Interactive comment on "Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 1. Ground-based validation of total ozone column data products" by Katerina Garane et al.

Anonymous Referee #1

Received and published: 21 November 2017

General Comments:

Data records of satellite borne instruments are only temporary in contrast to most of the ground based total ozone column (TOC) records. Thus the development of a method to compare the available satellite records and to merge them to create a long term, homogeneous TOC data set, is a very valuable contribution to the monitoring of the ozone layer. This publication gives a very good description of the validation of such merged data records with ground based records of Dobson, Brewer and SAOZ instruments.

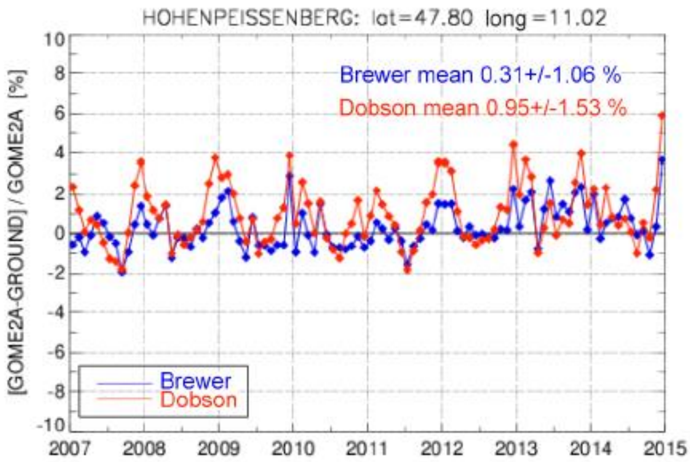
Specific Comments:

1. **Comment:**
   It should be mentioned that the used Dobson and Brewer TOC data records are still based on the "old" Bass and Paur ozone cross sections, whereas it seems that the satellite data are produced using the new ozone cross sections (Bremen, IUP?), good place for this explanation would be page 7 after line 25.
   REPLY:
   The explanation is added in section 2.3, as suggested. Thank you.

2. **Comment:**
   Dependence on effective temperature of the Dobsons (p 5- 6): Basher 1982 is not an appropriate reference, as it was written, when the ozone cross-sections after Vigroux had been valid. Current data sets are processed using Bass and Paur. Better and up to date references for this issue are: Koukouli et al., 2016 (cited later in the text, page 7) Scarnato et al., 2009: Temperature and slant path effects in Dobson and Brewer total ozone measurements, Journal of Geophysical Research: Atmospheres, Vol. 114, Issue D24 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, 11,129– 11,140, 1988. Kerr, 2002, New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct sun spectra, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 107, NO. D23
   REPLY:
   The references are added. Thank you for the suggestions.

3. **Comment:**
   The use of SAOZ might be seen a little bit problematically with its accuracy of 6% (page 6)
   REPLY:
   While this overall accuracy is poor in comparison to that of the direct-sun instruments, the added value provided by the SAOZ instruments is their ability to produce reference measurements at those locations and times-of-year where and when the satellite measurements occur under low-sun conditions and no reliable direct-sun measurements can be made. As such, they allow the validation of an otherwise inaccessible satellite measurement regime. This point was already made in the paper.
   Moreover, it must also be noted that a significant fraction of this 6% total accuracy is made up of the (systematic) uncertainty in the O_3 cross sections (3%) and by the impact of clouds (3.3 %, Hendrick et al., 2011), both of which are of minor importance in differential analyses of cloud-free data. This note was added in the paper, in Section 2.2.

4. **Comment:**
   On page 12 a correction for the Izana record due to the altitude is mentioned. Such a correction should make sense for other mountain stations too, especially when they are more or less isolated compared with the 150km footprint of the satellite data. A first guess of correction would be +0.1% per 100m difference of station altitude and environmental altitude. There are some mountain stations with significant differences (e.g. Arosa, Hohenpeissenberg, Mauna Loa). This information can be included in the tables S1 – S3.
   REPLY:
   The mentioned correction for the SAOZ measurements is an ERA-Interim-based estimate of the column below the instrument altitude in the immediate vicinity of the island and/or mountain top (see Verhoelst et al, 2015 for further details). For the SAOZ/ZSL-DOAS network, Izana and Jungfraujoch are the only stations for which a significant missing column was derived with this methodology (about 2.8% and 3.2% respectively, with some seasonal variation), due to their isolated mountain-top locations.
   As for the ground based measurements performed by Dobson and Brewer spectrophotometers that are used in this work, since they are downloaded from the WOUDC database we are not able to correct them for the altitude issue, as suggested. Nevertheless, we can use the information to identify any discrepancies seen in our figures.
   Furthermore, as seen in Koukouli et al. (2016), when a high altitude station like Hohenpeissenberg (where the gradient is not very steep and the instrument is exceptionally maintained and calibrated) is used, the satellite-to-ground comparison is excellent (Brewer bias ~0.3% and Dobson bias ~1%, see figure below). For the Mauna Loa station (10° - 20° N), on the other hand, where the gradient is much steeper, the satellite-to-ground comparison is about 2-4%. However, when considering zonal means of the differences, where all available stations in each belt are included in the calculations, the effect of the station altitude becomes less evident, which is the case for the 10-20° N belt in Figure 5 – panel (a) where Mauna Loa and Bangkok are co-calculated.
   Thank you for the suggestion about this issue, we will take it seriously under consideration and use it as basis for a future study.
   Some more information on the SAOZ measurements' correction is added in the manuscript, in section 2.3. We have also added the altitude information for each station in the Tables S1 – S3.



Koukouli et al. (2016) – Figure 1.

5. **Comment:**
   Addition information in these tables about the lengths of the records would be informative, as not all stations have measured from 1995 to 2017.
   REPLY:
   We have added the time period for each ground based station in the Tables S1 – S3.

6. **Comment:**
   The explanation on page 9, why the SZA-dependence for the Dobsons are not drawn is misleading. As reason a high correlation between Dobsons' large stratospheric effective temperature dependence and the SZA is mentioned. This correlation is physically not correct. The SZA of daily means of TOC is larger during winter season, when the sun is not very high. In addition in winter the Teff is lower than the used -46 degree Celsius. Thus it is an indirect correlation, which is e.g. not valid during summer season, when Teff is "normal" and Dobson TOCs drop at very high SZA (mue> than 3.5 depending on turbidity) -values because of straylight effects but not because of temperature dependence. In any case it is justified not to use Dobson data at SZA larger than 75 degrees, even if they were available.
   REPLY:
   Thank you for the suggestion. We agree with the comment and we have added two plots (SH and NH) in Figure 4 (and the respective comments) showing the dependence of the satellite-to-Dobson comparison on SZA. As for the cut-off at 75°, we did not apply it because the SZAs used for the binning and the plots are the satellite SZAs, since we use daily means of the ground based measurements. We have also added a sentence in section 2.3 making this clear.

7. **Comment:**
   In figures 4, 5 and 10 Brewer observations are drawn above SZA of 75 degrees. The slant path mue of these measurements are larger than 3.5. Observations with larger mue-values

are not accurate enough, especially when using single Brewers. Double Brewers might be able to measure up mue = 4, before the TOC drops (reason see Dobson explanation of straylight effects in the bullet point before).

REPLY:

Thank you for the suggestion. As explained in the previous comment, please note that the SZAs used for the binning and the plots are the satellite SZAs, since we use daily means of the ground based measurements.

8. **Comment:**

   Concerning the seasonality of SAOZ-difference mentioned on page 9 and seen in figure 3: Is there an explanation for this pattern?

   REPLY:

   It should be noted that seasonality seen on Figure 3 are observed at all latitudes and all instruments but,

   i. the amplitude is larger in NH compared to SH on both SAOZ and Dobson
   ii. the amplitude is larger with SAOZ compared to Dobson and Brewer
   iii. the amplitude varies with the satellites, the largest being with GOME and SCIAMACHY in Northern hemisphere.
   iv. the strongest minima are observed in the winter particularly on the difference between SAOZ and GOME.

   The seasonality can be attributed to:

   i) the cross sections dependencies on the effective temperature of the stratosphere impacting all measurements in the UV but not SAOZ analyzing in the visible.
   ii) the number of stations used in the statistics in winter, limited in latitudes for Dobson and Brewer but being possible at higher latitude for SAOZ.

   The SAOZ seasonality observed on panels (d) and (e) of Figure 3 comes from the latitude of the selected stations, which, in the case of SAOZ, allows to perform comparisons in winter at high latitude when the effect of the temperature on UV cross section is the largest.

   We have modified the respective paragraph commenting on panels (d) and (e) of Figure 3, in Section 2.3, to give a more clear explanation of the seasonality effect.

Technical corrections:

1. In references Serdyuchenko on page 26 “‐ Part 2” is written instead of “- Part 2“.
2. Kerr et al. 1988 is cited on page 5, line 18, but cannot be found in the references.

REPLY:

The references are corrected/added. Thank you!

Interactive comment on "Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 1. Ground-based validation of total ozone column data products" by Katerina Garane et al.

Anonymous Referee #2

Received and published: 18 December 2017

The manuscript presents the validation study for the merged satellite total ozone record obtained from the GOME-type satellite ozone sensors. In this work the new dataset is validated against the ground-based network of Dobson, Brewer an SAOZ instruments. Presented study fits well to the scope of the problems covered in the AMT. The manuscript is well written and organized.

Major comments:

1. **Comment:**
   Most of the results presented in the study are shown for the hemispheric monthly means. However, the description of the methodology to compute these hemispheric means is not provided. These means could be computed in a number of ways (e.g. with/without weights), thus it would be important to provide a brief description in the text or in the appendix/supplement. Scientific results should be reproducible, and the clear description of the methods is an important component to ensure the robustness of the results.
   REPLY:
   Thank you for this remark. We agree and we have added a few sentences explaining the methodology of our calculations in Section 2.3, where Figure 3 is commented.

2. **Comment:**
   In the Tables in the Supplement, please, indicate the time periods for which the data from individual ground-based stations were used in this study.
   REPLY:
   We have added the time period for each ground based station in the Tables S1 – S3.

3. **Comment:**
   Page 9, lines 15-18: I don't quite understand the reason for not showing a plot with the SZA dependence for Dobson comparisons. Authors stated that Dobson observations depend on the stratospheric temperature, which should produce an artificial dependence on SZA. At the same time, when authors discuss results for comparisons with Brewer and SAOZ, they claim that the observed dependence on SZA for satellite measurements doesn't matter because all satellites show consistent patterns relative to ground-based stations. If so, why don't to show results with Dobson if the goal is to check consistency among satellite records.

REPLY:

Thank you for the suggestion. We agree with the comment and we have added two plots (SH and NH) in Figure 4, showing the dependence of the satellite to Dobson comparison on SZA, and the respective comments in section 2.3.

4. **Comment:**

   Page 12, line 15: authors mentioned that the results of the comparison with measurements at Izana station were adjusted to account for the station's elevation. Have these adjustments been applied to any other station? This should be clearly described in the text.

   REPLY:

   The mentioned correction is an ERA-Interim-based estimate of the column below the instrument altitude in the immediate vicinity of the island and/or mountain top (see Verhoelst et al, 2015, for further details). For the SAOZ/ZSL-DOAS network, Izana and Jungfraujoch are the only stations for which a significant missing column was derived with this methodology (about 2.8% and 3.2% respectively, with some seasonal variation), due to their isolated mountain-top locations. Some more information is added in the manuscript, section 2.3.

5. **Comment:**

   Page 12-13, Table 1: I found that the quantities shown in Table 1 are not well described. I think this part needs some major revision, including terms that are used in the table. For example, I would recommend using a term "mean bias" instead of "monthly mean bias" and mentioning in the text that the mean bias was computed from monthly mean differences, because to me "monthly mean bias" would mean the bias in a specific month, while, if I understood this correctly, the biases shown in Table 1 were computed over the entire data record. I don't quite understand what is "monthly mean variability" in Table 1. The 1–sigma standard deviations for biases are shown along with the biases. In the text this quantity is explained as "the variability of the monthly mean standard deviation values". Is that the variability of the standard deviations of differences in individual months? I would also recommend replacing "Seasonality" with "seasonal bias". The two last lines in the table are very confusing (Latitude and SZA). At first, I thought they show the mean differences in latitude and sza between satellite and ground based observations. But according to the text these are the mean differences in ozone between satellite and ground-based observations, calculated by averaging all points shown in Figure 4 and Figure 5. I understand that if you try to bin differences in smaller bins (like you did with SZA) you can uncover some dependence on SZA. Then it would make sense to name these quantities as "Latitudinal biases" and "SZA biases". But why do you expect results to be different from the mean biases if you average over the entire latitude range? Please, explain.

   REPLY:

   Thank you for the comments and the suggestions. Please find below our replies:

   - The "Monthly mean bias" term was corrected in the text as suggested and a phrase explaining how it was calculated is added.
   - The "monthly mean variability" term is indeed the variability of the monthly mean standard deviation values and the text is modified so as to be clearer.

- "Seasonality": We consider the "bias" as a term to be the deviation of comparisons from the 0% line, which would mean 0% bias. In this statistic quantity we calculate the peak-to peak range in values, so we think changing the term would imply a different statistical quantity.
- Latitude and SZA statistics: we agree that we do not expect to see any major differences to the mean differences of the monthly means, but these statistical quantities have to be provided as requirements by the Users of the ESA CCI project (they will be uploaded soon at: http://www.esa-ozone-cci.org/?q=documents#). We have added the reference and changed the text accordingly. We also named the quantities "Latitudinal mean bias" and "SZA mean bias".

6. **Comment:**
   Page 15 lines 25-31 and Page 16 lines 1-2: In this part of the manuscript authors describe the correction factors against OMI that were used to correct individual satellite time series. It is stated that "we apply correction factors using the seasonal mean differences…". Then it mentioned that the drift in GOME 2A have been accounted for. Is it a static correction that depends on lat/lon and month of the year only? Or have you implemented time-dependent corrections? Please, explain that in the text.
   REPLY:
   For GOME and SCIAMACHY it is a static correction that depends on latitude and month of the year. For GOME-2A and GOME-2B it is a fully time- (and latitude-) dependent correction. We have improved the corresponding explanation in the text, Section 3.1.

7. **Comment:**
   Page 17, lines 10-20: you need to explain what was done in the merging process when data from two or more instruments are available. Did you simply average all available data? Did you use some weights?
   REPLY:
   The merging is done on a daily basis. When data from two or more instruments are available, we average all and use as weights the number of measurements per day and grid box for the corresponding sensor. We added the explanation in Section 3.1.

8. **Comment:**
   Looking at the results showing in Figure 11, it seems to me that the merged dataset almost fully overlaps with OMI. This is expected since all individual datasets have been corrected against OMI, and OMI has a very dense spatial and temporal coverage. My question here: what is the value of using GOME 2A or GOME 2B in the merged product? Please, provide an explanation in the text.
   REPLY:
   We agree with the reviewer that it is not unexpected that the results of the level-3 validation are dominated by the very dense OMI measurements. As already seen in the ground-based validation of the previous version of the merged product (Coldewey-Egbers et al., 2015) the quality of monthly mean level-3 data strongly depends on the spatio-temporal coverage of the input (level-2) measurements. Therefore, we think that it is more beneficial to include all

available measurements in order to further improve coverage and, thus, statistics and representativeness of the monthly mean values. We added the explanation in Section 3.1.

9. **Comment:**
   Figure 7: I would suggest to keep the range for the time scale (X-axis) the same for all 4 panels;
   REPLY:
   We changed the range for the time scale as suggested and we added a gray background to better distinguish between missing data and values close to zero.

10. **Comment:**
    Figure 9: it's very hard to see blue letters on the green background. I would suggest to move satellite timelines either to the top or bottom of the figure.
    REPLY:
    We changed the color of the letters to 'black' and moved the timelines to the middle of the plot in order to improve visibility.

11. **Comment:**
    Figure 10, left panel: why there is no point for Level 3 product in 80-90N latitude bin even though data for all individual instruments are shown;
    REPLY:
    Thank you for noticing this issue. The reason is that there is one ground-based station nearly at the limit between the 70°- 80° bin and the 80°- 90° bin, namely Eureka, at 79°.89 North. As a result, when allowing spatial collocations to the satellite central pixel within a certain radius, some of the collocations were allocated to the 80°- 90° bin.
    We have updated our validation chain to take this into account and changed Figure 10 accordingly.
    When analyzing the ground-based level-2 into gridded level-3 this issue did not come up, so there was from the start no data in the 80° -90° bin.

