JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, D07307, doi:10.1029/2005JD006376, 2007

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

D. Balis,[1] J.-C. Lambert,[2] M. Van Roozendael,[2] R. Spurr,[3] D. Loyola,[4] Y. Livschitz,[4]
P. Valks,[4] V. Amiridis,[1] P. Gerard,[2] J. Granville,[2] and C. Zehner[5]

Received 15 June 2005; revised 24 February 2006; accepted 24 August 2006; published 7 April 2007.

[1] The atmospheric chemistry instrument Global Ozone Monitoring Experiment (GOME) was launched in April 1995 on the ERS-2 platform. The GOME data processor (GDP) operational retrieval algorithm has produced total ozone columns since July 1995. With a data record of over ten years, GOME has become important for ozone trend analysis. In 2004, GDP was upgraded to version 4.0, a new validation was performed, and the entire GOME data record was reprocessed. In the preceding paper (Van Roozendael et al., 2006), the GDP 4.0 algorithm was described. In this paper, we deal with geophysical validation of the GDP 4.0 algorithm and the retrieved ozone products. We present results of a validation exercise involving comparisons of GDP 4.0 total ozone with the Network for Detection of Stratospheric Change (NDSC) and the World Meteorological Organization (WMO)/Global Atmospheric Watch (GAW) ground-based networks. We compare these results with similar validations of earlier GDP ozone products. We also present ground-based validation of TOMS versions 7 and 8 total ozone products, and we contrast these with GDP 4.0 values. On a global basis, GDP 4.0 total ozone results lie between −1% and +1.5% of ground-based values for solar zenith angles less than 70°; accuracy is now comparable to that obtainable from ground-based stations. At higher solar zenith angles in polar regions, larger discrepancies of up to +5% are found; in these regimes, errors on both satellite and ground-based measurements are higher. The validation also showed marked improvement in TOMS total ozone performance for the version 8 algorithm.

Citation: Balis, D., et al. (2007), 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.

1. Introduction

1.1. GOME Operation and Data Processor

[2] The Global Ozone Monitoring Experiment (GOME) is one of several instruments on board the European Remote Sensing Satellite-2 (ERS-2), launched in April 1995. GOME is the successful predecessor of a number of new generation atmospheric chemistry sensors; these include Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) [*Bovensmann et al.*, 1999], launched in 2002 as part of the Environmental Satellite (ENVISAT) payload, the Ozone Monitoring Instrument (OMI) on board the NASA's Aura platform launched in July 2004 [*Stammes et al.*, 1999], and the GOME-2 series of instruments, the first of which was launched in October 2006 on board the Meteorological Operational (METOP) platform. The primary aim of these atmospheric chemistry instruments is to provide global measurements of key ozone-related species for the assessment of current and future changes in atmospheric composition and chemistry atmosphere [*Burrows et al.*, 1999]. GOME provides a global picture of atmospheric ozone (O_3); it is also the first orbiting instrument with the ability to measure the vertical column amount of nitrogen dioxide (NO_2).

[3] GOME on ERS-2 is an across-track nadir-viewing spectrometer with four linear array detectors covering the spectral range 240–793 nm and with resolutions from 0.2 to 0.4 nm. The satellite has a sun-synchronous polar orbit at height ~790 km, and the instrument swath is 960 km, with three forward scans (footprint 320×40 km^2) in nominal viewing mode. For the correction of polarization effects and for use in cloud fraction determination, GOME also has three broadband detectors (PMDs or Polarization Measurement Devices) measuring at subpixel read-out times. For more details on the GOME instrument, see the GOME Users Manual [*ESA*, 1995]. Since August 1996, GOME total O_3 and NO_2 column data have been processed with the

---

[1] Aristotle University of Thessaloniki, Thessaloniki, Greece.
[2] Belgian Institute for Space Aeronomy, Brussels, Belgium.
[3] RT Solutions, Incorporated, Cambridge, Massachusetts, USA.
[4] German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Wessling, Germany.
[5] European Space Agency Centre for Earth Observation (ESA/ESRIN), Frascati, Italy.

0148-0227/07/2005JD006376

GOME data processor [*Loyola et al.*, 1997] at the German Processing and Archiving Facility (D-PAF) established at the German Aerospace Centre (DLR) on behalf of the European Space Agency (ESA).

[4] GOME has now been producing global distributions of total ozone for ten years. The length and the long-term stability of this data record make it desirable for use in long-term ozone trend monitoring. Following the ESA's call in summer 2002 for improved GOME total ozone algorithms to meet trend analysis accuracy and stability requirements, three new Differential Optical Absorption Spectroscopy (DOAS)–based algorithms were developed to reprocess the GOME total ozone record. These algorithms are (1) the WFDOAS algorithm [*Coldewey-Egbers et al.*, 2005; *Weber et al.*, 2005]; (2) the TOGOMI/TOSOMI algorithm [*Eskes et al.*, 2005]; and (3) the GDOAS algorithm (this work). GDOAS was implemented operationally in the D-PAF at DLR. This GDOAS implementation has become GDP 4.0 (version 4.0 of the GOME data processor). A description of the GDP 4.0 GDOAS algorithm was given in the preceding paper [*Van Roozendael et al.*, 2006].

[5] The GOME data processing environment at D-PAF was replaced recently by a fully redesigned system called the "Universal Processor for UV/visible Atmospheric Spectrometers" (UPAS). The greater flexibility of the UPAS environment has facilitated the implementation, verification and validation of the new GDP 4.0 operational total column retrieval algorithm. The complete reprocessing of the entire GOME total ozone record was finished in December 2004, and the reprocessed GOME data record, including historical data, is already available to the public via the ERS Help and Order Desk.

1.2. Summary of the GDP 4.0 Algorithm

[6] In this section, we summarize the GDP 4.0 algorithm; more details are given by *Van Roozendael et al.* [2006]. We shall also refer to the predecessor GDP 3.0 algorithm for operational total column products; a description and validation summary may be found in [*Spurr et al.*, 2005]. GDP 4.0 is a DOAS retrieval algorithm, comprising a least squares fitting for the slant column of ozone followed by an Air Mass Factor (AMF) computation to derive the corresponding vertical column amount.

[7] In addition to the ozone slant column, the DOAS part of the algorithm also retrieves an effective temperature to characterize the Huggins-bands ozone absorption temperature dependence, and two fitting amplitudes corresponding to additive reference spectra for undersampling and Fraunhofer filling-in (Ring) effect [*Van Roozendael et al.*, 2006, section 2.2]. There is also an explicit shift and squeeze fitting to improve the wavelength registration of the earthshine radiance spectrum, plus an additional cross-correlation for the Level 1b irradiance spectrum in the region of the DOAS fitting window (325–335 nm).

[8] GDP 4.0 uses an iterative computation for the AMF and vertical column density VCD, with the radiative transfer calculations based on the use of a column classified ozone profile climatology. This method was first used in GDP 3.0, where neural network functions were employed to deliver the AMFs at 325.0 nm. In GDP 4.0, AMFs are calculated directly with calls to a radiative transfer model. Also in GDP 4.0, the improved TOMS version 8 ozone profile climatology (P. K. Bhartia and C. Wellemeyer, TOMS version 8 Algorithm Theoretical Basis Document, available at http://toms.gsfc.nasa.gov, 2004-11-24, hereinafter referred to as Bhartia and Wellemeyer, 2004) was used in place of the version 7 data set for GDP 3.0, and the AMFs were calculated at 325.5 nm (the use of this wavelength has improved total ozone accuracy [see *Van Roozendael et al.*, 2006, Figure 5]).

[9] Perhaps the most important change from GDP 3.0 to GDP 4.0 has been the introduction of a completely new algorithm component to deal with the telluric filling-in interference effects in ozone absorption signatures caused by inelastic rotational Raman scattering (RRS). This "molecular Ring correction" generates a scaling on the effective slant column at each AMF/VCD iteration. It has resulted in a dramatic improvement in accuracy [*Van Roozendael et al.*, 2006, section 4, Figure 6]. New parameterizations of the molecular Ring effect contribution were also incorporated in the other two algorithms used for GOME total ozone reprocessing [*Coldewey-Egbers et al.*, 2005; *Eskes et al.*, 2005].

[10] Last, a new cloud preprocessing step has been introduced in GDP 4.0 for the treatment of clouds in the independent pixel approximation (IPA). This replaces the GDP 3.0 system, which relied on databased selections of cloud top albedo and height, and only retrieved the cloud fraction. In GDP 4.0 there are two new algorithms; the first (OCRA) uses data fusion techniques to derive the cloud fraction from the subpixel PMD measurements, while the second (ROCINN) derives the required cloud top information from the spectral fitting of reflectivity in and around the Oxygen A band. For details, see *Van Roozendael et al.* [2006, section 5].

[11] Improvements in total ozone accuracy were discussed by *Van Roozendael et al.* [2006] for a limited number of orbits, and a number of sensitivity tests were carried out [*Van Roozendael et al.*, 2006, section 6]. An important requirement for trend analysis is the lack of drift in the ozone record, and this has been a remarkable feature of the GOME record. This point was given by *Van Roozendael et al.* [2006] (see, e.g., Figure 11). In the present paper we will present a thorough documentation of the improvement in the accuracy of the new GDP 4.0 total ozone product. This validation will be based on detailed comparisons with ground-based networks and with results from the TOMS version 7 and 8 algorithms for total ozone.

1.3. Introduction to the GDP 4.0 Delta Validation

[12] Prior to the implementation of any major change in the operational GDP processing chain, it is essential to verify the accuracy and effectiveness of new modifications and to assess the quality of the new data product. Such "Delta Validations" of putative product improvements have been executed after every major GDP upgrade by a subgroup of the GOME Validation Group responsible for the investigation of GOME data product quality throughout the mission lifetime. In the context of the present GDP upgrade to version 4.0, a delta validation campaign was set up in 2004 with the main emphasis on the quality assessment of new ozone column amounts on the global scale and in the long term. The nitrogen dioxide (NO_2) column product has also been reprocessed. The GDP 4.0 validation

effort involved a team of scientists from the Belgian Institute for Space Aeronomy (IASB-BIRA, Brussels, Belgium), the Laboratory of Atmospheric Physics (AUTH, Aristotle University of Thessaloniki, Greece), and DLR's Remote Sensing Technology Institute (DLR-IMF, Oberpfaffenhofen, Germany).

[13] Ground-based data have been compared with various versions of GOME and TOMS satellite data: GDP 2.7, GDP 3.0, GDP 4.0, and TOMS versions 7 and 8 provided by NASA/GSFC. In the ground-based validation for GDP 2.7 total ozone [*Lambert et al.*, 1999, 2000], there were a number of marked dependencies in the total column differences on the season, the solar zenith angle, the ozone column value itself, and the latitude. Validation of the GDP 3.0 total ozone product revealed that the ozone column dependence had nearly disappeared and the amplitudes of the other dependencies were reduced by about 50% on average [*Lambert et al.*, 2002; *Spurr et al.*, 2005]. With TOMS V7 ozone product validation, seasonally varying errors showed lower amplitude, but a systematic offset affected the whole Southern Hemisphere, increasing with latitude from a few percent in the Tropics to about 10% in Antarctica [*Lambert et al.*, 2000]. Ozone column dependence was also found in TOMS V7. Effects of instrumental degradation were noted with both TOMS V7 and GDP 2.7, but these were corrected with GDP 3.0.

[14] This paper describes the main results for total ozone column validation from the GDP 4.0 Delta Validation Campaign in 2004. The new GDP 4.0 total ozone product is characterized first by comparison with correlative measurements from extensive ground-based networks archived in the World Ozone and UV Data Center (WOUDC) operated at Toronto (Canada), and in the database of the Network for the Detection of Stratospheric Change (NDSC) operated at NOAA (National Oceanic and Atmospheric Administration) and mirrored at NILU (Norwegian Institute for Air Research). Secondly, a similar characterization is performed for the Earth Probe (EP) TOMS total ozone overpass products as available from NASA Goddard Space Flight Center (GSFC), and the results compared with those from GDP 4.0. Although the choice of GOME validation orbits is oriented toward validating the ten-year ozone record from GDP 4.0, the advent of a recent major upgrade of the TOMS BUV total ozone algorithm gives us a unique opportunity to extend the GDP 4.0 validation with a series of comparisons against TOMS data.

