO-ECV GDP covered the period from 1995 to 2008, but this has now been extended to June 2013. In addition to the retrieval algorithm, the level 3 gridding method and the merging algorithm differ from the approach used for GTO-ECV CCI. Regarding the level 3 generation, only one measurement per day and grid cell is used for the GTO-ECV GDP product and the daily grid cells have a size of $0.33^\circ \times 0.33^\circ$. Regarding the merging approach, all available satellites are averaged instead of using only one at a time. GTO-ECV GDP was incorporated already in the preceding WMO scientific assessment of ozone depletion (WMO, 2011). Moreover, it has been used for chemistry-climate model evaluation (Loyola et al., 2009a) as well as the investigation of decadal ozone trends and variability (Loyola et al., 2009b; Coldewey-Egbers et al., 2014). Both GTO-ECV CCI and GDP data records agree very well

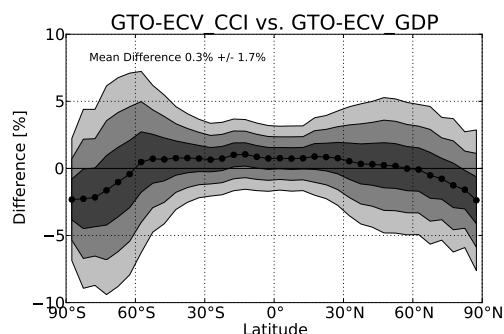


Figure 14. Percentage differences between GTO-ECV CCI and GDP $1^\circ \times 1^\circ$ monthly means 1996–2011 binned into 5° latitude belts (black dots). The grey shading denotes the 1-, 2-, and 3- σ standard deviations, respectively.

regarding the long-term trends, emphasizing their excellent decadal stability.

Figure 14 presents the percentage differences between GTO-ECV CCI and GDP $1^\circ \times 1^\circ$ monthly means binned into 5° latitude belts (black dots). The grey shading denotes the 1-, 2-, and 3- σ standard deviations, respectively. Both data records show a remarkable inter-consistency; the overall mean difference is $0.3\% \pm 1.7\%$. The deviations are slightly positive in low and middle latitudes, and negative in high latitudes. This latitudinal structure of the differences is mainly due to the usage of different level 2 retrieval algorithms. The application of different level 3 gridding methods leads to differences of up to $\pm 4\%$ in regions where two or more orbits per day overlap each other.

4.2 SBUV version 8.6 merged ozone data record

Within the framework of the NASA (National Aeronautics and Space Administration) program MEaSUREs (Making Earth System data records for Use in Research Environments) data from a series of nine BUV, SBUV, and SBUV/2 instruments have been reprocessed using the version 8.6 ozone retrieval algorithm (Labow et al., 2013; McPeters et al., 2013). From these data records a coherent long-term 5° zonal monthly mean ozone time series covering the periods 1970–1972 and 1979–2014 has been created which contains both profile and total ozone column information (Frith et al., 2014). Chiou et al. (2014) compared this merged data set (referred to as SBUV-MOD in the following) with GTO-ECV CCI and ground-based total ozone columns for the 16-year overlap period from March 1996 to June 2011. They found very good agreement in terms of monthly zonal mean total ozone and monthly zonal mean anomalies (their Figs. 6 and 8). The mean difference between both data sets is $0.3 \pm 1.1\%$.

Figure 15 shows the percentage difference between GTO-ECV CCI and SBUV-MOD 5° zonal mean ozone columns as a function of latitude. The black curve denotes the annual mean difference and its standard deviation (grey shaded

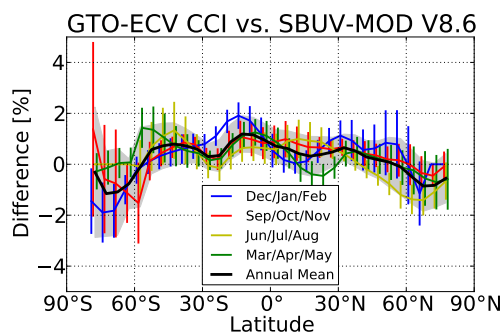


Figure 15. Percentage differences between GTO-ECV CCI and SBUV-MOD 5° monthly zonal mean ozone columns 1996–2011 as a function of latitude. Black: annual mean difference and its standard deviation (grey shaded area). Blue, red, yellow, and green lines denote the differences for northern hemispheric winter, autumn, summer, and spring, respectively.