12. **Comment:**
    I am puzzled why the Level 3 value in 70-80S latitude bin is higher than for any given individual instrument.
    REPLY:
    The time series shown in this Figure are not common collocations, as you can well imagine, since the different Level-2 instruments have a different time span, which again differs from the level-3 data. Hence, the variability shown by all six time series in the -70° to -80° S bin may be due to a number of factors, including ground-based data availability, as well as dynamic issues that may affect different years in a different manner, such as the polar vortex for example.

Minor comments:

1. All abbreviations should be spelled out when used for the first time in the text. For instance, there are many abbreviations in the Abstract and Introduction that are not explained: P. 1, line 15: "GOME-type" – please, spell out "GOME"; P. 1, line 18: "GODFIT", "ERS", "OMI",

"SCIAMACHY" –please, spell them out; P. 2, line 33: please, spell out "SAOZ"; P. 3, line 7: please, spell out "BIRA-IASB" and "DLR"; P. 3, line 10: please, spell out "LIDORT"

REPLY:

We have spelled all abbreviations in the main text (mostly in the introduction), as suggested. The satellite names in the Abstract are left as they were because we think that it would make it too extensive, but they are also spelled in the introduction.

2. P.3 lines 25-28: It is not quite clear from the context which quantity has "been estimated to rise up to +/-2%": systematic uncertainty in the ozone cross sections or ozone itself? Please, consider re-wording this statement.

   REPLY:

   We agree that it was not quite clear in the text that the quantity that is biased by +/-2 % due to the use of different cross sections, is ozone. The manuscript was rephrased. Thank you for the suggestion.

3. P. 14, line 23: I guess it should be "NOAA 18" (not NOAA 16) to match with the labels on the right panel of Figure 6.

   REPLY:

   Yes, thank you, it should be NOAA 18. We have made the correction.

4. There are several places in the manuscript where authors use words "excellent", "exceptional" etc. I would recommend to avoid these statements in the scientific publication and rather provide quantitative results like "the stability within +/-1%" or "biases less than 2%".

   REPLY:

   Thank you, we have scanned through the text and made all the necessary alterations.

Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 1. Ground-based validation of total ozone column data products

Katerina Garane[1], Christophe Lerot[2], Melanie Coldewey-Egbers[3], Tijl Verhoelst[2], Maria Elissavet Koukouli[1], Irene Zyrichidou[1], Dimitris S. Balis[1], Thomas Danckaert[2], Florence Goutail[4], Jose Granville[2], Daan Hubert[2], Arno Keppens[2], Jean-Christopher Lambert[2], Diego Loyola[3], Jean-Pierre Pommereau[4], Michel Van Roozendael[2] and Claus Zehner[5]

[1] Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece.
[2] Royal Belgian Institute for Space Aeronomy (BIRA-IASB), 3, Avenue Circulaire, B-1180 Brussels, Belgium.
[3] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik der Fernerkundung (IMF), 82234 Oberpfaffenhofen, Germany.
[4] LATMOS, CNRS, University Versailles St Quentin, Guyancourt, France
[5] European Space Agency, ESRIN, Frascati, Italy

Correspondence to: Katerina Garane (agarane@auth.gr)

Abstract. The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) is a Level-3 data record, which combines individual sensor products into one single cohesive record covering the 22 year period from 1995 to 2017, generated in the frame of the European Space Agency's Climate Change Initiative Phase-II. It is based on Level-2 total ozone data produced by the GODFIT (GOME-type Direct FITting) v4 algorithm as applied to the GOME/ERS-2, OMI/Aura, SCIAMACHY/Envisat and GOME-2/MetopA and /MetopB observations. In this paper we examine whether GTO-ECV meets the specific requirements set by the international climate-chemistry modelling community for decadal stability, long-term and short term accuracy. In the following, we present the validation of the 2017 release of the Climate Research Data Package Total Ozone Column (CRDP TOC), both at Level-2 and Level-3. The inter-sensor consistency of the individual Level-2 data sets have ~~show excellent inter-sensor consistency with~~ mean differences generally within 0.5 % at moderate latitudes (+/- 50°), whereas the Level-3 data sets show mean differences with respect to the OMI reference data record that span between -0.2 ± 0.9 % (for GOME-2B) and 1.0 ± 1.4 % (for SCIAMACHY). Very similar findings are reported for the Level-2 validation against independent ground-based TOC observations reported by Brewer, Dobson and SAOZ instruments; the mean bias between GODFIT v4 satellite TOC and ground instrument is well within 1.0 ± 1.0 % for all sensors, the drift per decade spans between -0.5 % to 1.0 ± 1.0 % depending on the sensor, and the peak-to-peak seasonality of the differences ranges between ~1% for GOME and OMI, to ~2% for SCIAMACHY. For the Level-3 validation, as a first step the aim was to show that the Level-3 CRDP produces consistent findings as the Level-2 individual sensor comparisons. We show a~~n excellent~~very good agreement with 0.5 to 2 % peak-to-peak amplitude for the monthly mean difference time series and a negligible drift per decade in the Northern Hemisphere differences at -0.11 ± 0.10 % per decade for Dobson and +0.22 ± 0.08 % per decade for Brewer collocations. The exceptional quality of the Level-3 GTO-ECV v3 TOC record temporal stability well satisfies the

requirements for the total ozone measurement decadal stability of between 1 – 3 % and the short term and long-term accuracy requirements of 2% and 3% respectively, showing an excellent remarkable inter-sensor consistency, both in the Level-2 GODFIT v4 as well as in the Level-3 GTO-ECV v3 datasets, and thus can be used for longer term analysis of the ozone layer, such as decadal trend studies, chemistry-climate model evaluation and data assimilation applications.

1 Introduction

The European Space Agency's *Climate Change Initiative* (CCI) Phases-I & -II focused on building consolidated climate-relevant Ozone data sets as Essential Climate Variables, ECVs. During Phase-I, the Ozone CCI mostly concentrated on developing and demonstrating improved algorithms and methods, with the aim to define new baselines for the generation of consistent, state-of-the-art and fully characterized long-term ozone data products derived from a complete suite of European nadir and limb-type sensors. For the first time, Earth Observation science teams consisting of leading experts from European ozone sensing communities were gathered in a single project working towards common objectives defined against requirements formulated by the scientific user community. This resulted in new synergies, exchanges of ideas, and overall significant progress in terms of data harmonisation and understanding of quality issues at Level-1, Level-2 and Level-3. Three lines of multi-sensor ozone data products were hence developed: (i). total ozone columns from Ultraviolet (UV) nadir instruments, (ii). low resolution ozone profiles from nadir sensors and (iii). stratospheric and upper tropospheric ozone profiles from limb and occultation types of sensors. During Phase-II, existing state-of-the-art ozone retrieval algorithms were further developed and applied to long time series of observations from all relevant ESA atmospheric chemistry sensors, with the aim to generate well characterized and validated ozone data products that meet as closely as possible the requirements formulated by the Global Climate Observing System (GCOS) as well as the Climate Modelling User Group (CMUG) climate modelling community, for ozone column and profile ECVs. The most important user requirements were identified as: (i). homogenized multi-decadal records, (ii). records with good vertical resolution in the (lower) stratosphere and (iii). records with good horizontal resolution in the troposphere, the main gap being the lack of multi-decadal high-vertical resolution ozone profile data sets that cover the full ozone depletion time period (1980-present) and provide a potential to cover the upcoming ozone recovery time period.

This work addresses the first of these requirements, the Level-2 and Level-3 homogenized multi-decadal total ozone Climate Research Data Package (CRDP), with two more companion papers (Keppens et al., 20187; Hubert et al., 2017) expanding on the limb and nadir ozone profile CRDPs. On total ozone, 21 years of harmonised Level-2 data records from GOME/ERS-2 (Global Ozone Monitoring Experiment instrument on board the second European Remote Sensing satellite), OMI/Aura (Ozone Monitoring Instrument on board Aura satellite), SCIAMACHY/Envisat (Scanning Imaging Absorption Spectrometer for Atmospheric Cartography on board Envisat) and GOME-2/MetopA -and /MetopB (Global Ozone Monitoring Experiment-2 on board MetopA and MetopB satellites) sensors have been produced using an advanced version of the direct-fitting GODFIT (GOME-type Direct FITting) v4 algorithm. The ESA-CCI total ozone CRDP includes the Level-2 products for each instrument (over the entire instrument lifetime) and a Level-3 merged monthly mean gridded data set using GOME and OMI as long-term stability reference.

In the following section, we briefly present the GODFIT v4 algorithm that creates the Level-2 CRDPs, followed by the validation against the Brewer, Dobson and SAOZ (Système d'Analyse par Observation Zénitale; Pommereau & Goutail, 1988) ground-based instruments and the comparison to the independent Solar Backscatter Ultraviolet measurements (SBUV) v8.6 long term TOC record. Thereafter, the algorithm that merges the individual Level-2 TOC records to create the Level-3 dataset is presented, followed by the validation to the ground-based records and inter-comparison to the individual Level-2 validation findings. Summary and conclusions are given in the last section.

2 Level-2 Total Ozone Columns

2.1 Satellite Total Ozone Column records

GODFIT (GOME-type Direct FITting) is an algorithm jointly developed by BIRA-IASB (Royal Belgian Institute for Space Aeronomy), RT Solutions and DLR (German Aerospace Center) to retrieve Total Ozone Columns (TOC) from satellite-borne nadir-viewing hyperspectral spectrometers, such as GOME(-2), SCIAMACHY and OMI. It relies on a non-linear least-squares minimization procedure, during which sun-normalized radiances simulated in the Huggins bands (325-335 nm) with the Radiative Transfer model LIDORT (Linearized Discrete Ordinate Radiative Transfer, (Spurr et al., 2013) are adjusted to the Level-1 measurements. As part of the phase-I of the ESA Ozone_cci project, version 3 of GODFIT has been successfully transferred to other nadir sensors and is comprehensively described in Lerot et al. (2014) and validated in Koukouli et al. (2015). During the second phase of this project, a number of algorithmic improvements have been realized and the full time series of GOME, OMI, SCIAMACHY and GOME-2A/B have been entirely reprocessed with the latest version (v4) of GODFIT. The most important update is the adaptation of the L1 soft-calibration scheme in order to restore the full independency of the satellite observations with respect to the ground-based measurements. This algorithm, described in detail in Danckaert et al. (2017), is also the future baseline for generating the offline operational total ozone from the TROPOMI/S5-p (TROPOspheric Monitoring Instrument on board the Copernicus Sentinel-5 Precursor satellite) instrument that launched in October 2017.

The radiance simulations require that the atmosphere is properly defined at each iteration within the retrieval and so a series of auxiliary data are also required. Ozone vertical profiles are prescribed by the total ozone classified climatology recently released by Labow et al. (2015) using MLS (Microwave Limb Sounder) and sondes data, combined with the tropospheric column database constructed by Ziemke et al. (2011). The ozone absorption is modelled using the temperature-dependent cross-sections measured by Serdyuchenko et al. (2014). The temperature in each atmospheric layer is prescribed by a priori profiles, allowed to be shifted by a constant offset, determined simultaneously to the total column. All cross-sections are preconvolved at the respective instrumental resolution and an improved correction for the so called solar I_0-effect (Aliwell et al., 2002) has been applied (Danckaert et al., 2017). GODFIT has the capability to characterize instrumental slit function on an orbit-basis by fitting pre-determined functions such as (Super-)Gaussian shapes (Beirle et al., 2017) or by stretching slit functions pre-measured on-ground. To account for contamination by clouds and/or aerosols, an effective scene approach is used (Coldewey-Egbers et al., 2005) in which the effective albedo of a scene located in between the cloud top height and the ground surface is fitted during the retrieval. The altitude of this effective scene depends on both the effective cloud fraction and cloud top altitude provided by independent cloud algorithms (FRESCO v7, Wang et al., 2008 or the O2-O2 product, Veefkind et al., 2016). Radiances are simulated on-the-fly with the scalar radiative transfer model LIDORT for GOME, SCIAMACHY and GOME-2. Because of the heavy computational burden of those simulations, the radiances may alternatively be extracted from a pre-computed look-up table, of which the granularity has been cautiously defined in order to limit interpolation errors while keeping a reasonable size (Danckaert et al., 2017). Once simulated, correction terms are applied to the radiances to correct for the impact of atmospheric polarization and inelastic scattering processes (Lerot et al., 2014).

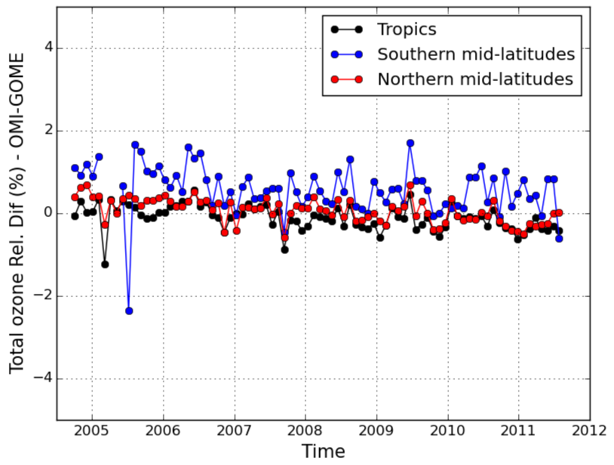


Figure 1. Time series of the relative differences between the total ozone columns retrieved from the GOME and OMI sensors for different latitude bands. Retrievals have been performed without any soft-calibration of the reflectances for both instruments.

When a common retrieval algorithm is applied to various instruments, systematic differences may remain due to calibration deficiencies or instrumental degradation effects affecting the Level-1 reflectance data. To generate the CCI total ozone data sets with the high inter-sensor consistency required for climate studies, an original soft-calibration scheme had been incorporated within GODFIT v3. This procedure, extensively described in Lerot et al. (2014), relied on reference total column measurements at selected Northern mid-latitude Brewer stations. Although it was shown to work well, this approach had the disadvantage to introduce a link between the satellite and ground-based measurements. As illustrated in Figure 1~~Figure 1~~, experience has shown that the GOME and OMI sensors perform in an extremely stable way and do not require any spectral soft-calibration procedure. Therefore it was decided to use these two instruments to soft-calibrate the spectra measured by

SCIAMACHY and GOME-2A/B. In practice, for every cloud-free satellite pixel falling into a reference sector between 40° S-50° N and 175° W-145° W, the closest reference clear-sky OMI (or GOME before 2005) column is used to simulate a radiance (using the GODFIT forward model), which is then compared to the Level-1 spectrum recorded by the sensor to be soft-calibrated. Such comparisons are done systematically for a large number of pixels (e.g. several hundreds of thousands for GOME-2A) spanning most of the observation geometries and the full time series, which allows to identify and correct for systematic issues in the Level-1 data. See Lerot et al. (2014) for more details on the soft-calibration approach.

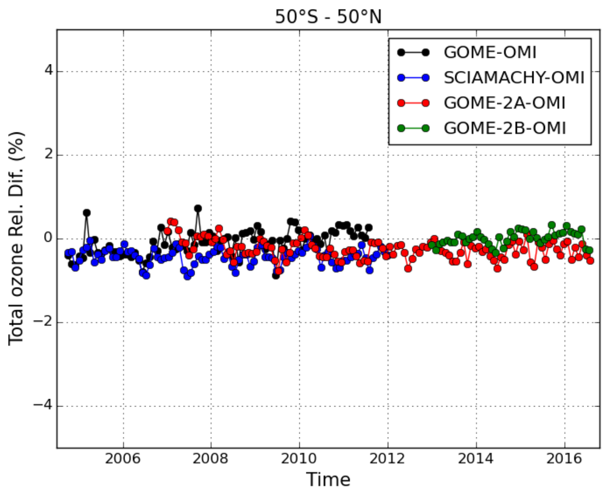


Figure 2. Time series of the relative differences between the total ozone columns retrieved from GOME, SCIAMACHY and GOME-2A/B with respect to OMI.

Using this new GODFIT v4 baseline, the time series of GOME, SCIAMACHY, GOME-2A/B and OMI have been entirely reprocessed. Figure 2~~Figure 2~~ illustrates the excellent consistency between the individual Level-2 data sets with mean differences generally within 0.5% at moderate latitudes (+/-50°). The Level-2 data sets are publicly available on the Ozone_cci website (http://www.esa-ozone-cci.org) and the time series are also regularly extended as part of the Copernicus Climate Change Service (C3S).