[15] In chapter 2, we discuss the reference data sets used for the GDP 4.0 delta validation, the selection procedure for ground-based stations (section 2.1), the GOME orbits chosen for the exercise (section 2.2) and a summary of the TOMS measurements selected for the intercomparisons (section 2.3). In section 3 we deal with uncertainties in ground-based data that might impact the validation of satellite ozone columns. Section 4 contains the validation results. In section 4.1, we present results from test case studies at an NDSC Alpine station and use them to illustrate how validation at the "percent level" is made feasible owing to the integrated exploitation of complementary measurement types (Brewer and Dobson). We demonstrate the large-scale capability of ground-based networks to investigate major characteristics of the GDP 4.0 and TOMS V8 total ozone products in terms of global agreement (section 4.2), seasonal dependency (section 4.3) and long-term stability (section 4.4). The remaining part of section 4 focuses on the major dependencies due to solar zenith angle (section 4.5), total ozone column, especially in ozone hole scenarios (section 4.6), and fractional cloud cover (section 4.7).

[16] Much more information is available in the GDP 4.0 Validation Report produced in December 2004 [*Lambert and Balis*, 2004]. This report may be downloaded from the Web page http://wdc.dlr.de/sensors/gome/gdp4.html (under "GDP 4.0 Validation"), which accompanies this paper and its predecessor [*Van Roozendael et al.*, 2006]. Older validation documentation may also be downloaded from this site.

2. Selection of Validation Data Sets

2.1. Ground-Based Correlative Data Sets

[17] The present study is based on archived total ozone measurements provided by two major contributors to World Meteorological Organization (WMO)/Global Atmospheric Watch (GAW): Dobson and Brewer total ozone data records, as deposited at the WOUDC in Toronto, Canada (http://www.woudc.org); and UV-visible DOAS, Dobson and Brewer total ozone data records acquired as part of the NDSC (public archive available via http://www.ndsc.ws). The WOUDC contains total ozone data mainly from Dobson and Brewer UV spectrophotometers and from M-124 UV filter radiometers. The NDSC database contains total ozone data from DOAS UV-visible and Fourier Transform infrared spectrometers, plus Dobson and Brewer instruments at selected stations that routinely upload their preliminary data to WOUDC. In general, coincidences offered by the Dobson, Brewer and UV-visible networks with the list of validation orbits are sufficient to investigate the effects of the current GDP upgrade. For this and other reasons (sparser records, fewer coincidences, some biases), ground-based measurements performed with FTIR spectrometers and with M-124 radiometers have not been considered in the GDP 4.0 total ozone validation. Total ozone data from a large number of the WOUDC and NDSC stations have already been used extensively both for trend studies [e.g., *WMO*, 1998, 2003] as well as for validation of satellite total ozone data [e.g., *Lambert et al.*, 1999; *Fioletov et al.*, 1999; *Lambert et al.*, 2000; *Bramstedt et al.*, 2003; *Labow et al.*, 2004].

[18] To prepare ground-based data sets for GOME and TOMS validation, we investigated the quality of the total ozone data of each station and instrument that deposited data at NDSC and WOUDC for any periods during 1995–2004. We examined for each station a series of plots and statistics based on GOME and TOMS comparisons. A sample of these plots is shown in Figure 1 for the Dobson station of Potsdam in Germany. Comparisons using direct sun ultraviolet measurements, which offer a greater accuracy (~1%) than those based on zenith-sky data, are performed separately (Figure 1, right). For each station we checked time series of the percent relative differences with the satellite data on a daily basis (first row) and monthly mean basis (second and third rows), the distribution and scatter of these differences (fourth and fifth rows), and finally the dependence of the difference on the solar zenith angle of the satellite observation for each season of the year

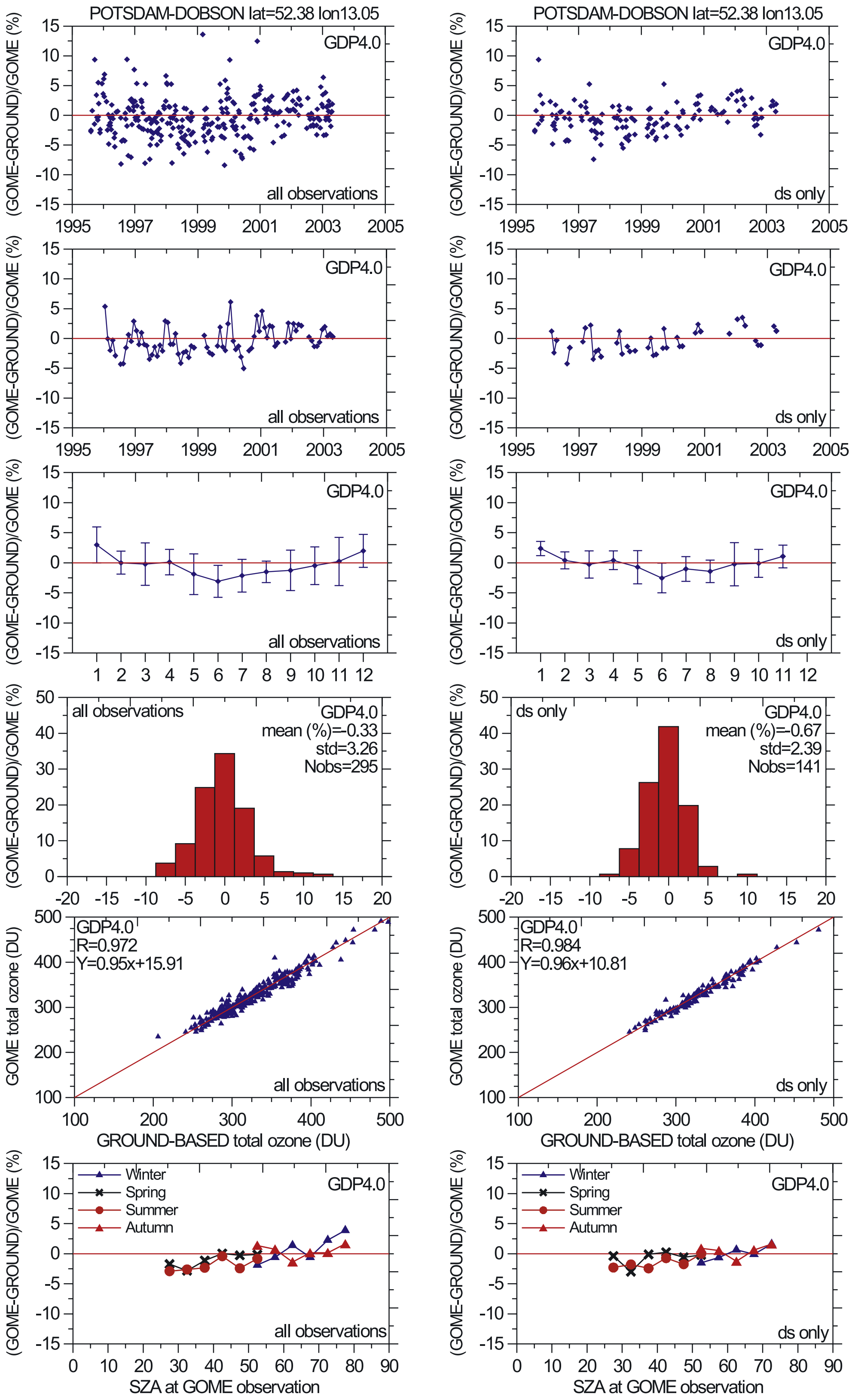


Figure 1. Sample of individual station statistics and plots carried out for all the stations and satellite data set examined (see section 2.1 for details). (left) All direct sun and zenith-sky Dobson data included. (right) direct sun data only.

(last row). In some cases, there were systematic inconsistencies between comparisons from neighboring stations. In certain cases, small correlation coefficients were found between the ground-based and satellite data.

[19] For the NDSC, there is a strict protocol on data verification, resulting in a 2-year delay between data acquisition and the upload of consolidated data to the central archive. For the longer-term validation considered in this paper, it was necessary to augment NDSC-endorsed data records with unconsolidated data reported in near real time. For the latter, an additional verification is based on climatological grounds. At ground stations where sufficiently long time series are available, the verification procedure consists in comparing fresh data to climatological means and standard deviations that were calculated from low-pass filtered time series acquired since 1995. The procedure singles out and records exceptions. First to be identified are the aberrant cases (e.g., Dobson data erroneously referred to as polar night data). Second, ozone column values deviating from the climatological mean by more than 2 or 3 standard deviations (2σ or 3σ) are singled out. Groups of consecutive values lying outside the $\pm 3\sigma$ interval are examined carefully to determine whether such persistent deviations are due to data quality issues, to natural atmospheric variability, or to unexpected atmospheric features such as the 2002 Antarctic vortex split. Single values outside the $\pm 3\sigma$ interval that do not belong to a justifiable 2σ train are flagged accordingly but are not rejected systematically since they could reflect natural atmospheric variability (e.g., high frequencies near the polar vortex boundary) or indicate special events such as tropospheric pollution episodes.

[20] From the above selection procedure, some 41 Brewer, 61 Dobson and 27 UV-visible DOAS instruments were considered for the comparisons with GDP 3.0, GDP 4.0, and TOMS V8. A complete listing of all instruments used in the validation may be found in the GDP 4.0 Delta-Validation Report [*Lambert and Balis*, 2004], wherein the Brewer, Dobson, and DOAS instruments are catalogued respectively in Tables 1, 2 and 3 in the report.

2.2. Selection of GOME Orbits

[21] In 2002, a list of 2257 validation orbits was compiled for the delta validation of the GDP upgrade to version 3.0 [*Lambert et al.*, 2002]. With a longer data record and additional validation required to test algorithmic upgrades in GDP 4.0, the selection will be larger. The selection is based on available ground-based data for the period 1995–2004, and it allows for delta-validation of total ozone from GOME from 1995–2004, of EP-TOMS total ozone from 1996–2004 and for the verification of GOME NO_2 column data from 1995–2004. The selection is large enough to characterize changes in cyclic errors, i.e., dependencies on the season, the latitude, the ozone column value and the solar zenith angle. Finally, results using the older GDP 3.0 algorithm are consistent between the new set of validation orbits and the previous set of 2257 orbits used in 2002.

[22] The current selection of about five thousand orbits based on histograms of GOME/ground comparisons at a list of 40 stations from pole to pole. Among these stations, there are 30 WMO/GAW stations equipped with Dobson and/or Brewer spectrophotometers, including the Canadian and NOAA/CMDL subnetworks and a few NDSC sites, and 20 NDSC sites operating UV/visible DOAS spectrometers. Orbits have been selected on the basis of proximity to the median value of the relative difference in total ozone. The selection has been constrained in such a way that the sampling of the column range and of its cyclic variations (with season, latitude and SZA) complies with both the Nyquist and Central Limit theorems [*Papoulis and Pillai*, 2002]. Another constraint is that there must be sufficient sampling of seasonal and meridian variations of the DOAS-derived effective temperature.

[23] Coincidences between GOME and ground-based observations are said to occur when any part of the GOME footprint covers the effective ground-based air mass. For Dobsons and Brewers, this is a maximum of 150 km between the footprint center and the stations, and within a temporal window of three hours. For UV-visible zenith-sky observations, the GOME footprint must intercept the air mass estimated with a ray tracing model (this is discussed further in section 3.1); observations must take place the same day (zenith-sky data are always taken at twilight).

[24] In Figure 2 we show time series of coincidences obtained at four stations from the Arctic, northern middle latitudes, the equator, and the Antarctic. This figure illustrates the degree to which selected coincidences capture geophysical variations of the ozone column (Figure 2, left), as well as major GDP 3.0 features (Figure 2, right) such as GDP 3.0 errors varying with the season. The set of selected orbits also offers adequate sampling of the Antarctic springtime, when the extreme ozone column range makes it easier to detect changes in ozone column dependence in the product. Sampling at single stations is not always constant with time. Selected orbits produce a total average value of about 3000 coincidences by station.