area), and the blue, red, yellow, and green lines denote the seasonal differences. On average, the differences are positive in middle and low latitudes, and negative in high latitudes, where largest deviations occur in the summer months. Largest scatter is found in the Southern Hemisphere poleward of 50° S. The amplitude of the seasonal cycle in the differences is about 1%.

In addition to total ozone columns we compare the standard deviations of the 5° zonal monthly means. Figure 16 indicates that the latitudinal and temporal structures of the standard deviations agree very well. The absolute differences (shown in the bottom panel) are small in low and middle latitudes, and reveal larger spread in the months and latitudes close to the polar night terminator.

5 Summary and outlook

In this paper, which is the third in a series of three on the ESA Ozone_cci total ozone products, we have described the new GTO-ECV CCI level 3 global monthly mean data record spanning the 15-year time period 1996–2011. The data record is composed of total ozone measurements from three European nadir UV backscatter sensors GOME/ERS-2, SCIAMACHY/ENVISAT, and GOME-2/MetOp-A. It is publicly available at <http://www.esa-ozone-cci.org>. The companion papers by Lerot et al. (2014) and Koukouli et al. (2015) introduced the ozone retrieval algorithm GODFIT_V3 and presented the validation of the level 2 total ozone products, which form the basis for the GTO-ECV CCI merged level 3 product described herein.

The merging approach relies on an inter-sensor calibration procedure using GOME as the reference. Small corrections have been applied to SCIAMACHY and GOME-2A in order to reduce the differences among the instruments. Special emphasis was placed on the analysis of sampling issues intrinsic