2.2 Ground-based Total Ozone Column records

For the purposes of this work, both direct-sun measurements (from Dobson and Brewer UV spectrophotometers) and zenith-sky scattered-light (ZSL-DOAS) measurements were used as ground-based reference data.

Total ozone column measurements from Dobson and Brewer UV spectrophotometers, were downloaded from the WOUDC (World Ozone Ultraviolet Radiation Data Center) archive (http://www.woudc.org), see Tables S2 & S3 for a complete list.

The measurement techniques and the data analysis methodology are extensively analyzed in Koukouli et al. (2015) and in references therein. It is important to point out that according to Van Roozendael et al. (1998), the estimated total uncertainty for the Dobson spectrophotometer is about 1 % for cloud-free direct Sun observations and 2 – 3 % for zenith-sky or cloudy observations, while the error of individual total ozone measurements for a well-maintained Brewer instrument is about 1 % (e.g. Kerr et al., 1988).

The main issues that have to be taken into account during the validation process with these direct-sun instruments are: (a) TOC measurements from Dobsons spectrometers depend on the stratospheric effective temperature, which is manifested in the comparisons as a seasonality effect (Kerr et al., 1988; Kerr, 2002; Basher, 1982; Bernhard et al., 2005; Scarnato et al., 2009; Koukouli et al., 2016), (b) even though the principles of operation between Dobsons and Brewers do not differ significantly, TOC measurements from the two types of instruments show small differences in the range of ± 0.6 % due to the use of different wavelengths and the different temperature dependence for the ozone absorption coefficients (Staehelin et al., 2003) and (c) due to the limited number and poor spatial distribution of stations with Brewer instruments in the Southern Hemisphere (all of them allocated in the Antarctic), the Dobson network is considered much more suitable to investigate spatial homogeneity of satellite products below the Equator.

TOC ground-based measurements from the abovementioned instruments have been extensively used in past publications for the purpose of analysis and validation of satellite data (see for e.g. Balis et al, 2007a; Balis et al, 2007b; Antón et al., 2009; Loyola et al., 2011; Koukouli et al., 2012; Labow et al., 2013; Bak et al. 2015; Koukouli et al., 2015). The ground-based stations were selected in accordance with the criteria discussed in detail in Balis et al. (2007a) and Balis et al. (2007b). Their measurements are thoroughly inspected once a year, in the aspect of quality assurance and stability, following the principles described in Fioletov et al. (1999); Vanicek (2006) and Fioletov et al. (2008), among others.

The GODFIT v4 total ozone columns were also compared against twilight zenith-sky measurements obtained with ZSL-DOAS (Zenith Scattered Light Differential Optical Absorption Spectroscopy) instruments. Most of these instruments form part of the SAOZ network (Système d'Analyse par Observation Zénitale; Pommereau & Goutail, 1988) of the Network for the Detection of Atmospheric Composition Change (NDACC). In NDACC, four slightly different ZSL-DOAS instruments are also routinely reporting data (see Table S1 for complete list of instruments used). To avoid confusion in the paper, hereafter they will all be referred to as "SAOZ measurements"

The total accuracy of SAOZ ZSL-DOAS measurements, including the cross-section uncertainties, is considered to be of the order of 6 % (Hendrick et al., 2011), including a 3 % systematic uncertainty of the absorption cross sections. However, since all NDACC SAOZ/ZSL-DOAS are using the same cross-sections, there is no systematic error between them. The random error of SAOZ spectral analysis is less than 2 % to which one should add the random error on the AMF (Air Mass Factor), mainly impacted by clouds (up to 3.3 %). Thus, significantly better performance, of the order of 2 %, can be expected in differential analyses of cloud-free data.. To avoid confusion in the paper, hereafter they will all be referred to as "SAOZ measurements".

These twilight zenith-sky measurements are complementary to the Brewer and Dobson measurements for several reasons: (a) they use spectral features of the visible Chappuis band, where the ozone differential absorption cross sections are temperature insensitive, (b) the long horizontal stratospheric optical path allows measurements of the column above cloudy scenes, and (c) measurements are always performed in the same, small, SZA range (86° - 91°). For further details on the measurement procedures and on the specific collocation approach, taking into account the actual area of measurement sensitivity, we refer to Balis et al. (2007a), Koukouli et al. (2015), and references therein. After quality control and the application of thresholds on the minimum number of collocated measurements, data from about 20 instruments were used, covering both the Northern and Southern hemisphere up to high latitudes and leaving only the equatorial region poorly sampled (see Figure S1 for the locations of all three types of instruments). In spite of the dedicated collocation method, some residual errors due to co-location mismatch may persist and must be kept in mind, in particular at high latitudes, as shown by Verhoelst et al. (2015).

2.3 Level-2 validation results and discussion

As a basis for the validation process of the satellite TOC measurements, pairs of co-located satellite and daily-mean ground-based measurements are formed and their percentage difference is calculated. Specific criteria are applied to minimize the noise of the comparison:

i. For the Dobsons and Brewers: (a) the maximum search radius between the ground-based stations and the center coordinates of the satellite pixel is set to 150 km and the spatially closest satellite observations are paired with the ground-based station's daily mean measurement and (b) only direct-sun ground-based measurements are used for the validation process, since they are deemed to be most accurate.

ii. For the SAOZ measurements, the large displacement (with respect to the instrument location) of the actual measurement sensitivity is taken into account by requiring satellite pixels to intersect with a 2-D (lat, lon) polygon describing the true area of measurement sensitivity, see Balis et al. (2007a) and Verhoelst et al. (2015) for full details.

Following those criteria, three timeseries (one for each type of ground-based instrument) of the percentage differences, are formed. Hereupon, a statistical analysis of the timeseries is performed, separately for each type of instrument, so as to study a variety of possible dependences on geospatial parameters such as the season, latitude, observation geometry, etc. The results of the analysis are shown in the following graphs and are summed up in Table 1. In the figures presented in this section, the dependency of the percentage difference between satellite and ground-based TOC measurements on parameters, such as the ones mentioned above, is displayed (the line colors used for Figure 3 to Figure 6~~Figure 6~~ are: GOME → black line; SCIAMACHY → blue line; OMI → cyan line; GOME-2A → green line and GOME-2B→ orange line). It should be noted that Southern Hemisphere GOME measurements are only shown before 2003, when it encountered downlink telemetry problems.

In Figure 3 the timeseries of the percentage difference between the ~~monthly mean~~ TOC measurements from five different satellites to the co-located Dobson, Brewer and SAOZ ground-based measurements are shown. In all panels the entire available timeseries from each satellite instrument is displayed (except for GOME for the Southern Hemisphere, as mentioned above) in the form of monthly mean difference (in %). The monthly means for each sensor were calculated using the percentage differences of all the available collocations from all stations for each month, without any weighting. The comparison with the Dobson measurements is presented in panel (a), which corresponds to the Northern Hemisphere (NH) stations, and panel (b), which presents the Southern Hemisphere (SH) percentage differences. It is shown that the NH timeseries are highly consistent and stable for all five satellites, with an amplitude of ~ 2 % for all sensors apart from SCIAMACHY, which shows a slightly increased variability with certain months under-estimating the ground-based mean (differences reaching -1 %). Part of the seasonality observed in Figure 3 – panels (a) and (b), is due to the known Dobson dependency on the effective temperature of the stratosphere (Koukouli et al., 2016). The ~ 1.5 % bias of the satellite TOCs compared to the Dobson TOCs is in agreement with the bias of ±2% found by the ‘‘Absorption Cross-Sections of Ozone” (ACSO) committee (Orphal et al., 2016) and might be related to systematic uncertainties in the different ozone absorption cross-sections used to retrieve satellite and ground-based measurements., which have been estimated to rise up to ±2% (Orphal et al., 2016). Dobson and Brewer TOC data records are based on Bass and Paur (1985) ozone absorption cross sections, whereas, as it is mentioned in the previous section, the respective satellite TOCs are produced using the cross-section measured by Serdyuchenko et al. (2014).

The comparison for the SH Dobson measurements (Figure 3, panel b) is showing higher variability due to the fact that the number of available stations in this part of the globe is limited and their measurements are greatly affected by the vigorous phenomena developing over the Antarctic. However, all timeseries present a rather consistent and stable behavior, similar to that shown in the NH, with a bias of the order of 1 - 1.5 % for OMI, GOME-2A and GOME-2B.

In Figure 3 - panel (c), the same plot of the percentage differences between the satellites and Brewer ground-based measurements performed at stations located in the NH, is shown. Due to the extremely limited number of stations with Brewer spectrophotometers in the SH, positioned exclusively on the Antarctic, it was decided not to present the respective plot. It is evident that tThe consistency and the stability of the satellite measurements is evidentexcellent for the whole time period of available data and for the whole set of five sensors: the overall bias of the comparison is up to 1 % for GOME, 0 % for SCIAMACHY and 1.5 % for the rest of the instruments, with peak-to-peak amplitude of the order of 1 – 2.5 %.

Panels (d) and (e) of Figure 3, depict the timeseries of the comparison to the SAOZ network, for the Northern and Southern Hemisphere, respectively. The knownEven though the seasonality effect, which is known to be present in comparisons between SAOZ and direct-sun measurements (Hendrick et al., 2011)-, is obviously stronger in these figures than in the other three panels., the inter-sensor consistency is evident here as well. Asides from the cross-sections’ stratospheric effective temperature dependence, affecting Dobson and lesser Brewer and satellite measurements, the SAOZ seasonality observed on panels (d) and (e), comes from the comparison performed up to high latitudes in winter, in contrast to Dobson and Brewer that are “blind” at that latitude in winter. In addition, SAOZ comparisons at high latitudes are known to be affected by co-location mismatch (Verhoelst et al., 2015). Finally, Tthe overall bias of the SAOZ comparison is fairly stable at 1.5% in the NH, but rather variable for the SH, which can be attributed to the large number of high-latitude stations contributing to the SH statistics.

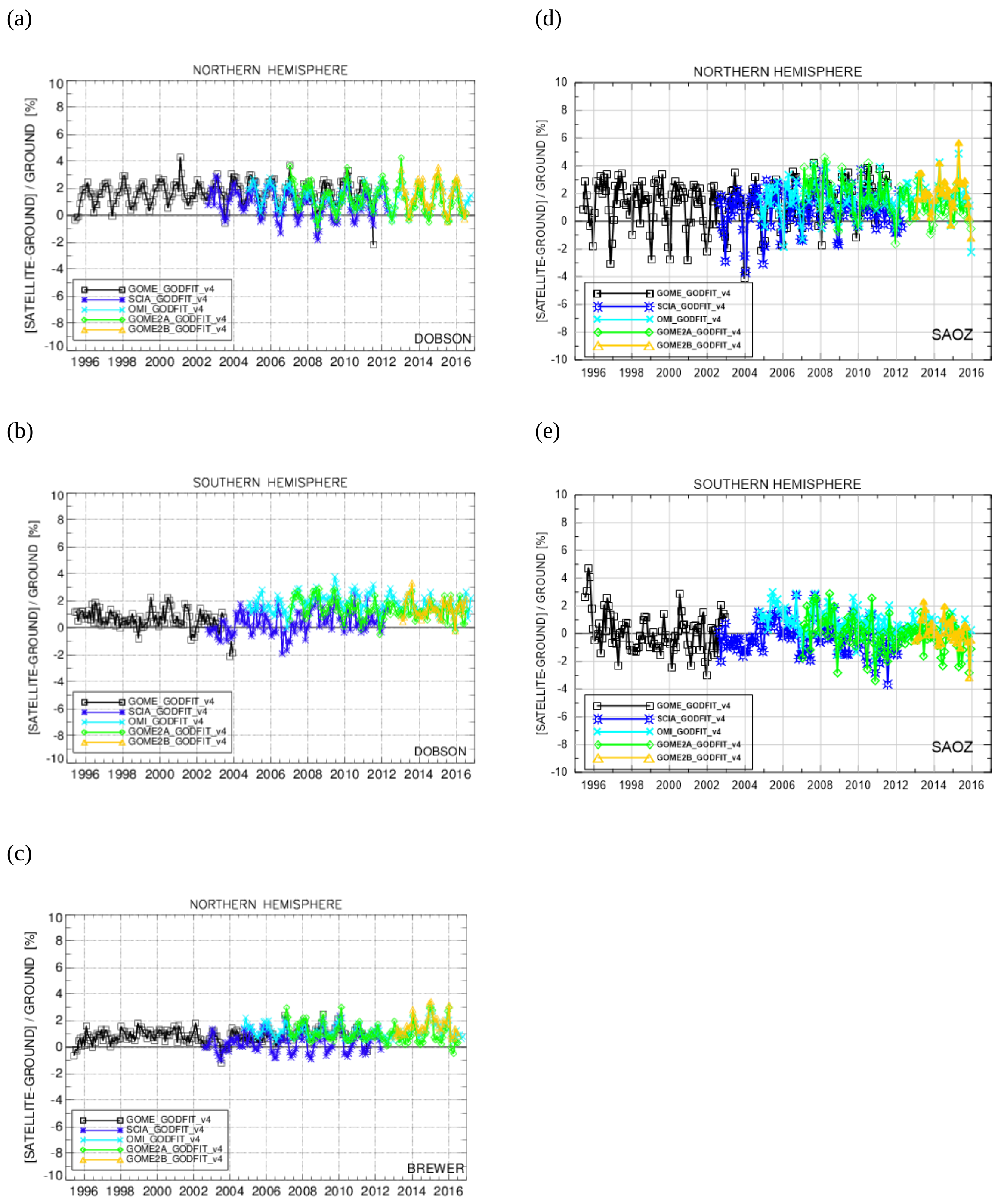


Figure 3. The time series of the monthly mean percentage differences between the five satellite instruments and the co-located ground-based TOC measurements performed by Dobsons (panel a: Northern Hemisphere and b: Southern Hemisphere), Brewers (panel c: Northern Hemisphere) and SAOZ (panel d: Northern Hemisphere and e: Southern Hemisphere) instruments.

Following, the dependence of the percentage differences of the five satellites measurements to the ground-based TOC measurements, on solar zenith angle (SZA) was investigated, as shown in Figure 4, where panels (a) and (b) depict Dobson NH and SH comparisons, panel (c) shows the Brewer – NH only comparisons and panels (d) and (e) shows the SAOZ NH and SH comparisons, respectively. It should be noted that the SZA values used for the grouping in the plots are the solar zenith angles of the satellite and not the ground based measurements, which are downloaded as daily means from the WOUDC database. We have chosen to present this plot only for Brewer (panel a – NH) and SAOZ (panels b – NH and c – SH) measurements, due to the aforementioned dependency of the Dobson measurements to the stratospheric effective temperature, which is highly correlated with SZA. Firstly, as it is seen in Figure 4, all curves in each plot have highly consistent dependencies on SZA, which proves that, irrespective of its magnitude, the dependence can be contributed mainly on the ground based measurements of each kind.

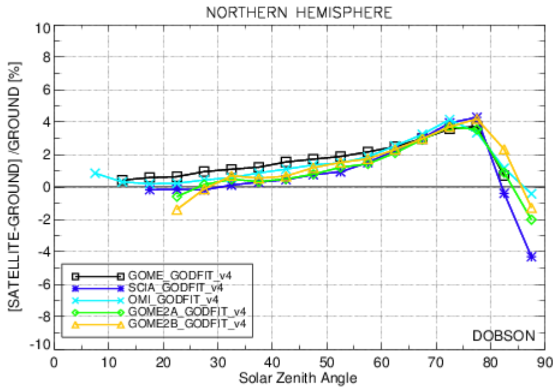
Specifically, in panel (a) where the NH comparison is shown, there is a strong but very consistent dependence on SZA for all five satellite instruments, whereas in the SH (panel b) almost no dependency is seen for SZAs < 80°. The first reason for this dissimilar behavior is the fact that in the NH most Dobson ground based stations are located in the middle latitudes, contrary to the SH stations that are much more homogeneously distributed. Additionally, since the measurements of the Dobson stations are affected by the variation of the stratospheric effective temperature, the data provided by NOAA/National Weather Service (http://www.cpc.ncep.noaa.gov/products/stratosphere/temperature/) were investigated to see whether there is a difference in the stratospheric temperature between the mid-latitudes of the two hemispheres. The results are very consistent with the two plots of the panels (a) and (b): the peak-to-peak amplitude of the stratospheric effective temperature annual variation above the mid-latitudes of the NH is about 3 – 4°C greater compared to the variation above the respective latitudes of the SH, which resulted to the stronger variability of the NH Dobson measurements, seen in panel (a).-Of course, further investigation on this issue is needed, but it is beyond the scope of this work.