2.3. TOMS Measurements Used in the GDP 4.0 Validation

[25] The TOMS instruments provide measurements of Earth's ozone total column by measuring the backscattered Earth radiance at a set of discrete 1-nm wavelength bands in the region 310–380 nm. The experiments use a single monochromator and scanning mirror to sample backscattered ultraviolet (BUV) radiation, with periodic solar measurements to provide radiance normalization, and remove some instrumental dependence. The TOMS scanning mechanism provides (except for Earth Probe) equatorial interorbit overlap so that the entire sunlit portion of the globe is sampled daily. The sun-synchronous near-polar orbits (except for Meteor-3) provide these measurements at the same approximate local time. The new version 8 (V8) total ozone algorithm (Bhartia and Wellemeyer, 2004) has corrected several errors that were discovered in the predecessor version 7 (V7) algorithm [*Labow et al.*, 2004]. TOMS V8 uses only two wavelengths (317.5 and 331.2 nm) to derive total ozone, while the other 4 wavelengths (depending on the instrument) are used for diagnostic and error corrections. Upgrades have been made to the set of a priori ozone profiles, and to the tropospheric ozone climatology, and the algorithm now includes fixed ozone temperature dependency corrections based on month latitude temperature fields derived from numerical weather analysis models. For the GDP 4.0 validation study, we used EP-TOMS data,

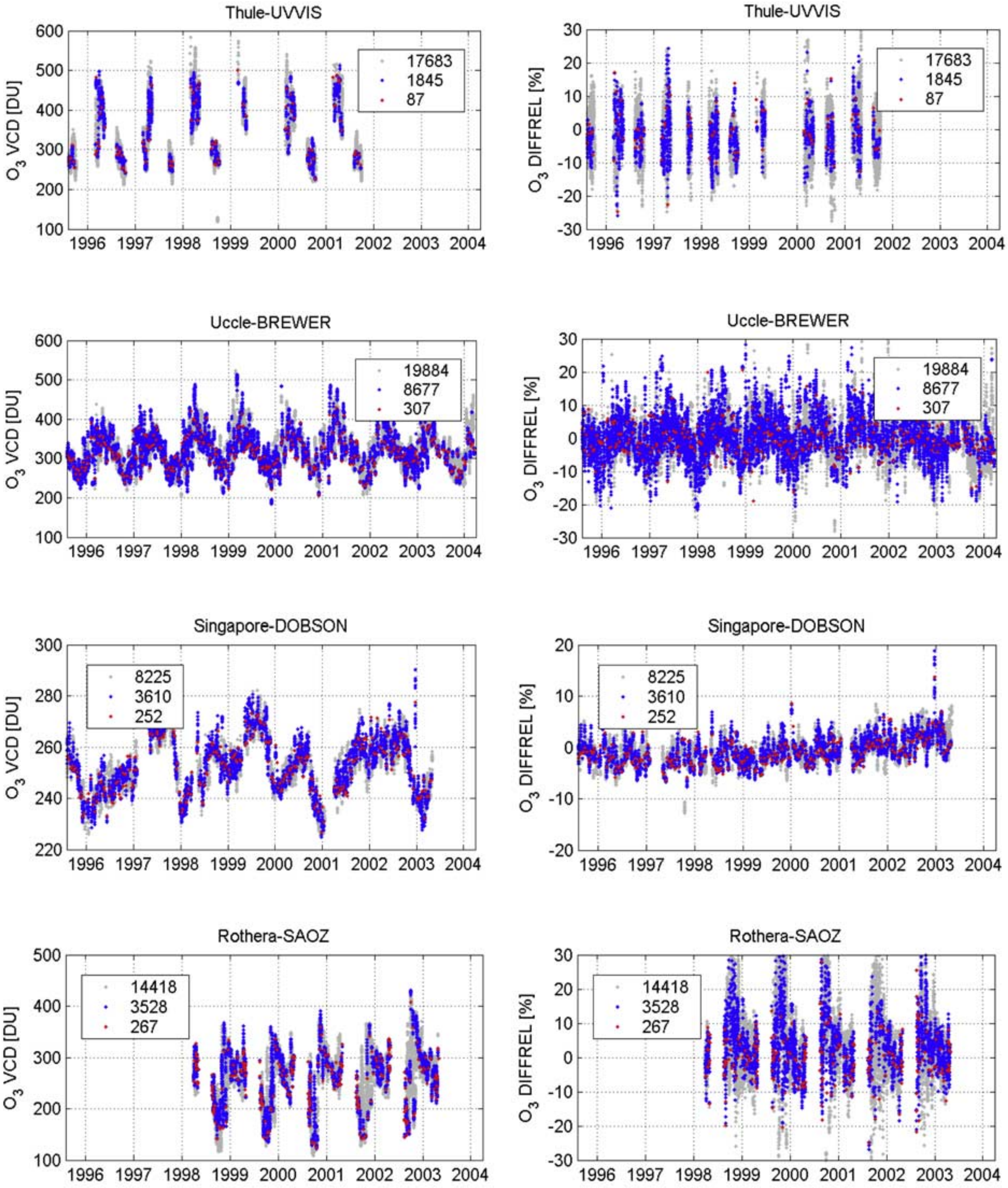


Figure 2. Time series of selected coincidences between GOME and ground-based ozone column observations at (top to bottom) the Arctic station of Thule (77°N, Danish Meteorological Institute SAOZ), the northern midlatitude station of Uccle (50°N, Royal Meteorological Institute of Belgium Brewer), the equatorial station of Singapore (1°N, Singapore Meteorological Office Dobson), and the Antarctic station of Rothera (68°S, British Antarctic Survey SAOZ). (left) Coincident ozone column measurements; (right) percent relative difference of GDP 3 versus ground-based ozone column data. Note that the vertical scale differs from one station to another. Gray dots show all coincidences between the available GOME and ground-based data records; the subset of coincidences plotted as red dots satisfies minimal sampling requirements (Nyquist, Shannon, etc.) for the known errors of GDP 3.0; blue dots show actual coincidences with the selected set of delta validation orbits. Total amount of coincidences for each subset of orbits is indicated in the legend.

processed at NASA/GSFC for the period 1996–2003. EP-TOMS time-dependent calibration is maintained (by using three on-board diffuser plates) to high accuracy by analyzing the degradation of the cover diffuser relative to the working and reference diffusers. The 150 km overpass criterion for GOME is also applied to the selection of TOMS data. As overpass data files provided by NASA/GSFC report only one value a day, corresponding to the closest ground pixel to the station, the total amount of coincidences by station is closer to 1000 for TOMS data, as compared with an average of 3000 for GOME data - still sufficient for reliable statistics (for details, see *Lambert and Balis* [2004]). In 2001, a bias in measurements made on one side of the orbital track relative to measurements from the opposite side of the scan was discovered in EP-TOMS data. A correction was implemented to stabilise the calibration relative to NOAA-16 SBUV/2 in the equatorial zone (Bhartia and Wellemeyer, 2004). However, there is still a latitude-dependent error that cannot be corrected by a simple calibration correction; at 50 degrees latitude, there are now −2% to −4% errors in TOMS. Because of this problem, NASA/GSFC recommends that data from 2002 should not be used for trend analysis. However, this validation exercise is a good opportunity to assess the performance of the TOMS corrections with respect to ground-based network data.

3. Error Sources in Ground-Based Data Sets

3.1. Ground-Based Errors and Mutual Consistency

[26] The present GDP 4.0 validation relies on the synergistic use of complementary ground-based network data contributing to the WMO GAW program. More than 300 Dobson [*Dobson*, 1957] and Brewer [*Kerr et al.*, 1983] stations form the world's primary total ozone monitoring network archived at WOUDC. Dobson and Brewer spectrophotometers derive total ozone from the ratio of direct sunlight intensities measured at standard wavelengths in the Huggins band. In the framework of the NDSC, there are also some 40 UV-visible DOAS spectrometers that monitor total ozone at twilight (this includes the automated network of 14 Système d'Analyse par Observation Zénithale (SAOZ) spectrometers [*Pommereau and Goutail*, 1988] measuring in the Chappuis bands). Here, ozone is derived from measurement of UV-visible sunlight scattered at the zenith of the instrument. Established procedures for maintaining high quality with Brewer and Dobson instruments are described in detail by [*Staehelin et al.*, 2003]. NDSC certification requires participation for UV-visible spectrometers in blind intercomparison campaigns [see, e.g., *Roscoe et al.*, 1999]. Ground-based stations delivering ozone column data regularly to WOUDC and/or NDSC are expected to provide well-controlled, long-term time series of data. However, there are several sources of uncertainty in the data that are known to interfere with GDP validation, and we now summarize these error sources.

[27] The work of [*Van Roozendael et al.*, 1998] has shown that mutual agreements between Dobson, Brewer and UV-visible data can reach the "percent" level when the major sources of discrepancy are properly accounted for. We summarize the main findings from this work. For Dobson instruments, the temperature dependence of the ozone absorption coefficients used in the retrievals might account for a seasonal variation in the error of ±0.9% in the Alps and ±1.7% at Sodankylä (Finland, 67.4°N), and for a systematic errors of up to 4% [*Bernhard et al.*, 2005]. The effect can dramatically increase in extremely cold conditions (winter polar vortex). The effect is smaller for Brewer instruments because of their use of wavelengths with reduced ozone absorption temperature dependency. Dobson and Brewer instruments might also suffer from long-term drift associated with calibration changes. Additional problems arise at solar elevations lower than 15°, for which diffuse and direct radiation contributions can be of the same order of magnitude. Assuming that the Dobson and Brewer instruments are well-calibrated, and that data are filtered to avoid air mass dependence, then the data can be corrected for temperature dependence using formulas given by [*Komhyr et al.*, 1993].

[28] For zenith-sky UV-visible spectrometers measuring ozone in the Chappuis band, the use of DOAS minimizes uncertainties associated with the temperature dependence of the absorption cross sections [*Burkholder and Talukdar*, 1994]. Major uncertainties on the vertical column are associated with the AMF conversion from spectrally derived slant columns to vertical columns. Real-time SAOZ data are based on the a single-profile standard AMF calculated at 60°N in winter and at sea level [*Sarkissian et al.*, 1995]. However, it is essential to calculate air mass factors at the correct station altitude, as otherwise offsets of up to 5% can appear in results from high mountain stations. Seasonal changes of the ozone profile and scattering geometry are responsible for a systematic bias of about 5–6% amplitude at 67°N, falling to 3–4% at 44°N. This is in addition to the ±1% scatter that might result from short-term fluctuations. The use of the standard SAOZ AMF also introduces an average meridian dependence of −3% at 67°N to +2.8% at the tropics.

[29] Accounting for these Dobson, Brewer and SAOZ effects, apparent seasonalities between Dobson/Brewer and SAOZ ozone column measurements vanish and the mutual consistency can reach the "1% level" on average. However, we do not include these corrections in the standard NDSC and WOUDC ozone data products as used in this GDP 4.0 validation; to do this would require a tremendous amount work that is beyond the scope of this paper. Consequently, ground-based uncertainties (and the same is true for smoothing errors described next) can be larger than differences between improved satellite ozone column data (from GOME GDP 4.0 and TOMS V8) and the ground-based results. Nevertheless, ground-based validation of satellites at the "1% level" is feasible provided that the interpretation includes an accounting of significant ground-based uncertainties. This will be done in sections 4 and 5 when discussing comparison results.