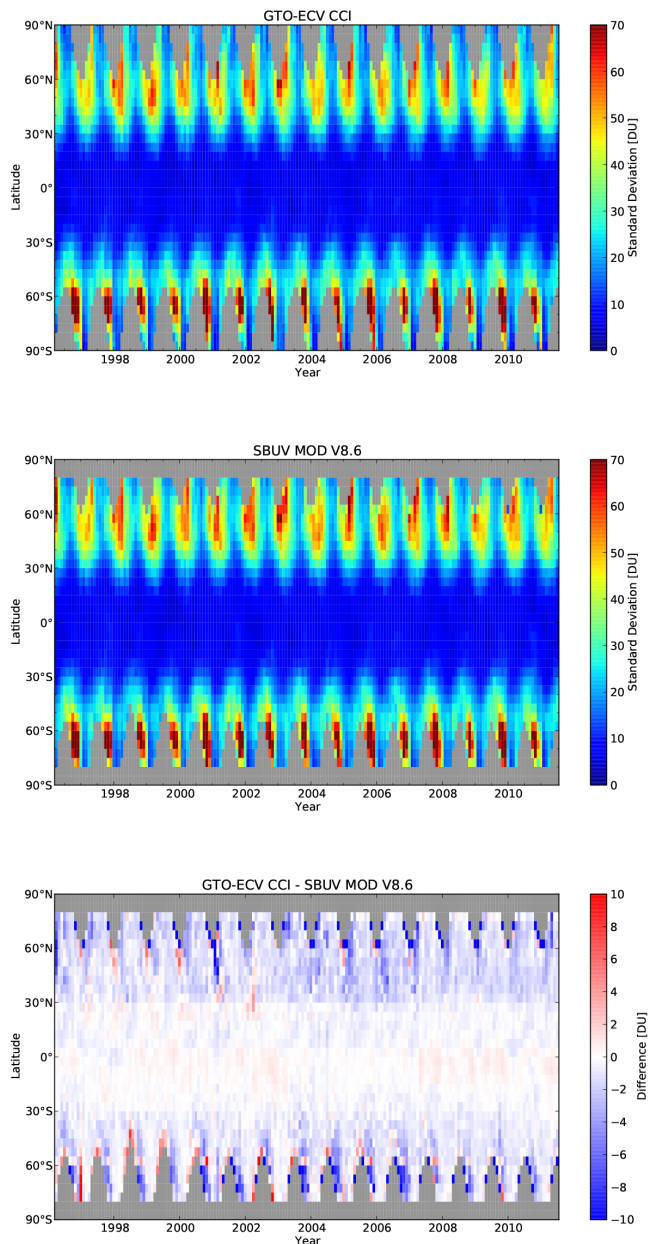


Figure 16. Standard deviation of 5° monthly mean ozone columns as a function of latitude and time: GTO-ECV CCI (top panel) and SBUV-MOD (middle panel). The bottom panel shows the absolute difference between GTO-ECV CCI and SBUV-MOD standard deviations.

to the satellite data and their impact on the final GTO-ECV CCI product.

We presented level 3 product geophysical validation results using as reference ground-based measurements with Brewer, Dobson, and UV-visible SAOZ instruments. The validation of the GTO-ECV CCI level 3 data record was found to be very consistent with the equivalent separate GOME, SCIAMACHY, and GOME-2A level 2 validation

(Koukouli et al., 2015). In particular, on a monthly mean basis, for the Dobson comparisons, both the Northern and Southern Hemisphere time series are in strong agreement. Similarly, for the Brewer comparisons (Northern Hemisphere), an excellent agreement is found apart from a handful of outliers. On a seasonal basis, both the Brewer and the Dobson level 3 comparisons show close agreement with the level 2 comparisons. We conclude that the current 15-year GTO-ECV CCI level 3 total ozone data product is of the same high quality as the equivalent individual level 2 data products that constitute it. This is due to a very high level of consistency among the level 2 products themselves and a robust merging approach. Both absolute agreement and long-term stability are excellent for almost all latitudes apart from a few outliers which are mostly due to sampling differences between the level 2 and level 3 data that cannot be completely eradicated.

This study demonstrates that the current GTO-ECV CCI data record is suitable for a variety of applications. In particular it is useful for the long-term monitoring of the past evolution of the ozone layer. Due to its excellent decadal stability – the relative drift compared to the ground-based reference is less than 1 % per decade – it is valuable for long-term trend analysis of the ozone field. The high spatial resolution of the level 3 data record of $1^\circ \times 1^\circ$ enables us to investigate ozone changes on global as well as regional scales as recently demonstrated by Coldewey-Egbers et al. (2014).

Furthermore, global long-term data records such as GTO-ECV CCI can be compared with chemistry–climate model simulations. One of the main purposes of these models is to identify and quantify relevant processes and forcings affecting the ozone layer and to project their future evolution. In particular, the simulations are analysed to assess the returning of ozone to historical levels and the complete recovery from ODSs as a consequence of the 1987 Montreal Protocol (UNEP, 1986). The satellite-based data records enable us to evaluate these model projections and to calibrate the efficacy of the model system (Loyola et al., 2009a).

Regarding total ozone, the second phase of ESA-CCI is dedicated to an improvement of the sampling errors (see Sects. 2.2 and 2.3) using spatio-temporal statistical tools and an extension of the GTO-ECV CCI data record. The GOME-2A sensor used in this study is the first of a series of three identical instruments. GOME-2 on MetOp-B was launched in September 2012 and the data will be included in the new version of GTO-ECV. In addition measurements performed with the Ozone Monitoring Instrument (OMI) onboard the NASA Aura satellite (2004–present) – which have been recently reprocessed with an adapted version of the GODFIT_V3 retrieval algorithm – and data from the Ozone Mapping and Profiler Suite (OMPS) onboard the NASA Suomi National Polar-orbiting Partnership satellite (2011–present) will be included. Thereby we can take advantage of OMI's excellent long-term stability over the 10 plus years of operation. The GOME-2 on MetOp-C is planned to be launched in 2018, and together with the Sentinel-5 Precursor (to be

launched in 2016) and the Sentinel-4 and Sentinel-5 sensors (to be launched by the end of this decade), these future instruments will contribute to the extension of this reference data set.