Figure 4 - panel (c) shows that the percentage difference of the measurements is almost constant for the Brewer comparison and it is only increasing for SZAs larger than 70°. SCIAMACHY however shows a slightly stronger dependence on SZA starting from low angles. Comparisons performed at SZAs over 75° and below 25° are affected by the limited number of observations and the uncertainties of the ground-based measurements themselves. Hence it is difficult to assess their significance level.

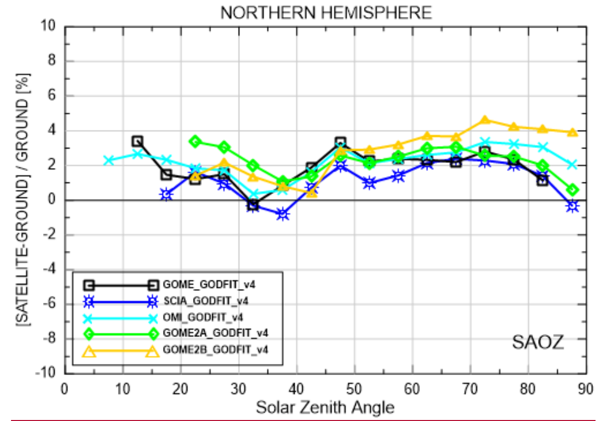
In Figure 4 – panels (db) and (ce), we show that the SZA dependence between satellite and SAOZ ground measurements was up to 4 % at the highest satellite-viewed SZAs (>80°) at all high-latitude stations, irrespective of season. There was also some minor dependence at very small SZAs in the Northern Tropics, but this is based on only a few tropical stations with limited data, and it is not confirmed by the Brewer comparisons. There are also some systematic inter-hemispheric differences for SAOZ measurements, which is obvious when comparing panels (bd) and (ce) of Figure 4, in particular due to comparisons at some Northern high-latitude stations being biased high (up to 5%), and those at Southern high-latitude stations being biased low (of the order of 2%), as shown in Figure 5 – panel (c) that will be commented on below.

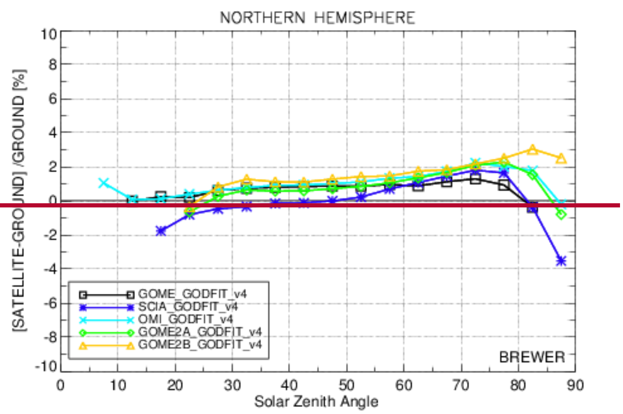
Additionally, the dependency of the satellite and ground-based measurements percentage difference on latitude, is presented. In Figure 5 - panel (a), the ground-based measurements are performed by Dobson spectrophotometers, in panel (b) Brewer data are used, while in panel (c) the comparison with the SAOZ data record is shown. It is obvious in this exercise too, that all five satellite sensors appear to be very consistent, regardless of the ground-based instrument type, which is the main concern of this work. It is also noticeable that, mainly for Brewer and Dobson ground-based measurements, the dependency on latitude is less eminent for the NH due to the much higher number of collocations found there. Specifically, the comparisons with Dobson measurements show differences between 0 and 2 % for latitudes between -40° and 0° as well as for the entire NH, similar to the Brewer comparisons. In the SH, especially Southwards of -40°, the comparisons show differences ranging between -2 and 4 %, depending on the satellite sensor, partially attributed to the small number of stations located in that part of the Earth and partially to the higher variability of the TOCs within the Southern polar vortex (see also Verhoelst et al.,

(a)

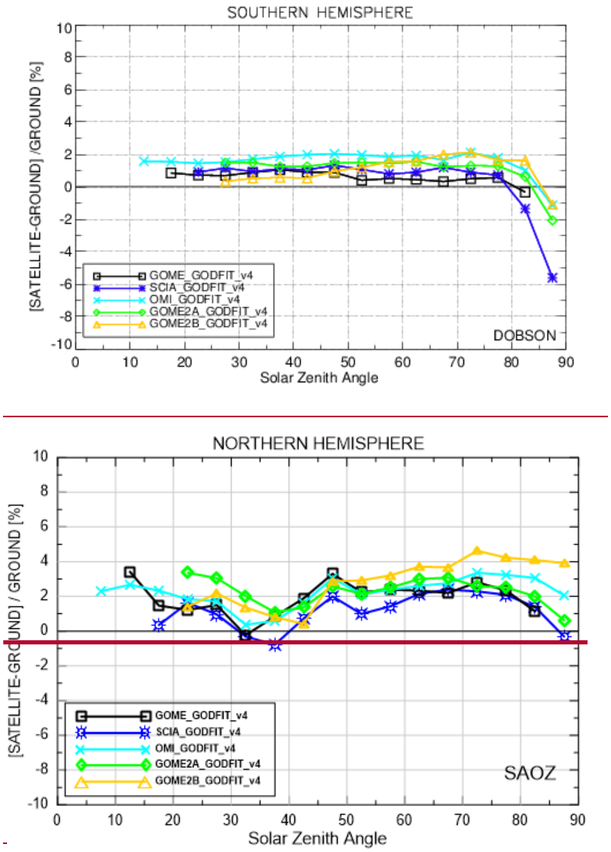

NORTHERN HEMISPHERE
[SATELLITE-GROUND] /GROUND [%]
Solar Zenith Angle
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
DOBSON

(d)

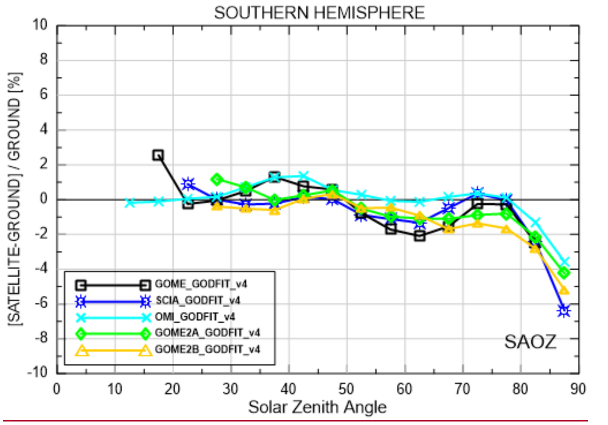

NORTHERN HEMISPHERE
[SATELLITE-GROUND] / GROUND [%]
Solar Zenith Angle
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
SAOZ


NORTHERN HEMISPHERE
[SATELLITE-GROUND] /GROUND [%]
Solar Zenith Angle
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
BREWER

(b)


SOUTHERN HEMISPHERE
[SATELLITE-GROUND] /GROUND [%]
Solar Zenith Angle
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
DOBSON
NORTHERN HEMISPHERE
[SATELLITE-GROUND] / GROUND [%]
Solar Zenith Angle
SAOZ

(e)


SOUTHERN HEMISPHERE
[SATELLITE-GROUND] / GROUND [%]
Solar Zenith Angle
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
SAOZ

(c)

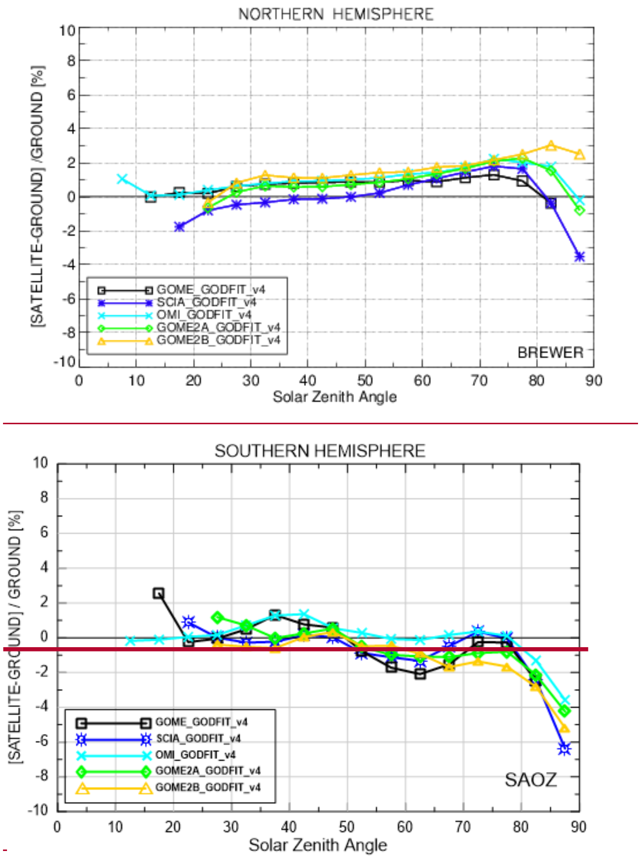
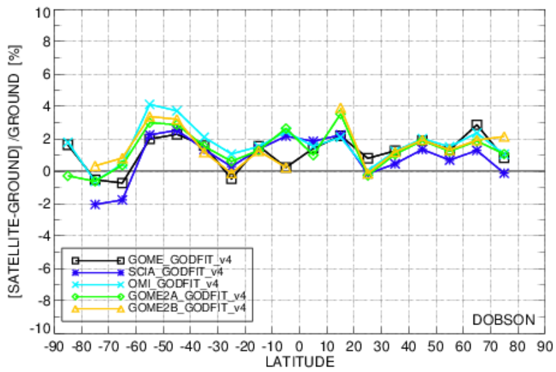
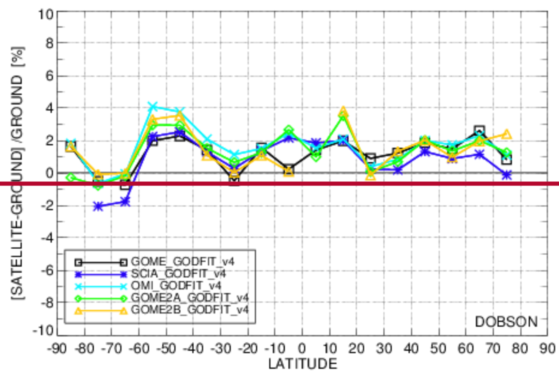


Figure 4. The dependence of the percentage differences between the satellites TOC measurements to the measurements of the <u>Dobson (panels a-NH and b-SH),</u> Brewer (panel ~~a~~<u>c</u>, NH only) and SAOZ (panels ~~b~~<u>d</u>-NH and ~~c~~<u>e</u>-SH) ground-based stations, on solar zenith angle.

(a)

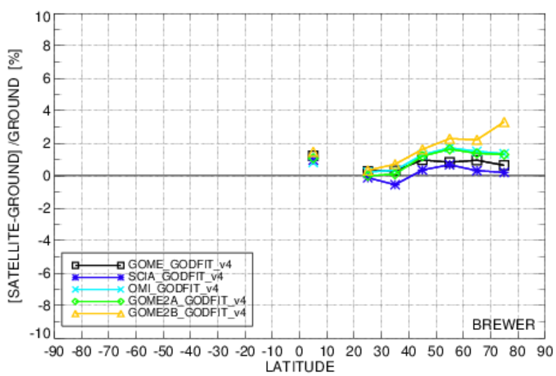


[SATELLITE-GROUND] /GROUND [%]
LATITUDE
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
DOBSON

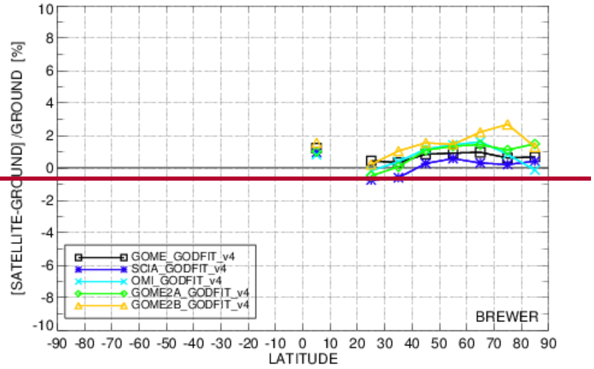


[SATELLITE-GROUND] /GROUND [%]
LATITUDE
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
DOBSON

(b)



[SATELLITE-GROUND] /GROUND [%]
LATITUDE
GOME_GODFIT_v4
SCIA_GODFIT_v4
OMI_GODFIT_v4
GOME2A_GODFIT_v4
GOME2B_GODFIT_v4
BREWER



(c)

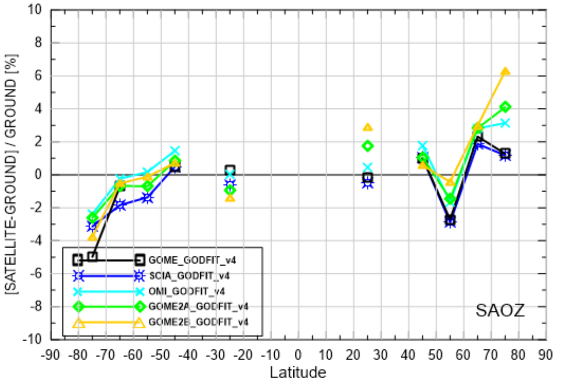


Figure 5: The percentage difference between the five satellites TOC measurements and ground-based measurements from Dobson (panel a), Brewer (panel b) and SAOZ (panel c) instruments, as a function of latitude.

2015). In Figure 5 - panel (c), where the comparison with the SAOZ measurements is shown, a higher dependency on latitude is eminent even for the NH, where the other two ground-based instruments have completely different performances. Nevertheless, the inter-sensor consistency is very satisfactory in this comparison too, except for the high altitude Izaña station located at 28°N (near the NH tropics), for which the differences were adjusted to take into account the missing column in the ground-based measurement but some residual effect due to different satellite pixel sizes is probably still present. The correction for the station's altitude is described in Verhoelst et al. (2015) and uses an ERA-Interim based estimate of the column below the instrument altitude in the immediate vicinity of the island and/or mountain, at the resolution of the reanalysis and not taking into account the exact satellite pixel size and location. For the SAOZ/ZSL-DOAS network, Izaña and Jungfraujoch are the only stations for which a significant missing column was derived with this methodology (about 2.8 % and 3.2 % respectively, with some seasonal variation), due to their isolated mountain-top locations. Any pixel-size dependence at Jungfraujoch is less evident in Figure 5 - panel (c) as that latitude bin contains three other stations not located on mountain tops. Moreover, the measurements performed by the stations located in the belt 70° - 80°N show larger differences between sensors, but these discrepancies are not confirmed by the Brewer or the Dobson networks and they are most probably related to the larger (and pixel-size dependent) horizontal smoothing difference errors between SAOZ and the satellite measurements.

According to the guidelines given at the Ozone_cci project's User Requirement Document (Version: 2.1) (van der A, 2011), Table 5, the stability of the total ozone column measurements must be among 1 and 3 %/decade, the evolution of the ozone layer (radiative forcing) has to be less than 2 % and the seasonal cycle and inter-annual (short-term) variability should be less than 3 %. To investigate whether the five satellite data records are compliant to those requirements, a statistical analysis of the percentage deviation between satellite and ground-based measurements was performed, with the statistics presented in Table 1. The first column enumerates the physical quantity studied, the second column differentiates between Brewer, Dobson and SAOZ collocations, the third column shows the results of the statistical analysis for GOME/ERS-2, the fourth column for SCIAMACHY/Envisat, the fifth for OMI/Aura, the sixth for GOME-2/MetopA and the seventh for GOME-2/MetopB sensor. The rows of Table 1 depict: (a) the ~~Monthly~~ Mean Bias and standard deviation (1 sigma), computed from the monthly mean differences of the entire record for each sensor, shown in Figure 3, (b) the Monthly mean variability, i.e. ~~the variability of the monthly mean standard deviation values calculated by the Root Mean Square (RMS)~~ the variability of the standard deviations of differences in individual months, calculated by the Root Mean Square (RMS), (c) the Drift per decade: the decadal drift and associated standard deviation, (d) the Seasonality: the peak-to-peak amplitude of the seasonal variability, (e) the Latitud~~e~~inal mean bias: the mean bias and standard deviation as calculated by the latitudinal variability plots (Figure 5) on a global scale, (z) the Solar Zenith Angle mean bias: the mean bias and standard deviation as calculated from the solar zenith angle ranges shown in Figure 4, on a global scale. The values of the Table are all measured in percent and all the quantities for the Brewer measurements, as well as quantities (a), (b)~~,~~ and (c) ~~and (d)~~ for the Dobson measurements are calculated for the NH only.