3.2. Smoothing Errors

[30] Validation comparisons require a sufficient degree of collocation of the air masses probed by the different sensors. In the GOME and TOMS ozone validation literature, collocation criteria are based on spatial windowing varying from 100 to 600 km and temporal windows from 1 hour to 1 day. The underlying assumptions in this procedure are that all collocated instruments are reporting the integral of the

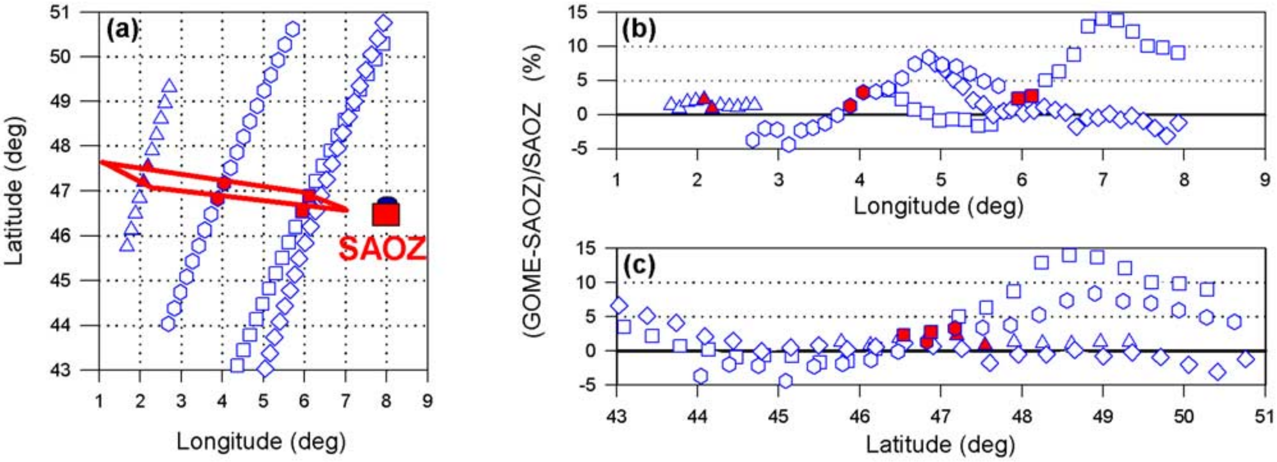


Figure 3. Effect of air mass differences on the GOME/SAOZ comparisons, illustrated at Jungfraujoch (Switzerland) on 4 September 1995. (a) Horizontal projection of the centers of individual GOME footprints recorded on this day in a radius of 500 km around the station (four kinds of blue open symbols represent four different GOME flight tracks); the red line indicates the horizontal projection of the SAOZ sunset air mass; GOME measurements crossing the SAOZ air mass with at least 25% of area overlap are highlighted in red. (b, c) Corresponding difference between GOME and SAOZ total ozone as a function of the longitude (Figure 3b) and latitude (Figure 3c) of the footprint centers; differences corresponding to the GOME footprints overlapping the SAOZ air mass are highlighted in red.

ozone distribution in a restricted vertical interval, and that the variability of the ozone field within this window is smaller than the anticipated data agreement. The reality is different: Remote sensing instruments actually report a three-dimensional weighted average of the ozone field. The concepts and formalism for vertical smoothing are described in detail by [*Rodgers*, 1990]. Horizontal smoothing has been much less studied.

[31] Vertical and horizontal smoothing properties depend not only on instrumental properties (e.g., viewing geometry, scanning strategy, field of view, etc.) but also on atmospheric and measurements parameters (solar elevation, aerosol loading, atmospheric profile of ozone, Rayleigh scattering geometry, surface albedo, presence of clouds, etc.). The GOME scanning strategy limits the horizontal resolution to the footprint dimensions (320 × 40 km^2 for the three forward scans). The TOMS Field of View is 40 × 40 km^2. The GOME and TOMS backscatter spectra are smoothed additionally since the measured radiation passes twice through the atmosphere (solar and scattered beams). Further, some solar and scattered radiation will pass through parts of the atmosphere that are outside the ground pixel box. Dobson and Brewer air masses are located between the measuring station and a maximum distance of 80 km in the direction of the sun. Zenith-sky air masses extend over several hundred kilometers in the azimuth of the rising and setting sun. Selection criteria based on simple time/space window limits can lead to comparisons based on very different perceptions of the atmospheric ozone burdens.

[32] Figure 3 illustrates this issue for a GOME/SAOZ total ozone comparison at Jungfraujoch on a day with large spatial gradients and high variability typical of the fall season over the European Alps. In this example, comparison of the SAOZ sunset data with all GOME pixels acquired within a radius of 500 km around the station would generate a systematic offset of nearly 5% and a 1σ scatter of 6%. The same GOME/SAOZ comparison based on daily averages of sunrise and sunset SAOZ data would generate even larger discrepancies, since sunrise SAOZ measurements are associated with high-ozone air masses located several hundred kilometres eastward from the station. On the other hand, if we compare only those GOME measurements with footprints overlapping by more than 25% with the effective air mass probed by SAOZ, then the mean agreement for this unstable day falls down to 2.5% and its 1σ scatter down to about 1%. The remaining 2.5% difference can be partly attributed to the use of older GDP 2.7 data in this illustration and partly to the 6-hour time interval between sunset (SAOZ acquisition) and mid-morning (GOME overpass). These horizontal smoothing uncertainties due to air mass extensions will in part be responsible for differences in the scatter patterns observed between GOME/SAOZ, GOME/Dobson, TOMS/SAOZ and TOMS/Dobson comparisons (see section 4). More generally, inappropriate selection of collocated data can dramatically increase the dispersion of the comparison results, from a few percent at middle latitudes to values approaching 60% percent at Antarctic stations located at the edge of the polar vortex and situated sometimes inside and sometimes outside the ozone hole. At stations experiencing stationary gradients, e.g., those near high mountain ranges with a north-south orientation or close to the vicinity of a stationary weather pattern, poor selection can even result in fictitious systematic biases up to 10% and beyond.

4. Ozone Column Validation

4.1. Test Case Studies in the Alps

[33] In order to assess the impact of the accuracy of the ground-based data on the validation of GOME and TOMS total ozone sensors, we have performed a comparative study

for a small region with collocated high-quality Brewer, Dobson and SAOZ ground-based total ozone measurements. Figure 4 shows for each season the percent relative difference between GOME and ground-based total ozone data, the latter recorded by three well-documented instruments at two NDSC Alpine stations: the Dobson and Brewer operated by MCH/ETHZ at Arosa (46.78°N, 9.68°E) and the SAOZ/UV-visible spectrometer operated by IASB-BIRA at Jungfraujoch (46.55°N, 7.98°E). Open circles represent results obtained with the earlier GDP versions 2.7 (gray open circles) and 3.0 (black open circles), while solid black circles indicate results for GDP 4.0. The ground-based data selected for the comparisons are based on recalibrated records of direct sun measurements only. In order to reduce the impact of air mass dependence, a filter is applied to the individual measurements to rule out Dobson and Brewer data recorded at solar elevations lower than 15°. For greater temporal coincidence, Dobson and Brewer time series are also filtered by a time window of two hours around the GOME overpass. With the Brewer (Figure 4, top), it is clear that the GDP 4.0 upgrade is an improvement: The fall-to-winter offset of +4% to −3% between GDP 2.7 and the Brewer, and of ±2% with GDP 3.0, has fallen with GDP 4.0 below the estimated accuracy limit of the ground-based data (this is about 1% for a well-calibrated Brewer instrument with direct sun observations [e.g., *Fioletov et al.*, 1999]). With the Dobson comparison (Figure 4, center), GDP 4.0 clearly outperforms GDP 2.7, but one cannot really conclude that the differences have improved from GDP 3.0 to GDP 4.0. In the latter case, a small seasonal variation persists, with a mean offset that is negligible in summertime but reaches a maximum of +3% in late fall and early winter (GOME SZA beyond 60°). With the SAOZ, the improvement with GDP 4.0 compared to GDP 3.0, and *a fortiori* to GDP 2.7, is clear: The yearly mean agreement does not exceed ±1%. From one 2.5° SZA bin to another, the mean GDP 4.0/SAOZ agreement ranges between +2.5% and −2%, while Brewer comparisons show only a limited range of ±1%. This difference in dispersion can be explained by the smaller number of coincidences with the SAOZ data (with the air mass matching technique, only one SAOZ value is relevant for each GOME pixel) compared to the larger number of GOME/Brewer coincidences (several Brewer values for each individual GOME pixel). Similar differences have been reported by [*Weber et al.*, 2005] when comparing GOME ozone data processed with the WF-DOAS algorithm with Brewer and Dobson data. The differences between Brewer, Dobson and SAOZ results are discussed in detail in the next paragraph.

[34] Dobson and Brewer total ozone data are processed with standard WMO recommendations [see, e.g., *Staehelin et al.*, 2003] at Arosa: Retrievals are based on absorption coefficients at a fixed temperature of −46.3°C. The sensitivity of the ozone absorption cross sections to temperature is typically 1.3% per 10K for Dobson spectrophotometers [*Komhyr et al.*, 1993], and 0.7% per 10K for Brewer instruments, leaving 0.6% per 10K for the difference between the two types of instruments [*Kerr et al.*, 1988]. It follows that the comparison results presented in Figure 4 are affected by the temperature dependence of the ozone absorption coefficients. Comparisons between collocated Brewer and Dobson measurements show a seasonally varying offset of about ±0.9% in amplitude for the Dobson data relative to the Brewer [*Staehelin et al.*, 2003], with a maximum of 2% offset in winter and no offset in summer. The seasonally varying overestimation presented in Figure 4, center, is comparable (but not equal) in phase and amplitude with the Dobson temperature effect (maximum offset in winter) calculated at this station [*Van Roozendael et al.*, 1998]. According to [*Staehelin et al.*, 2003], the temperature effect on Brewer data is less significant, and this is consistent with results in Figure 4, center. The data processing performed at IASB-BIRA on the Jungfraujoch SAOZ measurements includes a number of refinements to the standard SAOZ analysis procedure [*Lambert et al.*, 1998], and the resulting total ozone data does not show a significant AMF dependence. The agreement of the three closely collocated measurements at Arosa and at Jungfraujoch has been studied in detail [*Van Roozendael et al.*, 1998; *Staehelin et al.*, 2003]. When Dobson and Brewer data are corrected for their temperature and air mass dependencies, and provided that the SAOZ data is corrected for AMF variations and for station altitude, then the average agreement for the three ground-based data sets is better than 1%, and there is no detectable seasonal variation in the difference results. Taking into account the documented error budget of ozone column measurements at Arosa and Jungfraujoch, we can conclude that the average agreement between GDP 4.0 and ground-based measurements is close to ±1%.

4.2. Global Comparisons of Total Ozone

[35] GDP 2.7 overestimated ground-based total ozone during winter-spring by 2%, while during summer-autumn it underestimated ground-based ozone values by as much as 4%, the resulting amplitude is about 3%. In the tropics there was no significant seasonal dependence in the differences but in general, ozone was underestimated with GDP 2.7 by 1−2%. With GDP 3.0 the amplitude of the seasonal dependence over the middle and high latitudes was reduced by almost 50%. The remaining dependence was found to have almost the same phase with its counterpart in GDP 2.7 [*Lambert et al.*, 2002]. The change was mainly reflected in the winter-spring season where overestimation by GOME was minimized. In addition, GOME underestimated ground-based total ozone by 2% to 4% over desert areas. In the tropics, GDP 3.0 showed again a negative bias of about 2%, similar to that observed with GDP 2.7.