Acknowledgements. This work has been performed as part of ESA's Climate Change Initiative ozone project. The authors would like to thank ESA for the GOME and SCIAMACHY level 1 data and EUMETSAT for the GOME-2A level 1 data. Thanks to the EUMETSAT O3M-SAF project for providing the GOME-2A total ozone used in the GTO-ECV GDP data record. The ground-based data used for geophysical validation were obtained as part of WMO's Global Atmosphere Watch (GAW) and its contributing Network for the Detection of Atmospheric Composition Change and are publicly available (<http://woudc.org> and <http://ndacc.org>). We thank the various national and international agencies for funding the ground-based measurements and the databases, and the observers whose dedication made these data possible. Total ozone data utilized in the OSSE were provided by ECMWF.

Edited by: H. Worden

References

- Balis, D., Kroon, M., Koukouli, M. E., Brinkma, E. J., Labow, G., Veefkind, J. P., and McPeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, *J. Geophys. Res.*, 112, D24S46, doi:10.1029/2007JD008796, 2007a.
- Balis, D., Lambert, J.-C., van Roozendaal, M., Spurr, R., Loyola, D., Livschitz, Y., Valks, P., Amiridis, V., Gerard, P., Granville, J., and Zehner, C.: Ten years of GOME/ERS2 total ozone data. The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8, *J. Geophys. Res.*, 112, D07307, doi:10.1029/2005JD006376, 2007b.
- Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: mission objectives and measurement modes, *J. Atmos. Sci.*, 56, 127–150, 1999.
- Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V. V., Ladstädter-Weissenmayer, A., Richter, A., de Beek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., and Perner, D.: The Global Ozone Monitoring Experiment (GOME): mission concept and first scientific results, *J. Atmos. Sci.*, 56, 151–175, 1999.
- Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – Metop's second generation sensor for operational ozone monitoring, *ESA Bull.-Eur. Space*, 102, 28–36, 2000.
- Chiou, E. W., Bhartia, P. K., McPeters, R. D., Loyola, D. G., Coldewey-Egbers, M., Fioletov, V. E., van Roozendaal, M., Spurr, R., Lerot, C., and Frith, S. M.: Comparison of profile total ozone from SBUV (v8.6) with GOME-type and ground-based total ozone for a 16-year period (1996 to 2011), *Atmos. Meas. Tech.*, 7, 1681–1692, doi:10.5194/amt-7-1681-2014, 2014.
- Coldewey-Egbers, M., Loyola, D., Braesicke, P., Dameris, M., van Roozendaal, M., Lerot, C., and Zimmer, W.: A new health check of the ozone layer at global and regional scales, *Geophys. Res. Lett.*, 41, 4363–4372, doi:10.1002/2014GL060212, 2014.
- Frith, S. M., Kramarova, N. A., Stolarski, R. S., McPeters, R. D., Bhartia, P. K., and Labow, G. J.: Recent changes in total column ozone based on the SBUV Version 8.6 merged ozone data set, *J. Geophys. Res.*, 119, 9735–9751, doi:10.1002/2014JD021889, 2014.
- GCOS: Systematic Observation Requirements for Satellite-Based Data Products for Climate – 2011 Update, available at: <http://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf> (last access: 29 April 2015), 2011.
- Hao, N., Koukouli, M. E., Inness, A., Valks, P., Loyola, D. G., Zimmer, W., Balis, D. S., Zyrichidou, I., van Roozendaal, M., Lerot, C., and Spurr, R. J. D.: GOME-2 total ozone columns from MetOp-A/MetOp-B and assimilation in the MACC system, *Atmos. Meas. Tech.*, 7, 2937–2951, doi:10.5194/amt-7-2937-2014, 2014.
- Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendaal, M., and Wagner, W.: The ESA Climate Change Initiative: satellite data records for essential climate variables, *B. Am. Meteorol. Soc.*, 94, 1541–1552, doi:10.1175/BAMS-D-11-00254.1, 2013.
- Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrilat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, *Atmos. Chem. Phys.*, 13, 4073–4109, doi:10.5194/acp-13-4073-2013, 2013.
- Koukouli, M. E., Balis, D. S., Loyola, D., Valks, P., Zimmer, W., Hao, N., Lambert, J.-C., van Roozendaal, M., Lerot, C., and Spurr, R. J. D.: Geophysical validation and long-term consistency between GOME-2/MetOp-A total ozone column and measurements from the sensors GOME/ERS-2, SCIAMACHY/ENVISAT and OMI/Aura, *Atmos. Meas. Tech.*, 5, 2169–2181, doi:10.5194/amt-5-2169-2012, 2012.
- Koukouli, M. E., Lerot, C., Lambert, J.-C., Goutail, F., Granville, J., Pommereau, J.-P., Balis, D., Zyrichidou, I., van Roozendaal, M., Labow, G., Frith, S., Loyola, D., Coldewey-Egbers, M., Spurr, R., and Zehner, C.: Evaluating a new homogeneous total ozone climate data record from GOME/ERS-2, SCIAMACHY/Envisat and GOME-2/MetOp-A, *J. Geophys. Res.*, in review, 2015.
- Labow, G. J., McPeters, R. D., Bhartia, P. K., and Kramarova, N.: A comparison of 40 years of SBUV measurements of column ozone with data from the Dobson/Brewer network, *J. Geophys. Res.*, 118, 7370–7378, doi:10.1002/jgrd.50503, 2013.
- Lacan, A. and Lang, R.: Investigation on GOME-2 throughput degradation, Final report, EUM/LEO/REP/09/0732 Issue 1.1, 2011.
- Lerot, C., van Roozendaal, M., van Geffen, J., van Gent, J., Fayt, C., Spurr, R., Lichtenberg, G., and von Barga, A.: Six years of total