The percentages listed in Table 1 prove that the products of the GODFIT v4 algorithm for all five sensors fulfill the requirements set by the European Space Agency's Ozone_cci project (Lambert et al., 2018), since the amplitude of the short term variability (seasonality) is less than 2 % and the maximum drift per decade is equal to -1.37 ± 1.60 %/decade for GOME-

2/MetopB, whose time series is only 3.5 years long and as a result its drift/decade cannot be considered statistically significant. For the rest of the sensors the maximum drift per decade is less than ±1 %. In conclusion, the statistics presented in Table 1 indicate that the data sets produced by the Ozone_cci GODFIT v4 algorithm for all five sensors under validation are reliable, homogeneous and consistent.

Table 1: Statistics of the comparison between satellite ground-based TOC measurements.

			GOME/ ERS-2 (%)	SCIAMACHY/ Envisat (%)	OMI/ Aura (%)	GOME-2/ MetopA (%)	GOME-2/ MetopB (%)
Monthly mean bias and 1-sigma		Dobson*	1.62 ± 0.87	0.88 ± 1.01	1.26 ± 0.81	1.20 ± 1.04	1.45 ± 1.08
		Brewer*	0.83 ± 0.51	0.43 ± 0.80	1.18 ± 0.50	1.08 ± 0.75	1.59 ± 0.69
		SAOZ	1.07 ± 1.46	0.41 ± 1.00	1.00 ± 0.86	0.56 ± 1.10	0.57 ± 1.02
Monthly mean variability		Dobson*	±3.16	± 3.22	± 3.16	± 3.30	± 3.16
		Brewer*	± 3.06	± 2.92	± 2.82	± 2.92	± 3.08
		SAOZ	± 2.40	± 2.43	± 2.25	± 2.31	± 2.19
Drift per decade		Dobson*	0.08 ± 0.13	-0.61 ± 0.33	-0.41 ± 0.19	-0.71 ± 0.35	-1.37 ± 1.60
		Brewer*	0.21 ± 0.08	0.33 ± 0.26	0.01 ± 0.12	-0.61 ± 0.25	0.99 ± 1.02
		SAOZ	0.51 ± 1.92	-0.14 ± 2.43	0.48 ± 1.54	-1.32 ± 1.82	-1.00 ± 4.43
Seasonality (peak – to – peak)		Dobson	N/A	N/A	N/A	N/A	N/A
		Brewer*	0.85	2.00	0.97	1.56	1.22
		SAOZ	N/A	N/A	N/A	N/A	N/A
Latitudinal mean biase		Dobson	1.158 ± 1.040.99	0.85 ± 1.374	1.675 ± 1.2018	1.334 ± 1.204	1.5044 ± 1.3326
		Brewer*	0.745 ± 0.2935	0.2617 ± 0.507	1.030.74 ± 0.670	0.8895 ± 0.7164	1.570 ± 1.010.74
		SAOZ	0.69 ± 2.67	1.34 ± 3.14	0.22 ± 2.94	1.61 ± 4.55	0.82 ± 3.18
Solar Zenith Angle mean bias	<70°	Dobson	1.19 ± 0.48	0.79 ± 0.65	1.35 ± 0.67	1.02 ± 0.73	0.97 ± 1.06
	<70°	Brewer*	0.67 ± 0.35	-0.02 ± 0.89	0.88 ± 0.49	0.70 ± 0.63	1.17 ± 0.61
	<70°	SAOZ	0.84 ± 0.68	0.40 ± 0.57	1.17 ± 0.32	1.10 ± 0.49	0.91 ± 0.55
	>70°	Dobson	1.03 ± 1.14	-0.92 ± 3.29	1.37 ± 1.71	0.88 ± 1.92	1.77 ± 1.45
	>70°	Brewer*	0.61 ± 0.88	-0.12 ± 2.48	1.45 ± 1.12	1.27 ± 1.42	2.55 ± 0.36
	>70°	SAOZ	0.49 ± 0.87	-0.11 ± 1.94	1.02 ± 1.06	0.16 ± 1.18	0.78 ± 0.87

* NH only

In order to further demonstrate the long-term inter-sensor consistency of the GODFIT v4 Level-2 total ozone columns, comparisons to the Solar Backscatter Ultraviolet measurements (SBUV) data products are shown. Daily Level-2 overpass files of total ozone column measurements produced by the SBUV v8.6 algorithm for the locations of the ground-based stations, were downloaded from https://acd-ext.gsfc.nasa.gov/Data_services/merged/ and are described by McPeters et al. (2013) and Frith et al. (2014). The instruments and the respective time periods of measurements used for this comparison are: NOAA 14 SBUV/2 (February 1995 to March 2006), NOAA 16 SBUV/2 (October 2000 to May 2014), NOAA 17 SBUV/2 (July 2002 to March 2013), NOAA 18 SBUV/2 (July 2005 to November 2012) and NOAA 19 SBUV/2 (April 2009 to February 2017). As reported by Labow et al. (2013), their measurements were also validated against Brewer and Dobson ground-based measurements, showing an agreement of the order of ± 1 %.

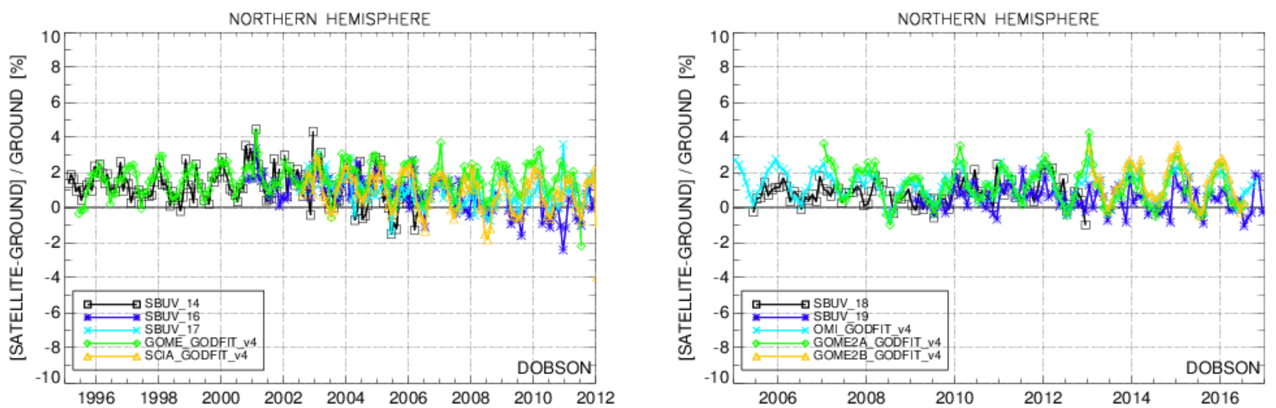


Figure 6: The timeseries of the percentage differences between satellite and ground-based monthly mean TOC measurements at Northern Hemisphere, separated into two time periods: 1995 – 2012 (left panel) and 2005 – 2017 (right panel). To the left: NOAA 14 SBUV/2 (black line), NOAA 16 SBUV/2 (blue line), NOAA 17 SBUV/2 (cyan line), GOME GODFIT v4 (green line) and SCIAMACHY GODFIT v4 (orange line). To the right: NOAA 18 SBUV/2 (black line), NOAA 19 SBUV/2 (blue line), OMI GODFIT v4 (cyan line), GOME-2A GODFIT v4 (green line) and GOME-2B GODFIT v4 (orange line).

In Figure 6Figure 6 the percentage deviation of Northern Hemisphere SBUV and GODFIT v4 satellite data sets from the respective ground-based measurements performed by Dobsons, is displayed. In the left panel, the time period 1995 to 2012 is shown, encompassing the available data sets from NOAA 14 SBUV/2, NOAA 16 SBUV/2, NOAA 17 SBUV/2, GOME and SCIAMACHY. In the right panel, the time series of NOAA 168 SBUV/2, NOAA 19 SBUV/2, OMI, GOME-2A and GOME-2B for the years 2005 to 2017 are shown. The purpose of these plots is to investigate the consistency, the stability and the homogeneity of ten completely different time series generated with two different algorithms. It is well shown that, for the two time periods under consideration, all sensors are in very good agreement, with very similar seasonality amplitudes and biases, further testifying to the homogeneity and stability of the GODFIT v4 products.

3 Level-3 Total Ozone Columns

3.1 The Level-3 GTO-ECV data record

One of the main aims of the ESA Ozone_cci project is to construct the homogeneous global long-term GOME-type Total Ozone Climate data record, hereafter termed GTO-ECV version3. The individual Level-2 observations (presented and validated above in Section 2) are converted into a Level-3 product and then combined into one single cohesive record spanning the entire 221-years period, from 1995 to 20176. This section summarizes the main characteristics of the merging methodology as well as the latest improvements and extensions implemented within the second phase of the Ozone_cci project. A detailed description of the predecessor of GTO-ECV v3 has been presented and validated in Loyola et al. (2009) and Coldewey-Egbers et al. (2015).

In short, at first, the individual Level-2 measurements processed with the GODFIT v4 retrieval algorithm are mapped onto a regular global grid of 1°x1° in latitude and longitude to construct daily averages for each sensor. Before combining the individual gridded data, adjustments are made in order to account for possible biases and drifts between the instruments. In the previous algorithm version, which spanned the 15-years period between March 1996 and June 2011 (Coldewey-Egbers et al., 2015), the GOME TOCs were used as a reference to the other sensors; in this version the OMI measurements serve as a baseline for the inter-sensor calibration. Their long-term stability with respect to ground-based observations data is ~~excellent~~ noteworthy (see Figure 3 – panels (a) and (c) and Table 1) and the periods of overlap with the other sensors sufficiently long, at least 4 years.

Figure 7 shows the percentage differences between OMI and the other four sensors for 1° zonal monthly mean ozone columns during overlap periods. These zonal means were computed for collocated daily gridded data in order to minimize the impact of differences in the sampling pattern for OMI and the corresponding second sensor. In general, the inter-sensor consistency is very good; mean differences are between -0.2 ± 0.9 % (for GOME-2B, lower right) and 1.0 ± 1.4 % (for SCIAMACHY, upper right). In the inner tropics the bias is slightly negative for all sensors and it increases toward higher latitudes. The differences between OMI and GOME show slightly larger scatter in the Southern Hemisphere due to significantly reduced spatial coverage of GOME as a consequence of the tape recorder failure in June 2003. The differences between OMI and SCIAMACHY indicate a positive bias for most parts of the Globe, with a maximum in the southern hemisphere around the polar night. For both GOME and SCIAMACHY we apply a static correction that depends on latitude and month of the year ~~factors~~ using the seasonal mean differences~~,~~ calculated from the seasonal mean average of all available years, with respect to OMI as a function of latitude. The differences between OMI and GOME-2A indicate a positive drift of ~ 0.15 % per annum in the middle latitudes of both hemispheres, which we take into account during the adjustment. ~~Likewise for~~ For both GOME-2A and GOME-2B, the correction factors with respect to OMI depend on time (month) and latitude. The adjustment is then applied to the daily gridded data for each individual sensor. Thereby the monthly correction factors are linearly interpolated in time.

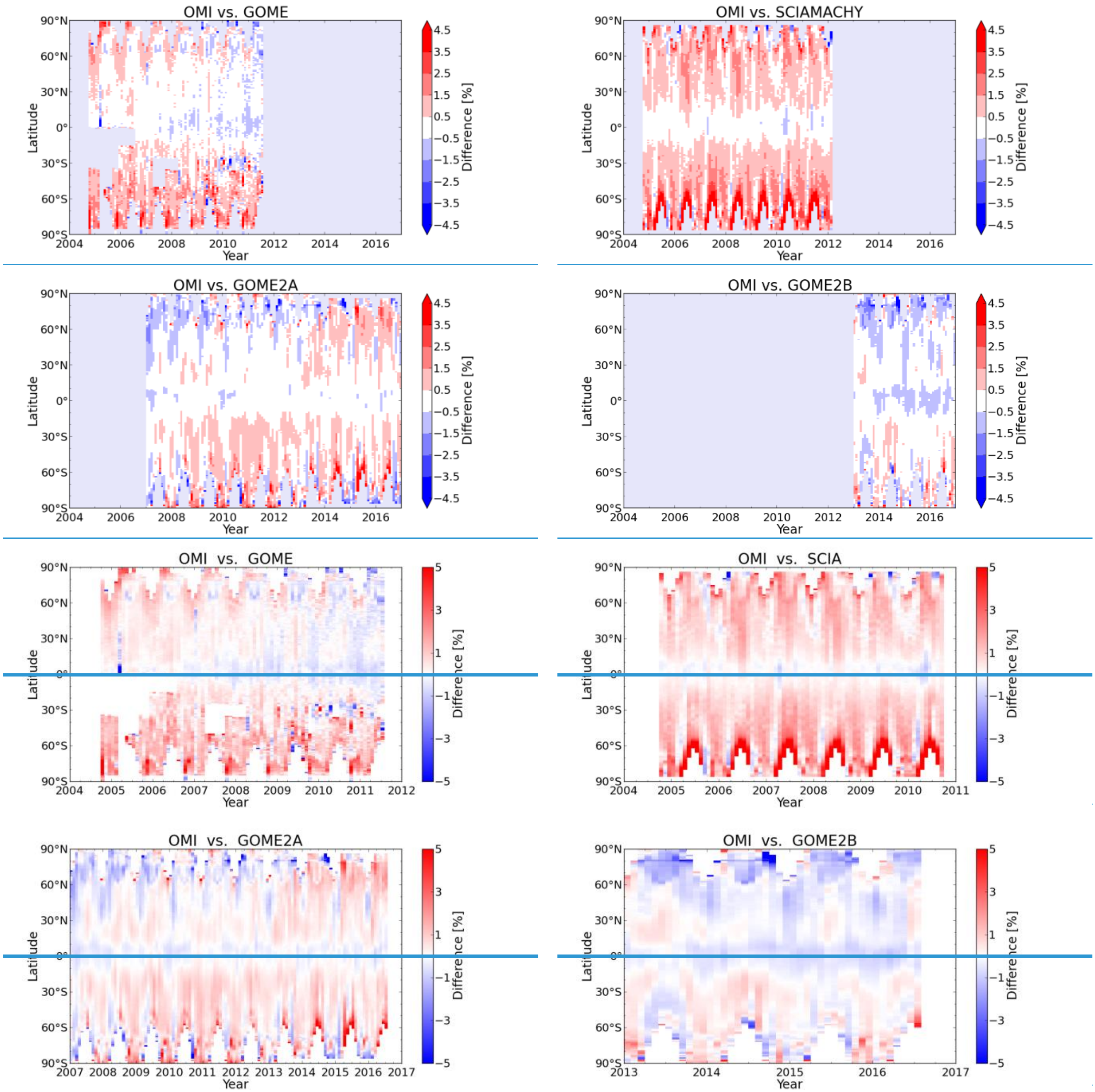


Figure 7. Percentage differences between OMI and the other four sensors for 1° zonal monthly mean ozone columns during overlap periods. Top left: GOME, top right: SCIAMACHY, bottom left: GOME-2A and bottom right: GOME-2B.

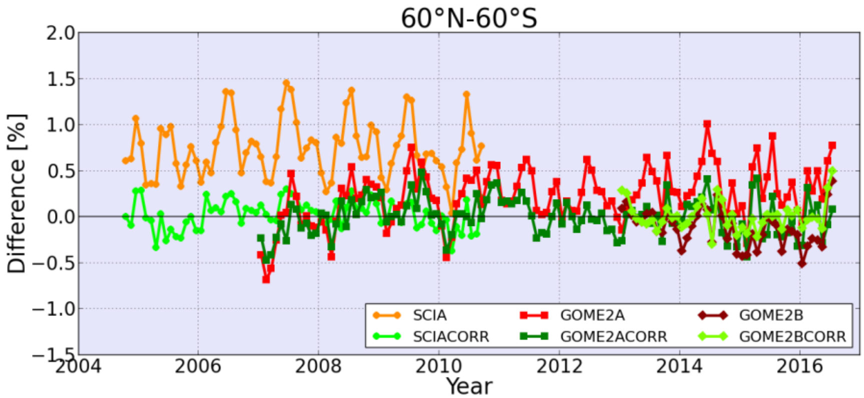


Figure 8. Percentage differences between SCIAMACHY and OMI (circles), GOME-2A and OMI (squares), and GOME-2B and OMI (diamonds) as a function of time for the periods of overlap. Orange-reddish curves denote the differences without adjustment to OMI, and greenish curves denote the differences after the adjustment to OMI.