[36] The algorithm upgrades introduced in GDP 4.0 (see section 1.2) have generated in a different picture of the month-latitude cross sections of the differences between GOME and ground-based stations. The comparison results are shown in Figures 5a and 5b separately for the Dobson and the Brewer, and they are based on the selection of validation orbits as described in section 2.2. Figures 5c and 5d show the same month-latitude cross sections, but here direct sun Dobson and Brewer data are compared to TOMS V8 overpass data records. TOMS V8 has almost no bias against the Dobson total ozone observations and shows no seasonality in nearly all latitudes. The systematic TOMS V7 overestimation of Southern Hemisphere total ozone values has disappeared. Both GDP 4.0 and TOMS V8 results show substantial improvements compared to total ozone from previous operational versions. Remaining structures in time

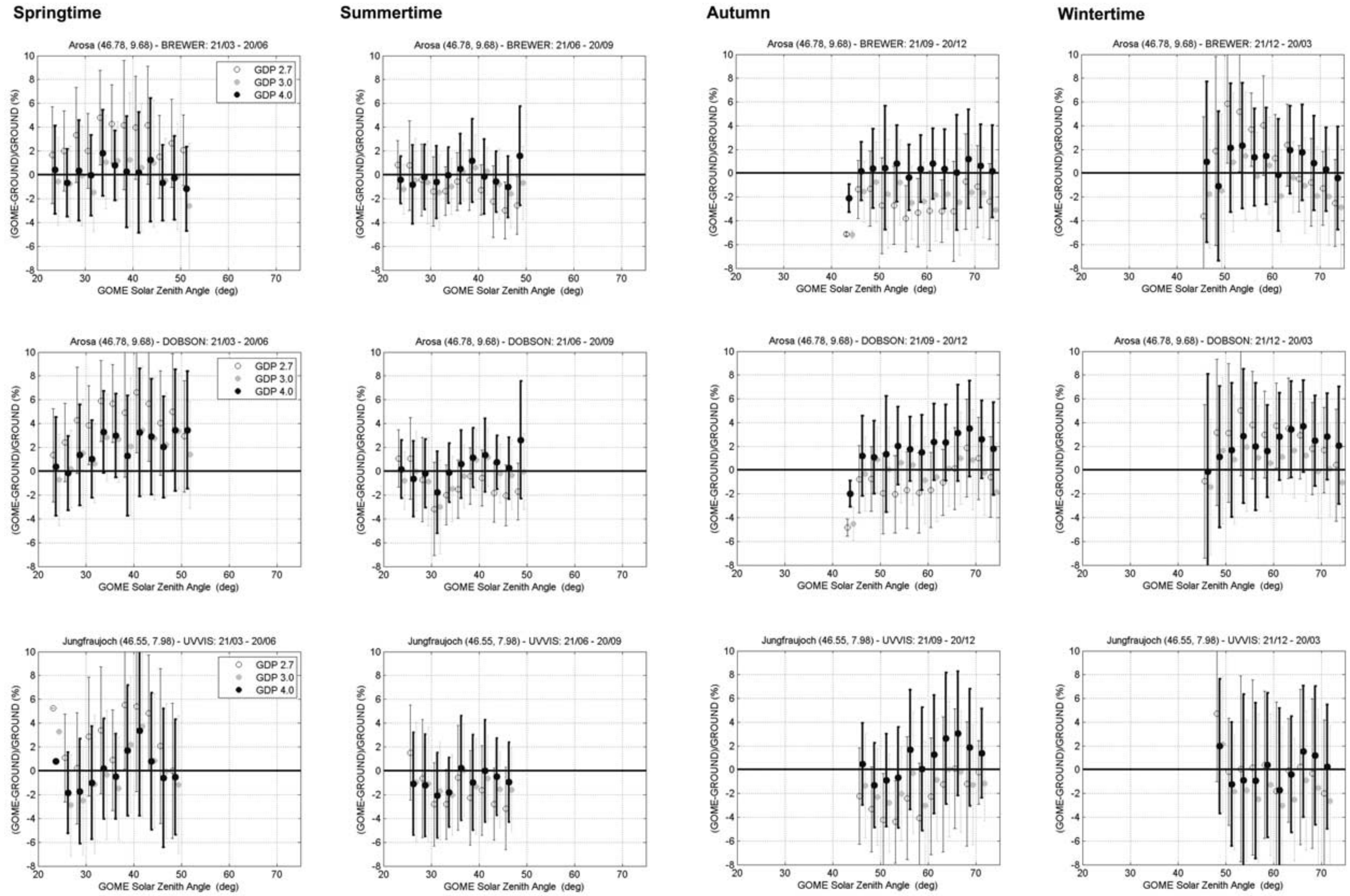
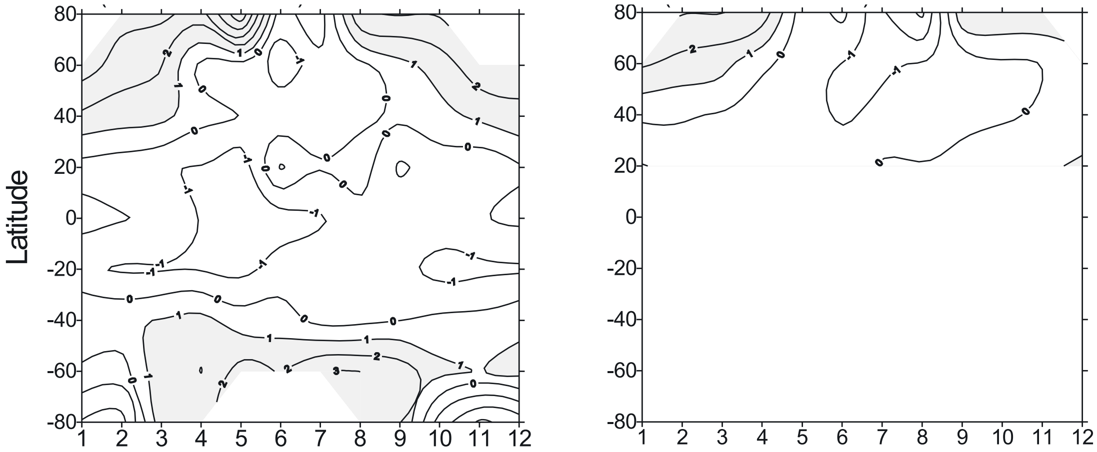


Figure 4. Percent relative difference between GOME and ground-based total ozone at two Network for Detection of Stratospheric Change (NDSC)/Alpine sites in Switzerland averaged over 2.5° solar zenith angle bins and plotted as a function of the GOME solar zenith angle. (top to bottom) Comparisons with MCH/ETH Brewer and Dobson at Arosa, and with BIRA-IASB SAOZ at Jungfraujoch. Seasons are indicated. (open black circles) GDP 2.7, (solid gray circles) GDP 3.0, and (solid black circles) GDP 4.0. Error bars show the 1σ standard deviation from the average agreement over the 2.5° SZA bin.

(a) (GDP 4.0 - **Dobson**)/Dobson [%]

(b) (GDP 4.0 - **Brewer**)/Brewer [%]



(c) (TOMS V8 - **Dobson**)/Dobson [%]

(d) (TOMS V8 - **Brewer**)/Brewer [%]

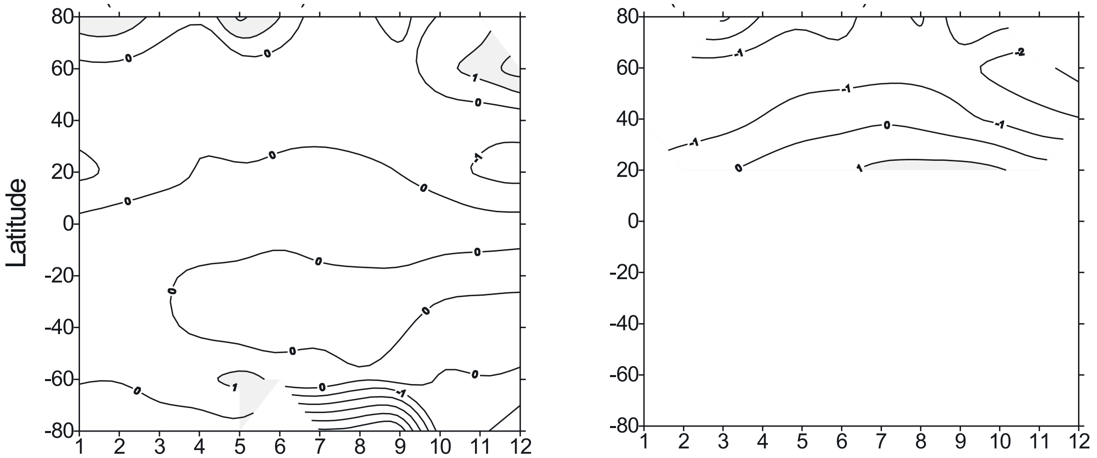


Figure 5. Month-latitude cross section of the relative difference between GOME and TOMS ground-based total ozone. GOME total ozone data are taken from the 4900 validation orbits processed with GDP 4.0, and the results obtained by comparison with (a) Dobsons and (b) Brewers are presented separately (note that nearly no Brewers operate in the Southern Hemisphere). TOMS version 8 total ozone overpass data are also compared separately with (c) Dobsons and (d) Brewers.

and latitude will be discussed below in sections 4.3 and 4.4 respectively. Systematic offsets of GDP 3.0 versus Brewer (−1.89%) and Dobson (−2.06%) data have all but vanished with GDP 4.0: to −0.17% for Brewer and 0.11% for Dobson comparisons. Zonal mean agreement varies within ±1−2%, with no marked meridian structure. Over the polar latitudes of both hemispheres, GDP 4.0 results are based on few measurements, especially at large SZA, and therefore they do not have the same significance with results at other latitudes. Studies relevant to high latitude and large SZA are detailed in section 4.5. TOMS has improved performance in all latitudes no meridian dependence of the differences. TOMS V8 data, when compared with Dobson instruments, have almost no offset (−0.3%), but they show a negative bias of about 1.5% compared with Brewer instruments.

4.3. Seasonal Dependence

[37] From Figures 5a−5d it can be concluded that there is still a small mean seasonal dependence remaining in the comparisons with the ground-based measurements north of 40°N and south of 40°S. The amplitude of this seasonality does not exceed 1%−1.5% for the Dobson comparisons and is even less for the Brewer comparisons. Over the tropics, the comparisons show results between 0 and 1% with nearly no seasonal variation. The GDP 4.0 seasonality is not in phase with the one observed in GDP 3.0, but it is rather in phase with the variation of the stratospheric temperatures (see 25°N−65°N zonal mean temperatures estimated by NOAA at 50 hPa (http://www.cpc.ncep.noaa.gov/products/stratosphere)): There is almost no bias to date during the warm period, while a positive bias of about 2% occurs during the cold period (including ozone hole conditions).

[38] As discussed in section 3 and in section 4.1 above, the use of absorption coefficients at a fixed temperature for both Dobson and Brewer total ozone measurements results in ozone column values dependent on stratospheric temperature variations. On the other hand, GDP 4.0 includes the spectral fitting of the effective temperature at which cross sections are calculated. Therefore GDP 4.0 should not

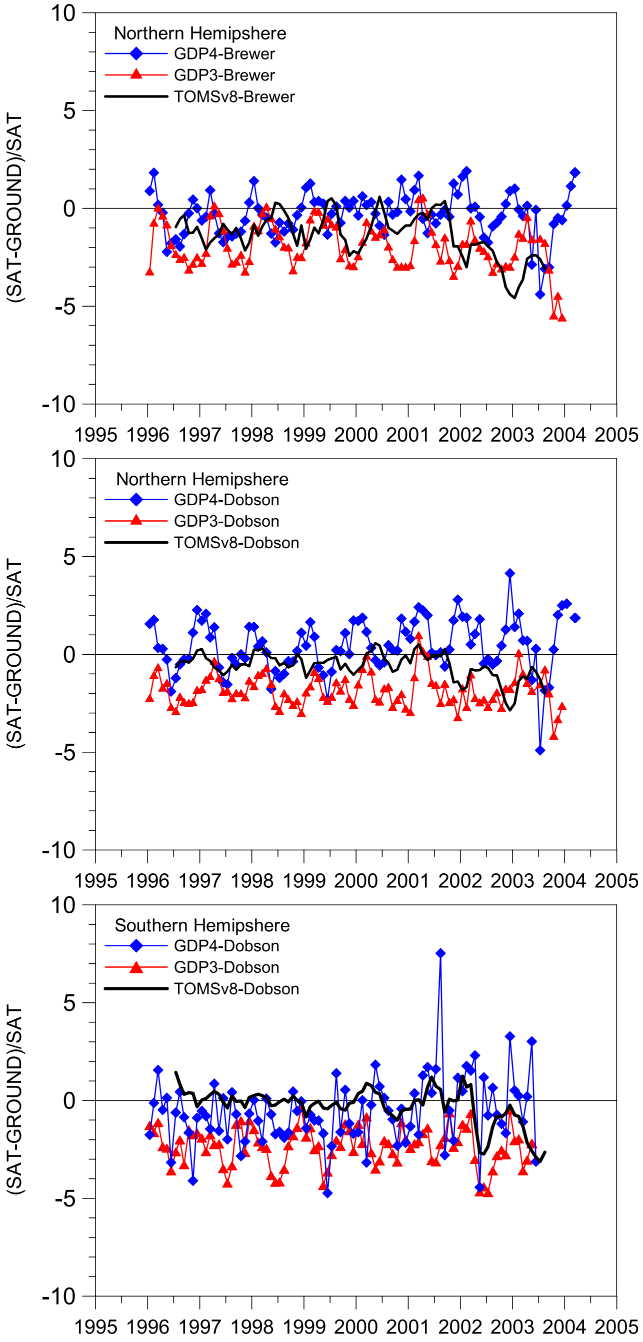


Figure 6. Time-series of the percent relative differences between satellite (GDP 3.0, GDP 4.0 and TOMS V8) and ground-based Dobson ozone data for the period 1996–2004, averaged over the Northern Hemisphere (top) and Southern Hemisphere (bottom).

depend significantly on the variability of stratospheric temperatures. A seasonal signature with similar phase and amplitude is found when comparing colocated Brewer and Dobson measurements from well-calibrated instruments for a long period (see examples at Hradec Kralove and Hohenpeiβenberg given by *Staehelin et al.* [2003]).