- ozone column measurements from SCIAMACHY nadir observations, *Atmos. Meas. Tech.*, 2, 87–98, doi:10.5194/amt-2-87-2009, 2009.
- Lerot, C., van Roozendaal, M., Spurr, R., Loyola, D., Coldewey-Egbers, M., Kochenova, S., van Gent, J., Koukouli, M., Balis, D., Lambert, J.-C., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat and GOME-2/MetOp-A, *J. Geophys. Res.*, 119, 1639–1662, doi:10.1002/2013JD020831, 2014.
- Loyola, D. and Coldewey-Egbers, M.: Multi-sensor data merging with stacked neural networks for the creation of satellite long-term climate data records, *EURASIP J. Adv. Sig. Pr.*, 2012, 91, doi:10.1186/1687-6180-2012-91, 2012.
- Loyola, D. G., Coldewey-Egbers, M., Dameris, M., Garny, H., Stenke, A., van Roozendaal, M., Lerot, C., Balis, D., and Koukouli, M.: Global long-term monitoring of the ozone layer – a prerequisite for predictions, *Int. J. Remote Sens.*, 30, 4295–4318, 2009a.
- Loyola, D. G., Coldewey-Egbers, M., Zimmer, W., Koukouli, M., Balis, D., Lerot, C., van Roozendaal, M., and Dameris, M.: Total ozone trends derived from the 14-years merged GOME/SCIAMACHY/GOME-2 data record, in: *Proc. ESA Atmospheric Science Conference, European Space Agency, Barcelona, Spain, 7–11 September 2009*, 1136, 2009b.
- Loyola, D. G., Koukouli, M. E., Valks, P., Balis, D. S., Hao, N., van Roozendaal, M., Spurr, R. J. D., Zimmer, W., Kiemle, S., Lerot, C., and Lambert, J.-C.: The GOME-2 total column ozone product: retrieval algorithm and ground-based validation, *J. Geophys. Res.*, 116, D07302, doi:10.1029/2010JD014675, 2011.
- McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G. L., and Flynn, L.: The version 8.6 SBUV ozone data record: an overview, *J. Geophys. Res.*, 118, 8032–8039, doi:10.1002/jgrd.50597, 2013.
- Pommereau, J. P. and Goutail, F.: O₃ and NO₂ ground-based measurements by visible spectrometry during Arctic winter and spring 1988, *Geophys. Res. Lett.*, 15, 891–894, doi:10.1029/GL015i008p00891, 1988.
- Sakazaki, T., Fujiwara, M., Mitsuda, C., Imai, K., Manago, N., Naito, Y., Nakamura, T., Akiyoshi, H., Kinnison, D., Sano, T., Suzuki, M., and Shiotani, M.: Diurnal ozone variations in the stratosphere revealed in observations from the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES) on board the International Space Station (ISS), *J. Geophys. Res.*, 118, 2991–3006, doi:10.1002/jgrd.50220, 2013.
- Scarnato, B., Staehelin, J., Peter, T., Gröbner, J., and Stübi, R.: Temperature and slant path effects in Dobson and Brewer total ozone measurements, *J. Geophys. Res.*, 114, D24303, doi:10.1029/2009JD012349, 2009.