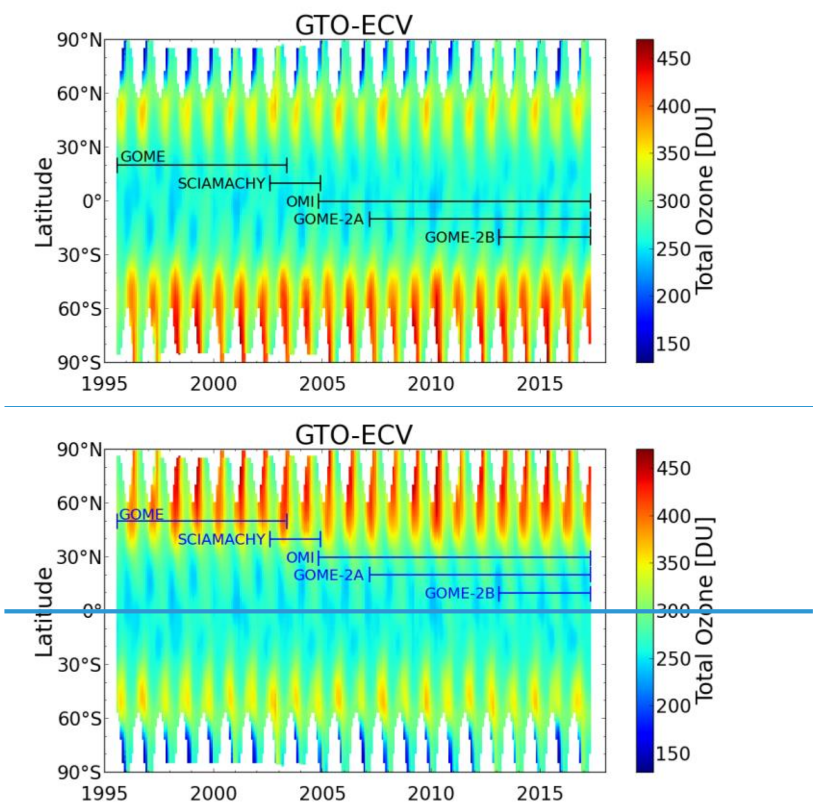


Figure 9. GTO-ECV total ozone column data record as a function of latitude and time from July 1995 to March 2017. Blue horizontal lines indicate the period for each sensor included in the merged product.

Figure 8~~Figure 8~~ shows the percentage differences between OMI and the other sensors without (orange-reddish curves) and with (greenish curves) the adjustment to OMI for the near global (60° N - 60° S) mean ozone column as a function of time during the periods of overlap. The comparison with GOME is omitted in this plot because we use these data only until June 2003 in the final product. After the application of the correction the mean biases are almost completely reduced, the scatter (standard deviation) decreased by 15 – 40 % and the drift in the differences between GOME-2A and OMI is eliminated. Subsequently, the individual (adjusted) data sets are combined into one single record. In contrast to the previous version (Coldewey-Egbers et al., 2015), where we used only one instrument at any given time, in GTO-ECV v3 we now average all available daily measurements (weighted by the number of measurements per day and grid box for the corresponding sensor), which improves the representativeness of the monthly averages. GOME data are restricted to up and until June 2003. As the ground-based validation of SCIAMACHY Level-2 data indicates some lingering issues with the Level-2 TOCs (see Sec. 2.3) we use SCIAMACHY only until October 2004 in order to fill the data gap between the GOME loss of global coverage and the launch date of OMI. For the calculation of monthly means we apply the same latitudinal constraints as defined in Coldewey-Egbers et al. (2015), see their Table 2, in order to provide representative averages that contain a sufficient number of measurements equally distributed over time. The complete merged GTO-ECV v3 data record with typical ozone characteristics is shown in Figure 9~~Figure 9~~. Highest ozone values occur in northern hemispheric springtime, whereas monthly mean values are below 200 D.U. from September to November southwards of 70° S. Blue horizontal lines indicate the period for each sensor included.

3.2 Level-3 validation results and discussion

The validation of the new Level-3 GTO-ECV v3 merged product was performed using as ground truth the Brewer and Dobson spectrophotometer network described in Section 2.2, as was applied in the validation of the previous Level-3 record (Coldewey-Egbers et al., 2015). In order to create the Level-3 TOC field, based on the WOUDC ground-based stations, the reported TOCs were gridded into the same 1°x1° grid as the GTO-ECV v3 data, on a monthly basis, with most grid points being represented by only one reporting station. In detail, direct Sun measurements were considered for the gridding of the ground-based TOCs into Level-3 grid points, even though in some cases this choice severely decreases the number of measurements. As also performed in Coldewey-Egbers et al., 2015, 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 per grid box was enforced so that the temporal representativeness errors are minimized. We note here that restricting the monthly collocated measurements with respect to their mean effective day, which is a measure for the temporal distribution of the daily measurements within a month, did not alter significantly the findings, whereas it excluded entire zones and months from the comparative process and we opted not to apply such a restriction here.

Figure 10 shows the percentage difference between the satellite and the Brewer (left) and Dobson (right) TOC records as a function of latitude. The five individual satellite TOCs are very consistent with each other for all latitudes and in very close agreement with the ground-based data. The Level-3 comparisons (purple line) show very good agreement with the individual Level-2 lines. In particular, over the NH, all Level-2 comparisons (apart from SCIAMACHY, in green) show a slight positive deviation of 0 – 2 % to the ground-based data for both ground-based instrument types. In the SH the Level-3 comparisons show a near-perfect agreement with the Level-2 comparisons, apart from the 70° - 80° S belt, where the spread in comparisons reaches the 3.0 % level, which may be attributed to sampling differences between the Level-2 and Level-3 data (see Coldewey-Egbers et al., 2015 for more in-depth discussion of this issue).

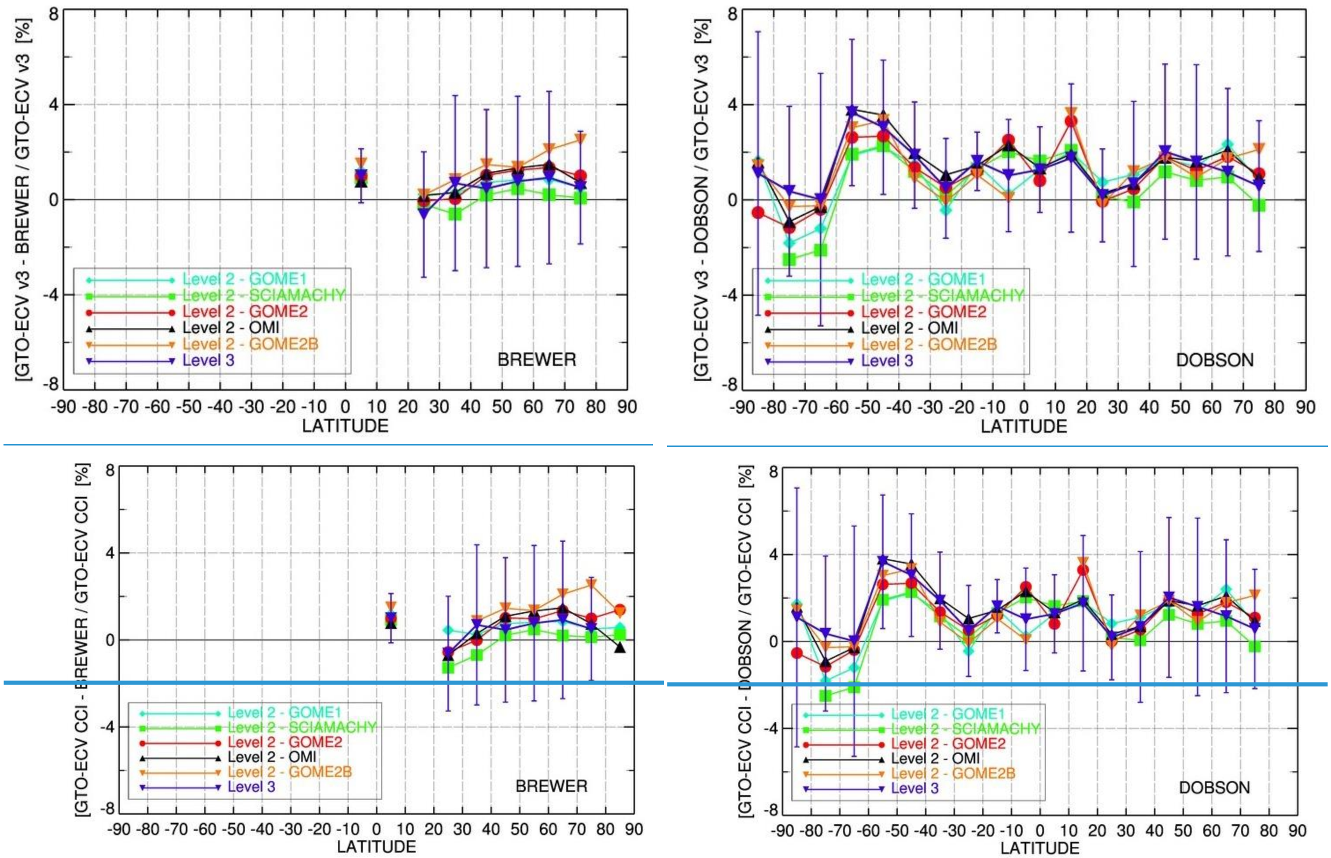


Figure 10: Latitudinal variability of the percentage difference between satellite observations and ground-based measurements. Left: for the Brewer network and right: for the Dobson network. Light blue line: GOME Level-2 comparison, green line: SCIAMACHY Level-2 comparison, red line: GOME-2A Level-2 comparison, black line: OMI Level-2 comparison, orange line: GOME-2B Level-2 comparison and purple line: Level-3 GTO-ECV v3 comparison. The 1-σ standard deviation of the average is also displayed only for the Level-3 lines.

In Figure 11 the NH and SH timeseries comparisons of the Level-2 and Level-3 data records with the Dobson and Brewer measurements are shown. The Dobson comparisons for SH (panel a) and NH (panel b) show very good agreement between Level-3 and individual Level-2 lines, within the 1 % difference level for most of the 21~~2~~-year data record, except for a small number of outliers. The Brewer comparison in the NH (panel c) shows less amplitude than the Dobson comparisons throughout the full time series, for reasons discussed already in Sections 2.2 and 2.3.

The agreement between the five datasets and the ground-based measurements is ~~excellent~~ outstanding, with 0.5 to 1.5 % peak-to-peak amplitude. For the entire time series of the Level-3 data record the mean difference remains mainly positive for all time-series comparisons shown in Figure 11. Concerning the Level-3 comparisons in the NH, the drift per decade of the differences with respect to ground-based data is negligible, -0.11 ± 0.10 % per decade for Dobson and +0.22 ± 0.08 % per decade for Brewer collocations. Similarly to Level-3, no long-term drift in the differences of the individual Level-2 data sets was found for either Dobson and Brewer comparisons, with OMI showing the smallest drift per decade (in the NH: +0.05 ± 0.12 % for Brewer and -0.39 ± 0.19 % for Dobson, in the SH: -0.15 ± 0.15 % for Dobson measurements). The good quality of the GTO-ECV v3 Level-3 TOC record temporal stability, which well satisfies the requirements for the long term stability for total ozone measurements of between 1 – 3 % per decade (van der A et al., 2011) and the excellent inter-sensor consistency, make the new Level-3 GTO-ECV v3 dataset suitable and useful for longer term analysis of the ozone layer, such as decadal trend studies (e.g. Coldewey-Egbers et al., 2015), the evaluation of chemistry-climate model projections, and data assimilation applications.

In order to assess and ensure the quality of the new Level-3 GTO-ECV v3 dataset, comparisons are performed against the solar backscatter ultraviolet (SBUV) merged data product, also shown above in the Level-2 TOC validation section and recently quality assured in Frith et al. (2017). In Figure 12, the time series comparison between GTO-ECV v3 and SBUV merged are presented for the NH and Dobson (panel a), the SH and Dobson (panel b) and the NH and Brewer (panel c) instrument types. The Level-3 GTO-ECV v3 (red line) and SBUV merged (black line) datasets show an ~~excellent~~ agreement of within ± 1.5 %, considering their individual instrumental and algorithm differences, as well as a very similar seasonal variability with a peak-to-peak amplitude between -1 % and +2 % in Dobson and -0.5 % and +1 % in Brewer cases over the entire time period. Furthermore, the two datasets show almost the same negligible drift per decade in the NH for both ground-based instrument networks, whereas in the SH for Dobson collocations the drift per decade is +0.23 ± 0.09 % and -0.09 ± 0.07 % for the Level-3 GTO-ECV v3 and the SBUV merged TOCs, respectively.

(a)
(b)
(c)

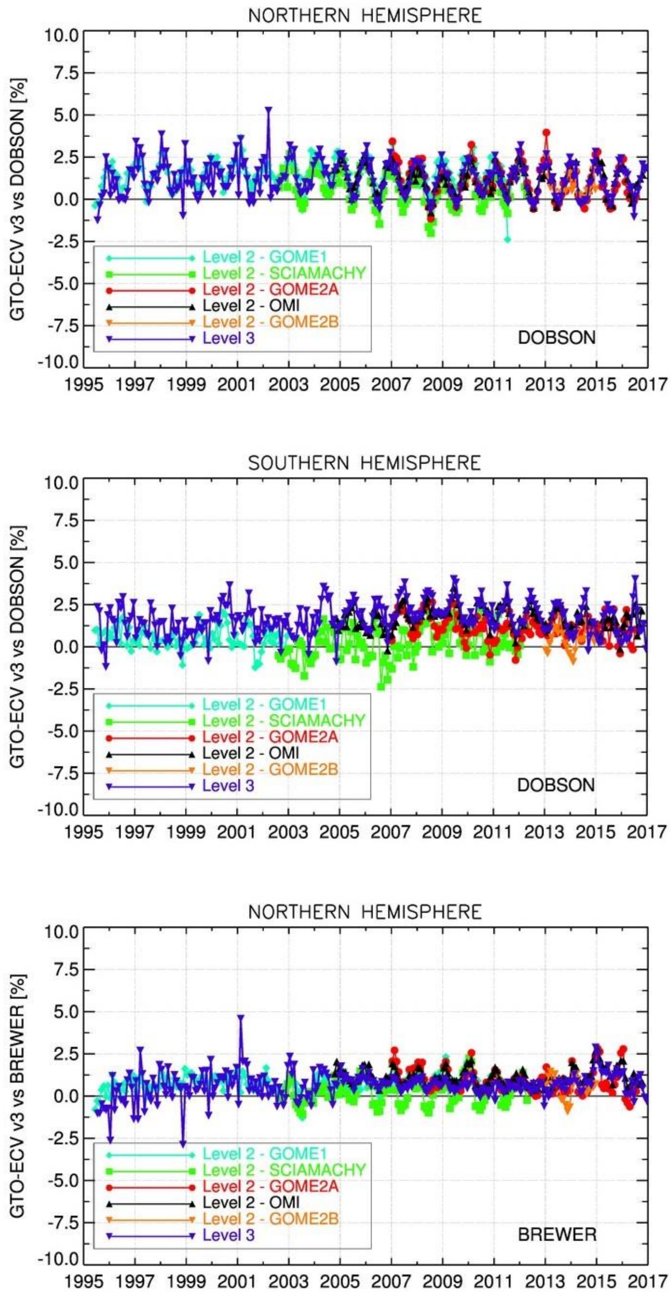
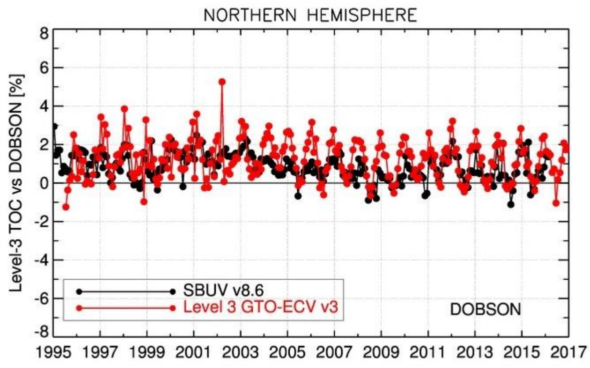
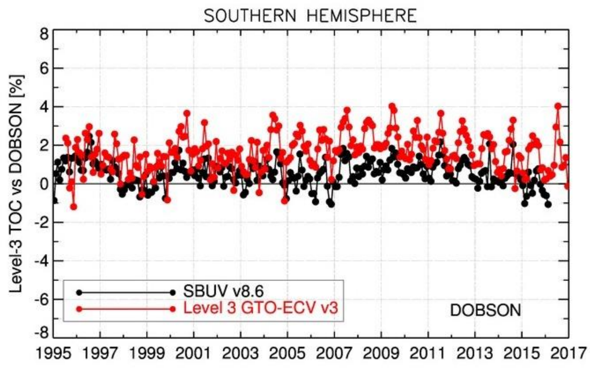


Figure 11: Time series of the percentage difference between satellite observations and ground-based measurements for the Dobson network in the NH (panel a) and in the SH (panel b) and for the Brewer network, NH only (panel c). Light blue line: GOME Level-2 comparison, green line: SCIAMACHY Level-2 comparison, red line: GOME-2A Level-2 comparison, black line: OMI Level-2 comparison, orange line: GOME-2B Level-2 comparison and purple line: Level-3 GTO-ECV v3 comparison.