[39] Brewer comparisons have smaller amplitudes than their Dobson counterparts. This indicates that a large part of the observed differences between GDP 4.0 and ground-based measurements can be attributed to characteristics of the ground-based retrieval algorithms rather than to aspects of the GDP 4.0 algorithm. This is supported by the fact that the seasonal behavior of the GDP4.0/ground differences is in phase with the variability of stratospheric temperature. This seasonal behavior is also consistent with comparisons between collocated Brewer and Dobson measurements, which show similar degrees of phase and amplitude seasonality [*Stahelin et al.*, 2003]. Further confirmation is provided by the comparisons between Dobson, Brewer and TOMS V8 total ozone. Comparison results displayed in Figure 5c and 5d indicate that TOMS has almost no bias against the Dobson total ozone observations and shows no seasonality almost in every latitude belt. On the other hand the Brewer comparisons over middle latitudes show a small negative bias of -1% and a seasonal behavior with an amplitude of 0.5% in phase with the one found in GDP 4.0. Given that (1) the forward model of TOMS V8 uses a single temperature profile for calculating radiances, and (2) the TOMS wavelengths are closer to those for the Dobson instruments, we can infer that TOMS V8 and Dobson measurements have similar dependence on the lower stratospheric temperature. In consequence, their temperature-related errors are likely to cancel out when calculating their differences. This is also partly valid in the Brewer comparisons.

4.4. Long-Term Stability

[40] In the previous paper [*Van Roozendael et al.*, 2006], we demonstrated that GDP 4.0 ozone record shows a weak sensitivity not only to instrumental degradation but also to trends in atmospheric parameters other than ozone - in other words, the lack of "atmospheric drifts" interfering with the targeted total ozone trend [see *Van Roozendael et al.*, 2006, Figure 11]. In order to examine and compare the long-term stability of GDP 3.0, GDP 4.0 and TOMS V8 on a global basis, we averaged the individual station comparisons (based on all available orbits) separately over the Northern and Southern Hemisphere. Results for Dobson and with Brewer instruments were also separated. The Dobson comparisons are shown in Figure 6 for both hemispheres (Brewer results only for the Northern Hemisphere). It is evident from Figure 6 that until June 2003, both GDP 3.0 and GDP 4.0 total ozone data do not show any drift with respect to the ground-based data. (Because of technical problems on ERS-2, GOME data lack global coverage from June 2003 onward). However, sampling issues (less ground-based data available during 2003–2004 and fewer validation orbits) might increase the noise at the end of the time series. There is still a residual seasonal dependence of differences; this is different in phase compared to that for GDP 3.0 and also different in amplitude when considering Brewer or Dobson comparisons by 1% and 1.5% respectively. This remaining seasonal dependence is related to the one discussed already in section 4.3.

[41] In the Southern Hemisphere, GDP 4.0 is consistent with Northern Hemisphere results; however there are fewer measurements available for comparison and therefore the noise in the time series is larger. The TOMS V8 time series for the Northern Hemisphere do not show any drift until mid-2001. After that, TOMS/ground differences show an increasing bias, evident both in the Dobson and Brewer

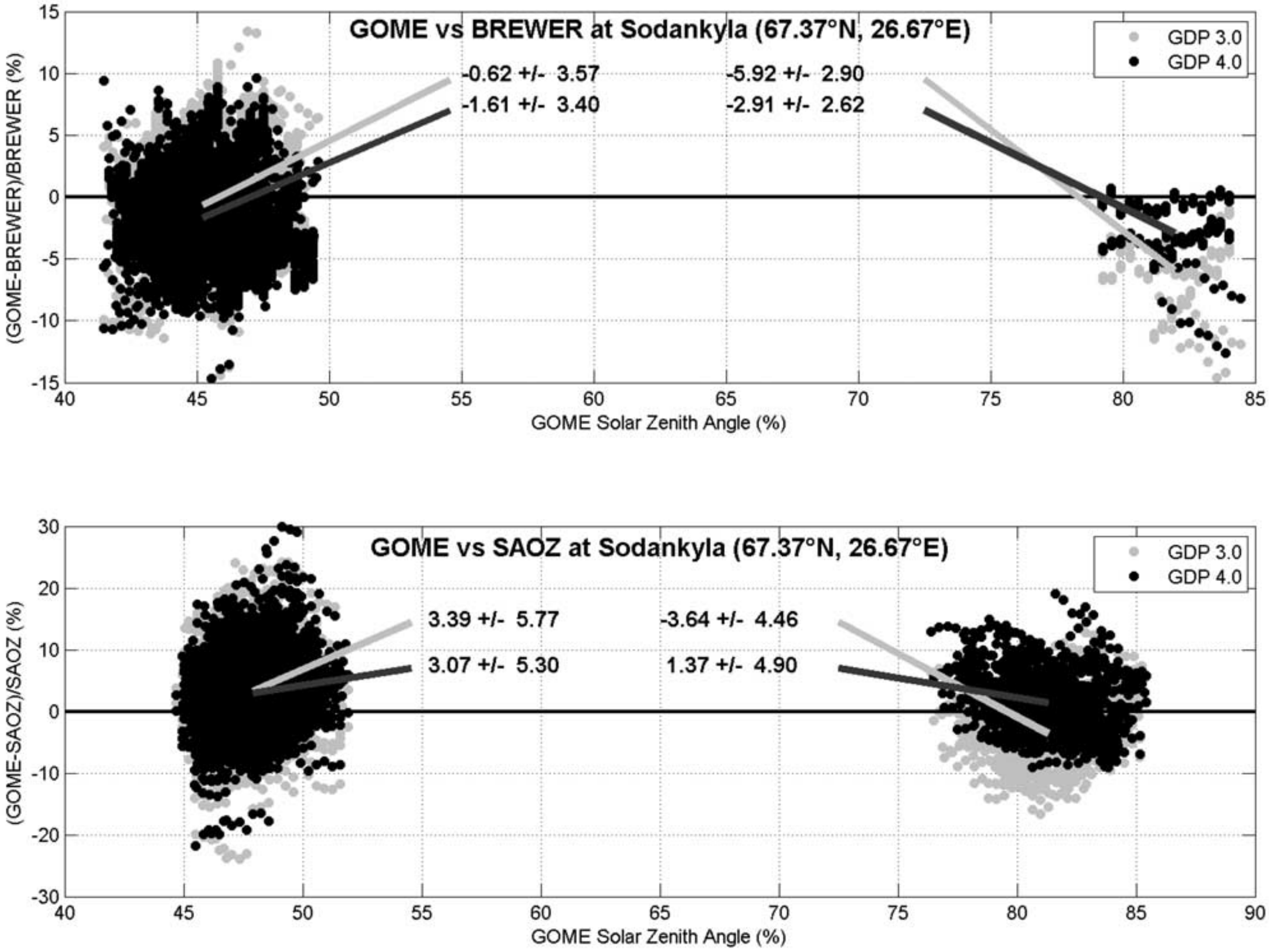


Figure 7. Comparison of the polar day SZA dependence of GDP 3.0 and GDP 4.0 total ozone at Sodankylä (Arctic Finland), using the FMI Brewer (top) and CNRS SAOZ (bottom) as standard transfers.

comparisons (see *Lambert and Balis* [2004] for details). The sizes of the GOME (320 × 40 km^2) and TOMS (40 × 40 km^2) footprints will generate differences in spatial smoothing of the ozone field, and this could partly explain why the scatter in the comparisons depends on the remote sensing instrument.

4.5. Dependence on Solar Zenith Angle

[42] Errors linked to the estimation of AMF and the correction for Raman scattering can result in a significant dependence of the vertical column product on the solar zenith angle (SZA) at which the observation was acquired. The SZA dependence in GDP 2.7 ozone columns possessed an average amplitude ranging from a few percent to ±15%, varying with latitude, season, vertical column amount, and sometimes from year to year [*Lambert et al.*, 2000, 2002]. With GDP 3.0, where the improved AMF calculation was based on an iterative neural network approach and the use of ozone profile climatology instead of modeling results, the amplitude of the SZA dependence was cut down by about 50% on average [*Lambert et al.*, 2002; *Spurr et al.*, 2005]. In GDP 4.0, with a new molecular Ring effect correction procedure and improvements in the AMF calculation, we expect to see a drastic reduction of the SZA dependence. This section assesses this reduction.

[43] We have already remarked in previous sections that it is more difficult to conclude that there is a clear reduction of the SZA/seasonal dependence at high latitudes. This difficulty arises not only as a result of the smaller number of delta validation coincidences with latitude, but also because of increasing uncertainties on both the satellite and ground-based measurements at such large latitudes and/or solar zenith angles. To detect and quantify possible improvements in the SZA dependences, we have first compared GOME data acquired during polar day at different solar elevations, to ground-based data acquired at fixed SZA values and therefore free from any air mass or SZA dependence. At high latitudes, the polar orbit of ERS-2 together with polar convergence of the meridians generates several daily overpasses. During polar summer, when the poles are illuminated permanently, GOME acquires measurements over the same high-latitude stations at least two times a day under different solar elevations. As the ozone field is usually relatively stable around the time of the summer solstice, this multiple daily overpass allows for the detection of the SZA dependence between GOME data acquired in the midmorning (medium SZA) and GOME data typical of midnight sun conditions (large SZA). Ground-based data are used in the comparison as a standard transfer.

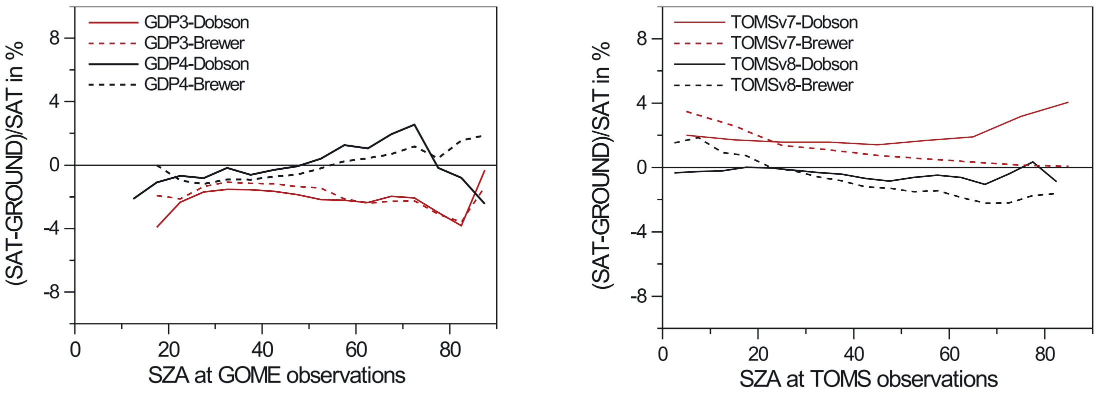


Figure 8. Percentage relative difference between GOME (GDP 3.0 and GDP 4.0) (left) and TOMS (v7 and v8) (right) and ground-based total ozone as a function of the satellite solar zenith angle (SZA). Dobson and Brewer comparisons are presented separately.