- Sofieva, V. F., Rahpoe, N., Tamminen, J., Kyrölä, E., Kalakoski, N., Weber, M., Rozanov, A., von Savigny, C., Laeng, A., von Clarmann, T., Stiller, G., Lossow, S., Degenstein, D., Bourassa, A., Adams, C., Roth, C., Lloyd, N., Bernath, P., Hargreaves, R. J., Urban, J., Murtagh, D., Hauchecorne, A., Dalaudier, F., van Roozendaal, M., Kalb, N., and Zehner, C.: Harmonized dataset of ozone profiles from satellite limb and occultation measurements, *Earth Syst. Sci. Data*, 5, 349–363, doi:10.5194/essd-5-349-2013, 2013.
- Sofieva, V. F., Kalakoski, N., Päiväranta, S.-M., Tamminen, J., Laine, M., and Froidevaux, L.: On sampling uncertainty of satellite ozone profile measurements, *Atmos. Meas. Tech.*, 7, 1891–1900, doi:10.5194/amt-7-1891-2014, 2014.
- UNEP: Handbook for the Montreal Protocol on Substances that Deplete the Ozone Layer, Ozone Secretariat – United Nations Environment Programme, available at: http://ozone.unep.org/Publications/MP_Handbook/index.shtml (last access: 29 April 2015), 1986.
- van der A, R.: User Requirement Document, Issue 2.1, available at: www.esa-ozone-cci.org/?q=webfm_send/37 (last access: 29 April 2015), 2011.
- van Roozendaal, M., Peeters, P., Roscoe, H. K., De Backer, H., Jones, A., Vaughan, G., Goutail, F., Pommereau, J.-P., Kyrö, E., Wahlstrøm, C., Braathen, G., and Simon, P. C.: Validation of ground-based UV-visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, *J. Atmos. Chem.*, 29, 55–83, 1998.
- van Roozendaal, M., Loyola, D., Spurr, R., Balis, D., Lambert, J.-C., Livschitz, Y., Valks, P., Ruppert, T., Kenter, P., Fayt, C., and Zehner, C.: Ten years of GOME/ERS-2 total ozone data. The new GOME data processor (GDP) version 4: 1. Algorithm description, *J. Geophys. Res.*, 111, D14311, doi:10.1029/2005JD006375, 2006.
- van Roozendaal, M., Spurr, R., Loyola, D., Lerot, C., Balis, D., Lambert, J.-C., Zimmer, W., van Gent, J., van Geffen, J., Koukouli, M., Granville, J., Doicu, A., Fayt, C., and Zehner, C.: Sixteen years of GOME/ERS-2 total ozone data: the new direct-fitting GOME Data Processor (GDP) version 5 – algorithm description, *J. Geophys. Res.*, 117, D03305, doi:10.1029/2011JD016471, 2012.
- Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J.-P., Redondas, A., Van Roozendaal, M., and Lambert, J.-C.: Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons, *Atmos. Meas. Tech. Discuss.*, 8, 8023–8082, doi:10.5194/amt-d-8-8023-2015, 2015.
- WMO: Scientific Assessment of Ozone Depletion: 2010, WMO Global Ozone Research and Monitoring Project – Report No. 52, Geneva, Switzerland, 2011.
- WMO: Scientific Assessment of Ozone Depletion: 2014, WMO Global Ozone Research and Monitoring Project – Report No. 55, Geneva, Switzerland, 2014.