(a)



(b)



(c)

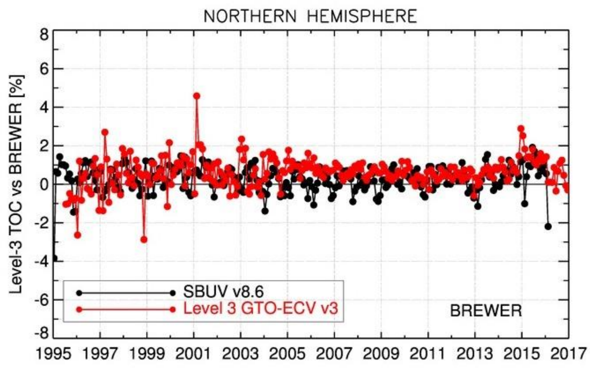


Figure 12: Same as in Figure 11. Black line: SBUV merged comparison and red line: Level-3 GTO-ECV v3 comparison.

4 Summary and conclusions

In this work, the Essential Climate Variable (ECV) Climate Research Data Package Total Ozone Column (CRDP TOC), refined and updated via the European Space Agency's *Climate Change Initiative* Phase-II, is presented and validated against independent ground-based TOC observations. Level-2 TOCs, produced by the GODFIT v4 algorithm as applied to the GOME/ERS-2, OMI/Aura, SCIAMACHY/Envisat and GOME-2/MetopA and /MetopB observations, form the basis for a 2~~2~~1-year long consistent, smooth and homogeneous CRDP. In addition, the individual sensor products have been combined and merged into one single cohesive Level-3 data record, GTO-ECV v3. Detailed quality control and assurance against specific requirements from the international climate-chemistry modelling community showed that the product more than meets the official User Requirements, i.e. that the stability of the TOC measurements has to be between 1 and 3 % per decade, that the radiative forcing introduced by the evolution of the ozone layer has to be less than 2 % and that the short-term variability has to be less than 3 %. In detail:

- The individual Level-2 data sets show excellent inter-sensor consistency with mean differences within 1.0 % at moderate latitudes (+/-50°), whereas the Level-3 data sets show mean differences with respect to the OMI reference data record that span between -0.2 ± 0.9 % (for GOME-2B) and 1.0 ± 1.4 % (for SCIAMACHY).
- For the Level-2 validation against ground-based measurements: the mean bias between GODFIT v4 satellite and Brewer, Dobson and SAOZ – reported TOCs is well within 1.5 ± 1.0 % for all sensors, the drift per decade spans between 0 % to 1.4 ± 1.0 % depending on the sensor and the peak-to-peak seasonality ranges between ~1 % for GOME and OMI, to ~2 % for SCIAMACHY.
- For the Level-3 validation against ground-based measurements shows ~~: an~~ remarkable ~~excellent~~ agreement with 0.5 to 1.5 % peak-to-peak amplitude for the monthly mean time series, ~~is found~~ as well as a negligible drift in the Northern Hemisphere with differences at -0.11 ± 0.10 % per decade for Dobson and +0.22 ± 0.08 % per decade for Brewer collocations.

We hence conclude that the ~~exceptional~~ quality of the GTO-ECV v3 Level-3 TOC record temporal stability satisfies well the requirements of 1 – 3 % per decade. The prominent~~excellent~~ inter-sensor consistency renders both the Level-2 GODFIT v4, as well as the Level-3 GTO-ECV v3 datasets, suitable and useful for longer term analysis of the ozone layer, such as decadal trend studies, the evaluation of model simulations, and data assimilation applications.

The Ozone_cci CRDP includes data products for total ozone columns, ozone profiles from nadir sensors and stratospheric ozone profiles from limb and occultation sensors. All data sets are reported in netCDF-CF format following CCI and GCOS standards, and are freely available on the Ozone_cci web site (http://www.esa-ozone-cci.org/?q=node/160).

Acknowledgements

The authors are grateful to ESA's Climate Change Initiative – Ozone project, Phase II, for providing the support and funding necessary for this work. The ground-based data used in this publication were obtained as part of WMO's Global Atmosphere Watch (GAW) and the Network for the Detection of Atmospheric Composition Change (NDACC). They are publicly available via the World Ozone and UV Data Centre (WOUDC) and the NDACC Data Host Facility (see http://woudc.org and http://ndacc.org, respectively). We would like to acknowledge and warmly thank all the investigators that provide data to these repositories on a timely basis, as well as the handlers of these databases for their upkeep and quality guaranteed efforts. We would also like to acknowledge the SBUV instrument team members for their work producing SBUV version 8.6 TOC data record.

References

Aliwell, S. R., M. Van Roozendael, P. V. Johnston, A. Richter, T. Wagner, D. W. Arlander, J. P. Burrows, D. J. Fish, R. L. Jones, K. K. Tørnkvist, J.-C. Lambert, K. Pfeilsticker, and I. Pundt; Analysis for BrO in zenith-sky spectra: An intercomparison exercise for analysis improvement. J. Geophys. Res., 107, 1–20. https://doi.org/10.1029/2001JD000329, 2002.

Antón, M., D. Loyola, M. López, et al., Comparison of GOME-2/MetOpA total ozone data with Brewer spectroradiometer data over the Iberian Peninsula, Annales Geophysicae, 27, 1377–1386, DOI:10.5194/angeo-27-1377-2009, 2009.

Bak, J., X. Liu, J. H. Kim, K. Chance, and D. P. Haffner (2015), Validation of OMI total ozone retrievals from the SAO ozone profile algorithm and three operational algorithms with Brewer measurements, Atmos. Chem. Phys., 15, 667–683, doi:10.5194/acp-15-667-2015.

Balis, D., et al., 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.

Balis, D., M. Kroon, M. E. Koukouli, et al., 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.

~~Basher, R. E., Review of the Dobson Spectrophotometer and Its Accuracy, Report No. 13 for. the WMO Global Ozone Research and Monitoring Project, Geneva, December 1982,~~ Bass, A. M. and Paur, R. J., The Ultraviolet Cross–sections of Ozone: I. The Measurements II. Results and Temperature Dependence, in Atmospheric Ozone. Proceedings of the Quadrennial Ozone Symposium, Halkidiki, Greece, edited by C. Zerefos and A. Ghazi, Reidel, Dordrecht, 606–616, 1985

Beirle, S., Lampel, J., Lerot, C., Sihler, H. and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10(2), 581–598, doi:10.5194/amt-10-581-2017, 2017.

Bernhard, G., Evans, R. D., Labow, G. J., and Oltmans, S. J., Bias in Dobson total ozone measurements at high latitudes due to approximations in calculations of ozone absorption coefficients and air mass, J.Geophys. Res, 110, D10305, doi:10.1029/2004JD005559, 2005.

Coldewey-Egbers, M., D.G. Loyola, M. Koukouli, D. Balis, J.-C. Lambert, T. Verhoelst, J. Granville, M. van Roozendael, C. Lerot, R. Spurr, S.M. Frith, and C. Zehner, The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative, *Atmos. Meas. Tech.*, **8**, 3923-3940, 2015.

Coldewey-Egbers, M., Weber, M., Lamsal, L. N., de Beek, R., Buchwitz, M. and Burrows, J. P.: Total ozone retrieval from GOME UV spectral data using the weighting function DOAS approach, Atmos. Chem. Phys., 5(4), 1015―1025, http://www.atmos-chem-phys.net/5/1015/2005/, 2005.

Danckaert T., Lerot, C., Van Roozendael, M., Spurr, R., Romahn, F., Zimmer, W., Heue, K.-P., Coldewey-Egbers, M., Loyola, D., Total ozone retrievals from TROPOMI/S5p using a direct-fitting approach: Algorithm Theoretical Basis, Atmos. Meas. Tech., in preparation, 2017.

Fioletov, V. E., et al., Performance of the ground-based total ozone network assessed using satellite data, J. Geophys. Res., 113, D14313, doi:10.1029/2008JD009809, 2008.

Fioletov, V. E., J. B. Kerr, E. W. Hare, G. J. Labow, and R. D. McPeters, An assessment of the world ground-based total ozone network performance from the comparison with satellite data, J. Geophys. Res., 104(D1), 1737–1747, doi:10.1029/1998JD100046, 1999.

Frith, S. M., N. A. Kramarova, R. S. Stolarski, R. D. McPeters, P. K. Bhartia, and G. J. Labow, Recent changes in total column ozone based on the SBUV Version 8.6 Merged Ozone Data Set, J. Geophys. Res. Atmos., 119, 9735–9751, doi:10.1002/2014JD021889, 2014.

Frith, S. M., Stolarski, R. S., Kramarova, N. A., and McPeters, R. D.: Estimating Uncertainties in the SBUV Version 8.6 Merged Profile Ozone Dataset, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-412, in review, 2017.

Hendrick, F., Pommereau, J.-P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M., and Van Roozendael, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, Atmos. Chem. Phys., 11, 5975-5995, https://doi.org/10.5194/acp-11-5975-2011, 2011.

Hubert et al.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) : 3. Ground-based validation of limb ozone profile data products, in preparation, 2017

Keppens, A., Lambert, J.-C., Granville, J., Hubert, D., Verhoelst, T., Compernolle, S., Latter, B., Kerridge, B., Siddans, R., Boynard, A., Hadji-Lazaro, J., Clerbaux, C., Wespes, C., Hurtmans, D. R., Coheur, P.-F., van Peet, J. C. A., van der A, R. J., Garane, K., Koukouli, M. E., Balis, D. S., Delcloo, A., Kivi, R., Stübi, R., Godin-Beekmann, S., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 2. Ground-based validation of nadir ozone profile data products, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-447, in review, 2018 Keppens et al.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) : 2. Ground-based validation of nadir ozone profile data products, in preparation, 2017.

Kerr, J. B., New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct sun spectra, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 107, NO. D23, 4731, doi:10.1029/2001JD001227, 2002

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), 11129–11140, doi: 10.1029/JD093iD09p11129, 1988

Koukouli, M. E., Zara, M., Lerot, C., Fragkos, K., Balis, D., van Roozendael, M., Allart, M. A. F., and van der A, R. J.: The impact of the ozone effective temperature on satellite validation using the Dobson spectrophotometer network, Atmos. Meas. Tech., 9, 2055-2065, https://doi.org/10.5194/amt-9-2055-2016, 2016.

Koukouli, M.E., Balis, D. S., Loyola, D., et al., 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, http://dx.doi.org/10.5194/amt-5-2169-2012, 2012.

Koukouli, M.E., D. S. Balis, I. Zyrichidou, et al., Evaluating a new homogeneous total ozone climate data record from GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, Journal of Geophysical Research, http://dx.doi.org/10.1002/2015JD023699, 2015.

Labow, G. J., R. D. McPeters, P. K. Bhartia, and N. Kramarova, A comparison of 40 years of SBUV measurements of column ozone with data from the Dobson/Brewer network, J. Geophys. Res. Atmos., 118, 7370–7378, doi:10.1002/jgrd.50503, 2013

Labow, G. J., Ziemke, J. R., McPeters, R. D., Haffner, D. P. and Bhartia, P. K.: A Total Ozone Dependent Ozone Profile Climatology based on Ozone-Sondes and Aura MLS Data, J. Geophys. Res. Atmos., n/a-n/a, doi:10.1002/2014JD022634, 2015.

Lambert, J.-C., D. Balis, A. Delcloo, K. Garane, F. Goutail, J. Granville, D. Hubert, A. Keppens, R. Kivi, ML. Koukouli, J.-P. Pommereau, R. Stübi, T. Verhoelst, D. Loyola, R. Siddans, R. van der A, C. Clerbaux, A. Laeng, V. Sofieva, M. Weber, M. Van Roozendael and C. Zehner, ESA Climate Change Initiative – Ozone_cci: Product Validation and Intercomparison Report (PVIR), Ozone_cci_Phase-II_PVIR_3.0, Issue 3.0 / Rev. 0, http://www.esa-ozone-cci.org/?q=documents#, January 2018.

Lerot, C., Van Roozendael, 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. Atmos., 119(3), 1639–1662, doi:10.1002/2013JD020831, 2014.

Loyola, D. G., M. E. Koukouli, P. Valks, et al., The GOME-2 Total Column Ozone Product: Retrieval Algorithm and Ground-Based Validation, J. Geophys. Res., doi:10.1029/2010JD014675, 2011.

Loyola, D.G., M. Coldewey-Egbers, M. Dameris, H. Garny, A. Stenke, M. van Roozendael, C. Lerot, D. Balis, and M. Koukouli, Global long-term monitoring of the ozone layer – a prerequisite for predictions, *Int. J. Remote Sens.*, **30**, 4295-4318, 2009.

McPeters, R. D., P. K. Bhartia, D. Haffner, G. J. Labow, and L. Flynn: The version 8.6 SBUV ozone data record: An overview, J. Geophys. Res. Atmos., 118, 8032–8039, doi:10.1002/jgrd.50597, 2013.

Orphal, J., Staehelin, J., Tamminen, J., Braathen, G., De Backer, M.-R., Bais, A., Balis, D., Barbe, A., Bhartia, P. K., Birk, M., Burkholder, J. B., Chance, K., von Clarmann, T., Cox, A., Degenstein, D., Evans, R., Flaud, J.-M., Flittner, D., Godin-Beekmann, S., Gorshelev, V., Gratien, A., Hare, E., Janssen, C., Kyrölä, E., McElroy, T., McPeters, R., Pastel, M., Petersen, M., Petropavlovskikh, I., Picquet-Varrault, B., Pitts, M., Labow, G., Rotger-Languereau, M., Leblanc, T., Lerot, C., Liu, X., Moussay, P., Redondas, A., Van Roozendael, M., Sander, S. P., Schneider, M., Serdyuchenko, A., Veefkind, P., Viallon, J., Viatte, C., Wagner, G., Weber, M., Wielgosz, R. I. and Zehner, C.: Absorption cross-sections of ozone in the ultraviolet and visible spectral regions: Status report 2015, J. Mol. Spectrosc., 327, 105–121, doi:10.1016/j.jms.2016.07.007, 2016.

Pommereau, J. & Goutail, F. O_3 and NO_2 ground-based measurements by visible spectrometry during Arctic winter and spring, Geophys. Res. Lett., doi:10.1029/GL015i008p00891, 1988.

Scarnato, B., J. Staehelin, T. Peter, J. Gröbner, and R. Stübi, Temperature and slant path effects in Dobson and Brewer total ozone measurements, J. Geophys. Res., 114, D24303, doi: 10.1029/2009JD012349, 2009.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W. and Burrows, J. P.: High spectral resolution ozone absorption cross-sections ~~&ndash~~ - ~~;~~ Part 2: Temperature dependence, Atmos. Meas. Tech., 7(2), 625–636, doi:10.5194/amt-7-625-2014, 2014.

Spurr, R., Natraj, V., Lerot, C., Van Roozendael, M. and Loyola, D.: Linearization of the Principal Component Analysis method for radiative transfer acceleration: Application to retrieval algorithms and sensitivity studies, J. Quant. Spectrosc. Radiat. Transf., 125, 1–17, 2013.