[44] For the present study, the polar day technique was applied to 6 polar stations in both hemispheres. Figure 7 shows GOME polar day SZA dependences with respect to two ground-based sensors operated at the NDSC/Arctic station of Sodankylä (67°N). Comparison with the FMI Brewer instrument (Figure 7, top) demonstrates that the average bias of about 6% observed between midmorning (moderate SZA) and midnight sun (large SZA) data with GDP 3.0 has reduced to a value of 1.5% with GDP 4.0. Comparison with the collocated CNRS/FMI SAOZ UV-visible instrument confirms this improvement. Similar comparisons were performed with the British Antarctic Survey (BAS) Dobson instruments operated at the Antarctic stations of Rothera (68°S) and Halley (76°S) [see *Lambert and Balis*, 2004]. At Rothera, the decrease of the polar day SZA dependence from 5.4% with GDP 3.0 to 0.9% with GDP 4.0 is similar to that reported in the Arctic. At Halley, where there was already no clear polar day SZA dependence with GDP 3.0, we observe no change with GDP 4.0.

[45] In Figure 8 (left), GOME total ozone data generated by both GDP 3.0 and GDP 4.0 are compared to Dobson and Brewer total ozone measurements on a global scale, as a function of the GOME SZA. In contrast to polar day studies where ground-based data are selected within fixed SZA ranges, it is important to take into account possible temperature/SZA dependences of the ground-based data that might interfere with dependences of the satellite data. It appears that in GDP 4.0 there is an overestimation of GOME for SZA between 60° and 70°, and a reversal of the SZA trend at 75° for the GDP 4.0/Dobson comparisons. These features were not evident in GDP 3.0, which had a completely different seasonal dependence. It is difficult to find unambiguous explanations for these dependences. However, when examining different latitude zones, we note that these results are associated with measurements of the cold period, and therefore temperature dependence effects in these differences might be enhanced; this is consistent with the fact that the Brewer comparisons show a smaller overestimation and no significant trend reversal. For SZA larger than 80°, there were only occasional data available from the ground stations for the GDP 4.0 comparisons, so the results should not be considered as statistically significant.

[46] The polar day technique could not be applied successfully to TOMS overpass data files provided by NASA/GSFC, since multioverpass data sets acquired during polar day are already filtered to provide only the ozone value acquired at the lowest SZA. In order to investigate the TOMS SZA dependence, we have carried out global studies relying on the statistical significance offered by network data. Figure 8 (right) shows SZA dependencies for TOMS V7 and TOMS V8 compared to Dobson and Brewer observations. Dobson comparisons show no significant SZA dependence, but this is not the case with the Brewer comparisons, where TOMS overestimates by 2% for small SZA and underestimates by −2% for large SZA. This behavior is similar when comparing Dobson and Brewer observations as a function of air mass without any temperature correction and therefore it could be probably attributed also partly to the temperature dependence of the comparisons.

4.6. Dependence on Ozone Column

[47] Both GDP 2.7 and TOMS V7 made significant overestimations of low ozone column values observed during Antarctic springtime ozone depletion. A similar effect was also detected in the Arctic, although this was more difficult to quantify because of the less stable conditions of the Arctic winter [*Lambert et al.*, 1998]. With the GDP upgrade to version 3.0, the large 10–20% overestimation of low column values (below 200 DU) by GDP 2.7 was greatly reduced to the 5% level for very low column values (below 130 DU) [*Lambert et al.*, 2002]. In this section, we examine changes in the ozone column dependence as a result of the GDP 4.0 upgrade. We expect also that the TOMS V7 column dependence is reduced in TOMS V8.

[48] In Figure 9, the column dependence of the difference between GOME (GDP 3.0 and GDP 4.0) and ground-based total ozone is presented for three different instruments (two Dobson and one SAOZ) operating in the Antarctic. Comparisons are limited to the ozone hole period (from 1 September to 1 November) in order to avoid interferences with possible SZA and seasonal dependencies. At all stations, the upgrade to GDP 4.0 generates an increase of the ozone column value by about 3% over the entire ozone column range; there is thus no perceptible change in the ozone column dependence. Similar comparisons were performed with TOMS V7 and TOMS V8 data, again using three

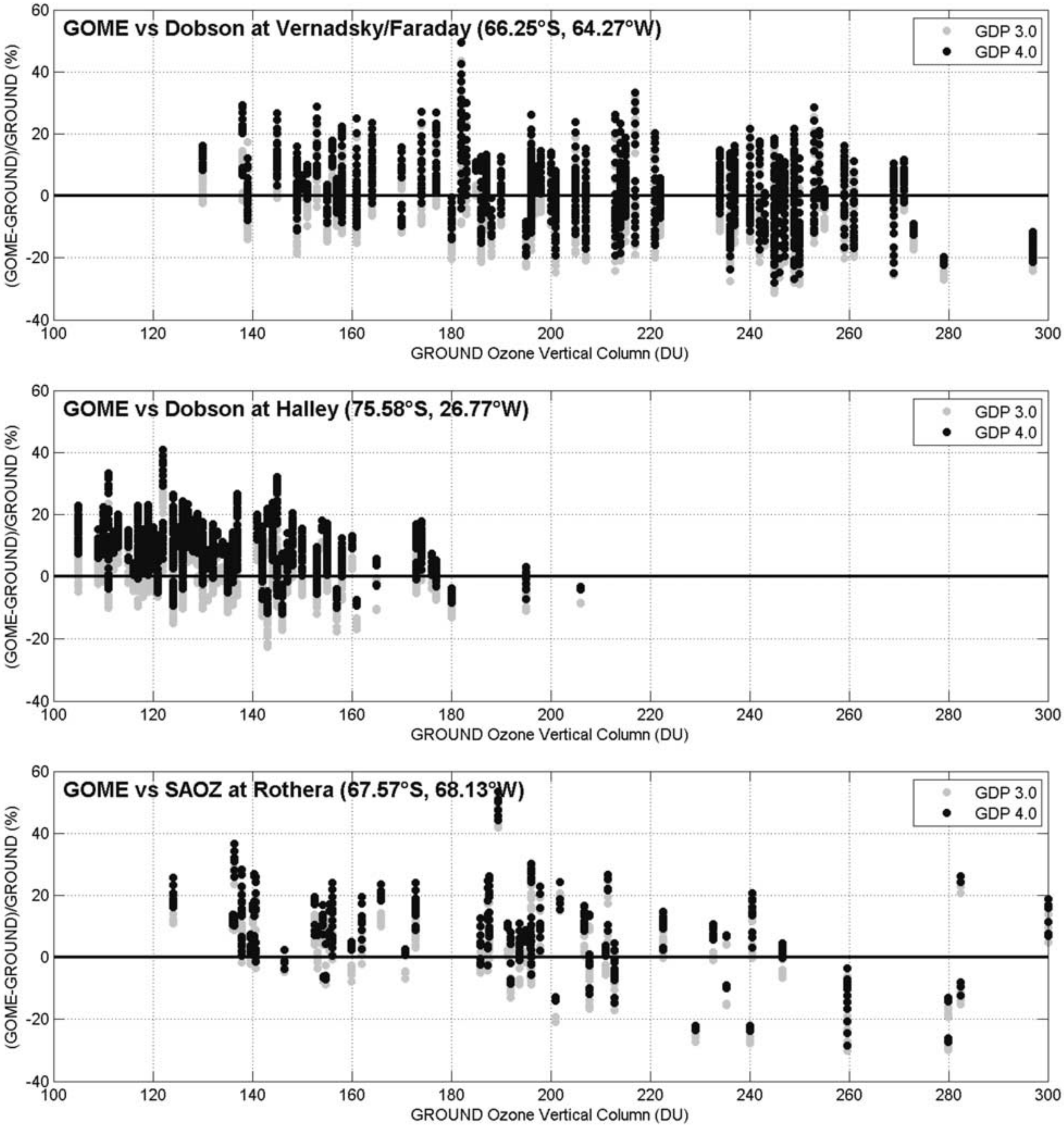


Figure 9. Percentage relative difference between GOME and ground-based total ozone during ozone hole conditions at three NDSC/Antarctic sites, as a function of the ground-based ozone column value: Dobson at Vernadsky (formerly Faraday), Dobson at Halley, and SAOZ at Rothera (ground-based data courtesy of the British Antarctic Survey). The main difference between GDP 3.0 and GDP 4.0 total ozone consists in an offset of a few percent, but there is no change to date in the GDP column dependence from one version to another.

different instruments operating in the Antarctic [*Lambert and Balis*, 2004]. In this case, the strong column dependence of TOMS V7 (about 15% below 130 DU) is much reduced, but the level of improvement is difficult to quantify because of the paucity of TOMS/ground coincidences.

4.7. Dependence on Fractional Cloud Cover

[49] In the GDP total ozone retrieval, reflecting clouds play an explicit role in the calculation of the AMF and in the estimation of the so-called ghost column of ozone below cloud top. Clearly, uncertainties in the GOME fractional cloud cover can generate significant errors in the retrieved

total ozone value. To date, the largest effect with GDP 2.7 was the 6% average offset identified in the Antarctic springtime [*Lambert et al.*, 2000, 2002]. This was due to the combined effects of (1) the use of an unsuitable ozone profile database with too high tropospheric ozone values for the given conditions; (2) a fractional cloud cover systematically set to 1, thus ensuring maximum application of the ghost column correction; and (3) a significant contribution of tropospheric ozone to the vertical column when stratospheric ozone is depleted. With GDP 3.0, this particular offset disappeared, thanks to the use of a more suitable ozone profile database for both AMF and ghost column estimations [*Lambert et al.*, 2002; *Spurr et al.*, 2005].

[50] Compared to GDP 3.0 results, the dispersion between GDP 4.0 and ground-based total ozone data has improved sometimes by a factor of two for clear sky scenarios. The situation is unchanged for cloudy scenes. (For examples of the changes in cloud fraction dependence between GDP 3.0 and GDP 4.0, see Figure 40 in [*Lambert and Balis*, 2004] for comparisons at the NDSC/Alpine station of Arosa and Figure 41 at high-latitude stations). The reduced clear-sky dispersion is likely the result of the new molecular Ring effect correction and improved AMF calculations in GDP 4.0.

5. Conclusion

[51] For the upgrade to GDP 4.0, improvements in the algorithm and their effects on the GOME total ozone products have been validated using a variety of methods based on comparisons with correlative measurements from extensive ground-based networks. In general, it is confirmed that modifications implemented in GDP 4.0 produced expected changes in the data products. However, it must be kept in mind that reported studies rely on a representative but limited set of orbits and unverified effects cannot therefore be ruled out.

[52] Total ozone data processed with different versions of the GOME data processor (GDP) have been validated from pole to pole through comparisons with ground-based measurements from Brewer and Dobson UV spectrophotometers and SAOZ/DOAS UV-visible spectrometers, as available from the WOUDC and NDSC data archives. TOMS V7 and V8 ozone column data have been studied similarly. Special attention has been paid to the quality control and documentation of ground-based data sets.

5.1. GOME Total Ozone

[53] The main achievement with GDP 4.0 is the drastic reduction of nearly all the remaining dependences on the latitude, SZA and season that persisted with GDP 3.0. The reduced ozone column dependence of GDP 3.0 has not changed with GDP 4.0. In general, the average agreement of GDP 4.0 with correlative ozone column measurements is now at the "percent level" at low and moderate SZA, that is, within the precision level of ground-based sensors when the latter are corrected for their own dependences on the season, solar elevation, temperature etc. At polar latitudes, and at GOME solar zenith angles larger than 80°, preliminary validation indicates that the agreement is slightly worse; however, average differences at low solar elevation usually do not exceed 5%. As noted already, the total column products do not suffer from any long-term drift of quality from 1995 to 2003, despite instrument degradation; the stability of the GDP 4.0 ozone data record enables it to be used confidently for ozone trend monitoring. In addition, the greatly improved quality of the GDP 4.0 ozone products makes them suitable for a wide variety of other geophysical research applications, such as polar process studies.

5.2. TOMS Total Ozone

[54] The upgrade to TOMS V8 is a marked improvement (with respect to the seasonal, meridian, and ozone column dependences) over V7. The systematic offset observed with TOMS V7 over the whole Southern Hemisphere has been eliminated. The TOMS V7 overestimation of extremely low ozone column values observed in the Antarctic springtime has also been corrected with TOMS V8. Seasonal and latitudinal dependencies still persist with TOMS V8 but their amplitudes are reduced to within 1% of ground-based stations. Both TOMS algorithms exhibit clear sensitivity to instrumental degradation: TOMS reports systematically lower ozone column values by a few percent from the second part of 2001 onward, when compared to more stable measurement systems. This result confirms the recommendation expressed by NASA/GSFC that EP TOMS ozone data acquired after 2000 should not be used for trend assessments.

5.3. GDP 4.0 Product

[55] The complete GOME data record from July 1995 onward has been reprocessed with GDP 4.0 and is available to the public via the ERS Help and Order desk. Documentation on GDP algorithms and products has been updated: a new algorithm theoretical basis document (ATBD) and Validation Report, a revised Product Specification Document (PSD), and GOME Data Disclaimer 2004. The GOME validation Web site has also been updated. To order GOME products or for further information, please contact EO Help Desk, ESA ESRIN, Via Galileo Galilei, I-00044, Frascati, Italy, +39 06 94180 777 (phone); +39 06 94180 272 (fax); eohelp@esa.int; http://earth.esa.int/gome).

[56] **Acknowledgments.** We would like to acknowledge the numerous observers who have taken systematic total ozone measurements at ground-based stations of the Global Ozone Observing System for sometimes more than 40 years and who have made their data available through the WOUDC and NDSC archives. Their sustained support in interpreting satellite/ground comparisons, and consequently in improving satellite data products, is greatly appreciated. Many scientists at a number of institutions in Europe and America have made contributions to the GOME data processing effort. In particular, we would like to thank colleagues from Uni. Bremen (Germany), KNMI (Netherlands), Uni. Heidelberg (Germany), BIRA-IASB (Belgium), DLR (Germany), SAO (USA), SRON (Netherlands), NASA Goddard (USA), RAL (Great Britain) and AUTH (Greece). We would also like to thank colleagues from the following agencies: ESA-ESTEC (Netherlands), ESA-ESRIN (Italy), and EUMETSAT (Germany). We would like to thank our reviewers for detailed comments; the paper has benefited significantly as a result. This work was funded mainly through an ESA contract: Upgrade of the GOME data processor for improved total ozone columns, ERSE-ESPR-EOPS-SW-04-0001. GOME-related activities at BIRA-IASB have also been supported by ProDEx and the Belgian Science Policy (PEA 15151/01/NL/SFe(IC)). In addition, V. Amiridis acknowledges the support of GSRT-Greece.

References

Bernhard, G., R. D. Evans, G. J. Labow, and S. J. Oltmans (2005), 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.

Bovensmann, H., J. Burrows, M. Buchwitz, J. Frerick, S. Noel, V. Rozanov, K. Chance, and A. Goede (1999), SCIAMACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, *56*, 127–150.

Bramstedt, K., J. Gleason, D. Loyola, W. Thomas, A. Bracher, M. Weber, and J. P. Burrows (2003), Comparison of total ozone from the satellite instruments GOME and TOMS with measurements from the Dobson network 1996–2000, *Atmos. Chem. Phys.*, *3*, 1409–1419.

Burkholder, J. B., and R. K. Talukdar (1994), Temperature dependence of the ozone absorption spectrum over the wavelength range 410 to 760 nm, *J. Geophys. Res.*, *21*, 581–584.

Burrows, J., et al. (1999), The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, *J. Atmos. Sci.*, *56*, 151–175.

Coldewey-Egbers, M., M. Weber, L. N. Lamsal, R. de Beek, M. Buchwitz, and J. P. Burrows (2005), Total ozone retrieval from GOME UV spectral data using the weighting function DOAS approach, *Atmos. Chem. Phys.*, *5*, 1015–1025.

Dobson, G. M. B. (1957), Observer's handbook for the ozone spectrophotometer, in *Annales International Geophysical Year, V, Part I: Ozone*, pp. 46–89, Elsevier., New York.

ESA (1995), *GOME Global Ozone Monitoring Experiment Users Manual*, edited by F. Bednarz, *ESA SP-1182*, Eur. Space Agency, Paris.

Eskes, H. J., R. J. van der A, E. J. Brinksma, J. P. Veefkind, J. F. de Haan, and P. J. M. Valks (2005), Retrieval and validation of ozone columns derived from measurements of SCIAMACHY on Envisat, *Atmos. Chem. Phys. Discuss.*, *5*, 4429–4475.

Fioletov, V., J. Kerr, E. Hare, G. Labow, and R. McPeters (1999), An assessment of the world ground-based total ozone network performance from the comparison with satellite data, *J. Geophys. Res.*, *104*, 1737–1747.

Kerr, J. B., C. T. McElroy, and W. F. J. Evans (1983), The automated Brewer spectrophotometer for measurement of SO_2, O_3, and aerosols, in *Proceedings of the WMO/AMS/CMOS Symposium on Meteorological Observations and Instrumentation*, pp. 470–472, Am. Meteorol. Soc., Boston, Mass.

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

Komhyr, W. D., C. L. Mateer, and R. D. Hudson (1993), Effective Bass-Paur 985 ozone absorption coefficients for use with Dobson ozone spectrophotometers, *J. Geophys. Res.*, *98*, 20,451–20,465.

Labow, G. J., R. D. McPeters, and P. K. Bhartia (2004), A comparison of TOMS, SBUV and SBUV/2 version 8 total column ozone data with data from ground stations, in *Ozone: Proceedings of the XX Quadrennial Ozone Symposium, 1–8 June 2004, Kos, Greece*, edited by C. S. Zerefos, pp. 123–124, Univ. of Athens, Greece.

Lambert, J. C. (2000), Combined characterization of GOME and TOMS total ozone measurements from space using ground-based observations from the NDSC, *Adv. Space Res.*, *26*, 1931–1940.

Lambert, J.-C., et al. (1998), GOME and TOMS total ozone in northern winter 1996/1997: Comparison with SAOZ/UV-visible ground-based measurements in the Arctic and at middle latitude, in *Polar Stratospheric Ozone 1997*: *Air Pollution Research Report 66. Proceedings of the Fourth European Symposium 22 to 26 September 1997, Schliersee, Bavaria, Germany*, edited by N. R. P. Harris, I. Kilbane-Dawe, and G. T. Amanatidis, pp. 696–699, Eur. Comm., Brussels.

Lambert, J.-C., M. Van Roozendael, M. De Mazière, P. C. Simon, J.-P. Pommereau, F. Goutail, A. Sarkissian, and J. F. Gleason (1999), Investigation of pole-to-pole performances of spaceborne atmospheric chemistry sensors with the NDSC, *J. Atmos. Sci.*, *56*, 176–193.

Lambert, J.-C., et al. (2002), *ERS-2 GOME GDP3.0 Implementation and Delta Validation Report for GOME Level-1-to-2 data processor Upgrade to Version 3.0*, *ESA Tech. Note ERSE-DTEX-EOAD-TN-02-0006*, Eur. Space Agency, Paris.

Lambert, J.-C., and D. Balis (2004), *Delta Validation Report for ERS-2 GOME data processor Upgrade to Version 4.0*, *ESA Tech. Note ERSE-CLVL-EOPG-TN-04-0001*, Eur. Space Agency, Paris.

Loyola, D., et al. (1997), Ground segment for ERS-2 GOME data processor, in 3rd Symposium on Space in the Service of our Environment, Florence, Italy, *Eur. Space Agency Spec. Publ., ESA SP*, *414*, 591–597.

Papoulis, A., and S. Pillai (2002), *Probability, Random Variables and Stochastic Processes*, 4th ed., McGraw-Hill, New York.

Pommereau, J.-P., and F. Goutail (1988), Ground-based measurements by visible spectrometry during Arctic winter and spring 1988, *Geophys. Res. Lett.*, *15*, 891–894.

Rodgers, C. D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*, 5587–5595.

Roscoe, H. K. (1999), Slant column measurements of O3 and NO2 during the NDSC intercomparison of zenith-sky UV-visible spectrometers in June 1996, *J. Atmos. Chem.*, *32*, 281–314.

Sarkissian, A., H. Roscoe, D. Fish, M. Van Roozendael, M. Gil, A. Dahlback, L. Perliski, J.-P. Pommereau, and J. Lenoble (1995), Ozone and NO2 air-mass factors for zenith-sky spectrometers: Intercomparison of calculations with different radiative transfer models, *Geophys. Res. Lett.*, *22*, 1113–1116.

Spurr, R., et al. (2005), GOME level 1-to-2 data processor version 3.0 (2005): A major upgrade of the GOME/ERS-2 total ozone retrieval algorithm, *Appl. Opt.*, *44*, 7196–7209.

Staehelin, J., J. Kerr, R. Evans, and K. Vanicek (2003), *Comparison of Total Ozone Measurements of Dobson and Brewer Spectrophotometers and Recommended Transfer Functions*, *WMO TD 1147*, *GAW Rep. 149*, 39 pp., World Meteorol. Org., Switzerland.

Stammes, P., P. Levelt, J. de Vries, H. Visser, B. Kruizinga, C. Smorenburg, G. Leppelmeier, and E. Hilsenrath (1999), Scientific requirements and optical design of the ozone monitoring instrument on EOS-CHEM, in *Earth Observing Systems IV*, vol. SPIE 3750, edited by W. L. Barnes, pp. 221–232, Int. Soc. for Opt. Eng. (SPIE), Bellingham, Wash.

Van Roozendael, M., et al. (1998), Validation of ground-based UV-visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, *J. Atmos. Chem.*, *29*, 55–83.

Van Roozendael, M., et al. (2006), Ten years of GOME/ERS-2 total ozone data—The new GOME data processor (GDP) version 4: 1. Algorithm description, *J. Geophys. Res.*, *111*, D14311, doi:10.1029/2005JD006375.

Weber, M., L. N. Lamsal, M. Coldewey-Egbers, K. Bramstedt, and J. P. Burrows (2005), Pole-to-pole validation of GOME WFDOAS total ozone with ground based data, *Atmos. Chem. Phys.*, *5*, 1341–1355.

WMO (1998), Scientific Assessment of Ozone Depletion: 1998, *Global Ozone Res. and Monit. Proj. Rep. 44*, World Meteorol. Org., Geneva.

WMO (2003), Scientific Assessment of Ozone Depletion: 2002, *Global Ozone Res. and Monit. Proj. Rep. 47*, World Meteorol. Org., Geneva.

V. Amiridis and D. Balis, Aristotle University of Thessaloniki, Box 149, 54124 Thessaloniki, Greece.

P. Gerard, J. Granville, J.-C. Lambert, and M. Van Roozendael, Belgian Institute for Space Aeronomy, 3 Avenue Circulaire, B-1180 Brussels, Belgium.

R. Spurr, RT Solutions, Inc., 9 Channing Street, Cambridge, MA 02138, USA. (rtsolutions@verizon.net)

Y. Livschitz, D. Loyola, and P. Valks, German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), P.O. Box 1116, D-82234 Wessling, Germany.

C. Zehner, ESA/ESRIN, Via Galileo Galilei CP.64, 00044 Frascati, Italy.