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 No. 1147, World Meteorol ogical Organization, Global Atmosphere Watch, No. 149, http://www.wmo.ch/web/arep/reports/gaw149.pdf, 2003

van der A, et al., User Requirement Document, Issue: 2.1 – Date of issue: 21/11/2011 Reference: Ozone_cci_URD_2.1, Ozone-CCI, 2011, http://www.esa-ozone-cci.org/?q=webfm_send/37, last accessed: 10.09.2017.

Van Roozendael, M., P. Peeters, H. K. Roscoe, et al., Validation of Ground-Based Visible Measurements of Total Ozone by Comparison with Dobson and Brewer Spectrophotometers, Journal of Atmospheric Chemistry, 29, 1, 55-83, 1998.

Vanicek, K., Differences between ground Dobson, Brewer and satellite TOMS-8, GOME-WFDOAS total ozone observations at Hradec Kralove, Czech, Atmos. Chem. Phys., 6, 5163-5171, doi:10.5194/acp-6-5163-2006, 2006.

Veefkind, J. P., Haan, J. F. de, Sneep, M. and Levelt, P. F.: Improvements to the OMI O2–O2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9(12), 6035–6049, doi:10.5194/AMT-9-6035-2016, 2016.

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.

Wang, P., Stammes, P., van der A, R., Pinardi, G. and van Roozendael, M.: FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8(21), 6565—6576,http://www.atmos-chem-phys.net/8/6565/2008/, 2008.

Ziemke, J. R., Chandra, S., Labow, G. J., Bhartia, P. K., Froidevaux, L., and Witte, J. C.: A global climatology of tropospheric and stratospheric ozone derived from Aura OMI and MLS measurements, Atmos. Chem. Phys., 11, 9237-9251, https://doi.org/10.5194/acp-11-9237-2011, 2011.

Quality assessment of the Ozone_CCI Climate Research Data Package (release 2017): 1. Ground-based validation of total ozone column data products

Katerina Garane[1], Christophe Lerot[2], Melanie Coldewey-Egbers[3], Tijl Verhoelst[2], Maria Elissavet Koukouli[1], Irene Zyrichidou[1], Dimitris S. Balis[1], Thomas Danckaert[2], Florence Goutail[4], Jose Granville[2], Daan Hubert[2], Arno Keppens[2], Jean-Christopher Lambert[2], Diego Loyola[3], Jean-Pierre Pommereau[4], Michel Van Roozendael[2] and Claus Zehner[5]

[1] Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece.
[2] Royal Belgian Institute for Space Aeronomy (BIRA-IASB), 3, Avenue Circulaire, B-1180 Brussels, Belgium.
[3] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik der Fernerkundung (IMF), 82234 Oberpfaffenhofen, Germany.
[4] LATMOS, CNRS, University Versailles St Quentin, Guyancourt, France
[5] European Space Agency, ESRIN, Frascati, Italy

Correspondence to: Katerina Garane (agarane@auth.gr)

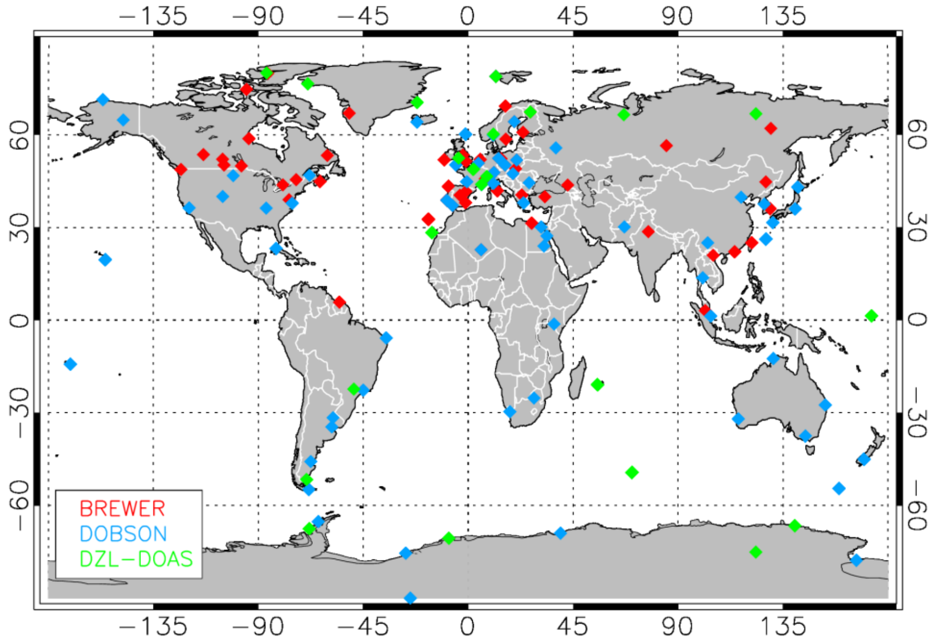


Figure S 1. The locations of the WOUDC/NDACC ground-based instruments reporting total ozone columns used in this study. For further information on these station, refer to Table S 1, Table S 2 and Table S 3.

Table S 1. The NDACC DOAS UV-Visible instruments selected for this study.

Station ID	Station name	Station Location	Latitude	Longitude	Elevation (m.a.s.l.)	Instrument	Start date	End date
315	Eureka	Canada	80.05°N	86.41°W	610	SAOZ	2004	2016
089	Ny Alesund	Spitsbergen	78.91°N	11.88°E	15	DOAS	1991	2012
460	Thule	Western Greenland	76.53°N	68.74°W	30	SAOZ	1991	2016
459	Scoresbysund	Eastern Greenland	70.48°N	21.95°W	68	SAOZ	1993	2016
262	Sodankylä	Finland	67.37°N	26.63°E	179	SAOZ	1990	2016
752	Zhigansk	Eastern Siberia	66.79°N	123.35°E	50	SAOZ	1994	2013
821	Salekhard	Western Siberia	66.50°N	66.70°E	137	SAOZ	2002	2009
658	Harestua	Norway	60.20°N	10.80°E	596	DOAS	1994	2013
601	Aberystwyth	Great Britain	52.45°N	4.07°W	50	SAOZ	1998	2017
049	Paris	France	48.85°N	2.35°E	65	SAOZ	2005	2016
	Guyancourt	France	48.78°N	2.05°E	165	SAOZ	2010	2016
041	Jungfraujoch	Switzerland	46.55°N	7.98°E	3580	SAOZ	1990	2012
040	Observatoire de Haute Provence	France	43.94°N	5.71°E	650	SAOZ	1993	2016
300	Izaña	Canaries Island	28.30° N	15.50° W	2367	DOAS	2000	2013
728	Tarawa	Kiribati	1.35°N	172.92°E	0	SAOZ	1993	1999
614	Bauru	Brazil	22.34°S	49.03°W	640	SAOZ	1996	2016
436	Reunion	Reunion Island	20.90°S	55.48°E	85	SAOZ	1994	2015
674	Kerguelen	Kerguelen Island	49.35°S	70.26°E	10	SAOZ	1996	2016
817	Rio Gallegos	Argentina	51.60°S	69.31°W	650	SAOZ	2008	2016
028	Dumont d'Urville	Antarctica	66.67°S	140.02°E	20	SAOZ	1988	2016
709	Rothera	Antarctic Peninsula	67.57°S	68.12°W	30	SAOZ	1996	2010
323	Neumayer	Antarctica	70.68°S	123.31 °E	42	DOAS	1999	2013
641	Dome Concordia	Antarctica	75.10°S	123.31 °E	3233	SAOZ	2007	2016

Table S 2. The WOUDC Dobson instruments selected for this study.

Station ID	Station Name	Station location	Latitude	Longitude	Elevation (m.a.s.l.)	Start date	End date
111	Amundsen Scott	Antarctica	-89.98	-24.8	2835	1995	2016
268	Arrival Heights	Antarctica	-77.83	166.4	250	1995	2016
57	Halley Bay	Antarctica	-75.52	-26.73	31	1995	2017
101	Syowa	Antarctica	-69	39.58	21	1995	2017
232	Vernadsky Faraday	Antarctica	-65.25	-64.27	7	1995	2017
339	Ushuaia	Argentina	-54.85	-68.31	7	1995	2016
29	Macquarie Island	Australia	-54.48	158.97	6	1995	2017
342	Comodoro Rivadavia	Argentina	-45.78	-67.5	43	1995	2016
256	Lauder	New Zealand	-45.03	169.68	3701	1995	2017
253	Melbourne	Australia	-37.48	144.58	125	1995	2017
91	Buenos Aires	Argentina	-34.58	-58.48	25	1995	2014
159	Perth	Australia	-31.95	115.85	2	1995	2016
343	Salto	Uruguay	-31.58	-57.95	31	1996	2013
340	Springbok	South Africa	-29.67	17.9	1	1995	2016
27	Brisbane	Australia	-27.47	153.03	5	1995	2017
265	Irene	South Africa	-25.25	28.22	1524	1995	2016
200	Cachoeira Paulista	Brazil	-22.68	-45	573	1995	2017
191	Samoa	USA	-14.25	-170.57	82	1995	2017

84	Darwin	Australia	-12.47	130.83	0	1995	2017
219	Natal	Brazil	-5.83	-35.2	32	1995	2017
175	Nairobi	Kenya	-1.27	36.8	1710	1995	2012
214	Singapore	Singapore	1.33	103.88	14	1995	2012
216	Bangkok	Thailand	13.73	100.57	2	1995	2017
31	Mauna_loa	USA	19.53	-155.58	3397	1995	2017
2	Tamanrasset	Algeria	22.8	5.52	1395	1995	2017
311	Havana	Cuba	23.17	-82.33	50	2005	2015
245	Aswan	Egypt	23.97	32.45	193	1995	2017
209	Kunming	China	25.02	102.68	1917	1995	2014
190	Naha	Japan	26.2	127.67	29	1995	2017
409	Hurghada	Egypt	27.25	33.72	22	2001	2017
10	New Delhi	India	28.63	77.22	216	1995	2015
152	Cairo	Egypt	30.08	31.28	35	1995	2017
11	Quetta	Pakistan	30.18	66.95	1799	1995	2013
7	Kagoshima	Japan	31.63	130.6	283	1995	2005
14	Tateno	Japan	36.05	140.13	31	1995	2017
106	Nashville	USA	36.25	-86.57	182	1995	2017
341	Hanford	USA	36.32	-119.63	73	1995	2017
213	El Arenosillo	Spain	37.1	-6.73	41	1995	2012
252	Seoul	Korea	37.57	126.95	84	1995	2013
107	Wallops Island	USA	37.87	-75.52	4	1995	2017
293	Athens	Greece	38	23.7	15	1995	2017
82	Lisbon	Portugal	38.77	-9.13	105	1995	2002
208	Shiangher	China	39.77	117	13	1995	2017
67	Boulder	USA	40.02	-105.25	1634	1995	2016
12	Sapporo	Japan	43.05	141.33	19	1995	2017
40	Haute Province	France	43.92	5.75	580	1995	2017
201	Sestola	Italy	44.22	10.77	1030	1995	2002
226	Bucharest	Romania	44.48	26.13	92	1995	2015
419	Bordeaux	FRA	44.81	-0.56	58	1995	2003
19	Bismarck	USA	46.77	-100.75	511	1995	2017
35	Arosa	Switzerland	46.77	9.67	1860	1995	2013
20	Caribou	USA	46.87	-68.02	192	1995	2017
100	Budapest	Hungary	47.43	19.18	140	1995	1998
99	Hohenpeissenberg	Germany	47.8	11.02	975	1995	2017
96	Hradec Kralove	Czech_Republic	50.18	15.83	285	1995	2017
36	Camborne	UK	50.22	-5.32	88	1995	2003
53	Uccle	Belgium	50.8	4.35	100	1995	2009
68	Belsk	Poland	51.83	20.78	180	1995	2017
50	Potsdam	Germany	52.38	13.05	89	1995	2003
116	Moscow	Russia	55.75	37.57	187	1995	2004
165	Oslo	Norway	59.92	10.72	50	1995	1998
43	Lerwick	UK	60.15	-1.15	90	1995	2017
51	Reykjavik	Iceland	64.13	-21.9	60	1995	2014
284	Vindeln	Sweden	64.25	19.77	0	1995	2017
105	Fairbanks	USA	64.8	-147.89	138	1995	2017
199	Barrow	USA	71.32	-156.6	11	1995	2017
89	Ny Alesund	Norway	78.93	11.88	0	1995	1997

Table S 3. The WOUDC Brewer instruments selected for this study.

Station ID	Station Name	Station location	Latitude	Longitude	Elevation (m.a.s.l.)	Start date	End date
322	Petaling Jaya	Malaysia	3.1	101.65	46	1999	2017
435	Paramaribo	Surinam	5.78	-55.2	5	1999	2016

330	Hanoi	Vietnam	21	105	0	2012	2017
468	Cape d'aguilar	Hong Kong	22.18	114.23	75	2003	2010
2	Tamanrasset	Algeria	22.8	5.52	1395	2011	2017
95	Taipei	Taiwan	25.03	121.52	22	2006	2013
376	Mrsa_mtrouh	Egypt	31.33	27.22	35	1998	2017
287	Funchal	Portugal	32.65	-17.05	59	1995	2002
332	Pohang	Korea	36.03	129.38	0	1995	2016
213	El Arenosillo	Spain	37.1	-6.73	41	2000	2012
346	Murcia	Spain	38	-1.17	69	1995	2017
82	Lisbon	Portugal	38.77	-9.13	105	2000	2002
447	Goddard	USA	38.99	-76.83	100	2000	2010
348	Ankara	Turkey	39.95	32.88	891	2006	2013
308	Madrid	Spain	40.45	-3.55	0	1995	2017
261	Thessaloniki	Greece	40.52	22.97	4	1995	2017
411	Zaragoza	Spain	41.66	-0.94	235	2000	2017
305	Rome University	Italy	41.9	12.52	0	1995	2015
405	La Coruna	Spain	43.33	-8.5	62	1999	2017
282	Kislovodsk	Russia	43.73	42.66	2070	1995	2016
65	Toronto	Canada	43.78	-79.47	198	1995	2014
326	Longfenshan	China	44.75	127.6	0	1995	2015
321	Halifax	Canada	44.9	-63.5	0	1995	2003
319	Montreal	Canada	45.47	-73.75	0	1995	2001
479	Aosta	Italy	45.71	7.33	585	2007	2017
301	Ispra	Italy	45.8	8.63	0	1995	2005
35	Arosa	Switzerland	46.77	9.67	1860	1995	2013
100	Budapest	Hungary	47.43	19.18	140	1999	2013
99	Hohenpeissenberg	Germany	47.8	11.02	975	1995	2017
290	Saturna	Canada	48.78	-123.13	0	1995	2016
331	Poprad Ganovce	Slovakia	49.03	20.32	0	1995	2017
320	Winnipeg	Canada	49.91	-97.24	0	1995	2002
96	Hradec Kralove	Czech Republic	50.18	15.83	285	1995	2017
338	Regina	Canada	50.21	-104.67	0	1995	2005
53	Uccle	Belgium	50.8	4.35	100	1995	2017
353	Reading	UK	51.42	-0.96	51	2002	2017
68	Belsk	Poland	51.83	20.78	180	1995	2005
318	Valentia	Ireland	51.93	-10.25	0	1995	2017
316	Debilt	Netherlands	52	5.18	0	1995	2017
241	Saskatoon	Canada	52.1	-105.28	550	1995	2001
174	Lindenberg	Germany	52.22	14.12	98	1995	2014
50	Potsdam	Germany	52.38	13.05	89	1995	2003
76	Goose	Canada	53.32	-60.38	44	1995	2016
352	Manchester	UK	53.45	-2.26	61	2000	2017
21	Edmonton	Canada	53.57	-113.52	668	1995	2016
481	Tomsk	Russai	56.48	84.97	170	2003	2012
279	Norkoping	Sweden	58.58	16.12	0	1995	2017
77	Churchill	Canada	58.75	-94.07	35	1995	2016
404	Jokioinen	Finland	60.8	23.5	103	1999	2001
123	Yakutsk	Russia	62.08	129.75	98	1995	2005
284	Vindeln	Sweden	64.25	19.77	0	1996	2017
267	Sondrestrom	Greenland	67	-50.98	150	1995	2017
262	Sodankyla	Finland	67.37	26.65	179	1995	2010
476	Andoya	Norway	69.25	15.97	395	2000	2016
24	Resolute	Canada	74.72	-94.98	64	1995	2016
89	Ny Alesund	Norway	78.93	11.88	0	2007	2009
315	Eureka	Canada	79.89	-85.93	10	2001	2016