

**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.**

M. E. Koukouli^{1*}, D. S. Balis¹,
D. Loyola², P. Valks², W. Zimmer², N. Hao²,
J-C. Lambert³, M. Van Roozendael³, C. Lerot³
and R. J. D. Spurr⁴.

1. Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki,
Thessaloniki, Greece.

2. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik der
Fernerkundung (IMF), 82234 Oberpfaffenhofen, Germany.

3. Belgian Institute for Space Aeronomy (BIRA-IASB), 3, Avenue Circulaire, B-1180
Brussels, Belgium.

4. RT-Solutions, Inc., Cambridge, MA, United States.

*Corresponding author: mariliza@auth.gr

Abstract. The main aim of the paper is to assess the consistency of five years of Global Ozone Monitoring Experiment-2/Metop-A [GOME-2] total ozone columns and the long-term total ozone satellite monitoring database already in existence through an extensive inter-comparison and validation exercise using as reference Brewer and Dobson ground-based measurements. The behaviour of the GOME-2 measurements is being weighed against that of GOME (1995-2011), Ozone Monitoring Experiment [OMI] (since 2004) and the Scanning Imaging Absorption spectroMeter for Atmospheric CartograpHY [SCIAMACHY] (since 2002) total ozone column products. Over the background truth of the ground-based measurements, the total ozone columns are inter-evaluated using a suite of established validation techniques; the GOME-2 time series follow the same patterns as those observed by the other satellite sensors and in particular, on the average, GOME-2 data underestimate GOME data by about 0.80%, and underestimate SCIAMACHY data by 0.37% with no seasonal dependence of the differences between GOME-2, GOME and SCIAMACHY. The latter is expected since the three data sets are based on similar DOAS algorithms. This underestimation of GOME-2 is within the uncertainty of the reference data used in the comparisons. Compared to the OMI sensor, on the average GOME-2 data underestimate OMI_DOAS (collection 3) data by 1.28%, without any significant seasonal dependence of the differences between them. The lack of seasonality might be expected since both the GOME Data Processor [GDP] 4.4 and OMI_DOAS are DOAS-type algorithms and both consider the variability of the stratospheric temperatures in their retrievals. Compared to the OMI_TOMS (collection 3) data, no bias was found. We hence conclude that the GOME-2 total ozone columns are well suitable to continue the long-term global total ozone record with the accuracy needed for climate monitoring studies.

1 Introduction

With the launch of Metop-A on October 19th 2006, a polar orbiting satellite carrying both environmental and meteorological instruments, a new global total ozone data record was started with the GOME-2 instrument. In *Loyola et al.*, 2011, the GOME

Data Processor (GDP) version 4.4 retrieval algorithm for GOME-2 total ozone, as well as the first global validation results for three years (2007-2009) using Brewer and Dobson ground-based measurements, were described. Several algorithm improvements were introduced in GDP 4.4 compared to previous versions [van Roozendaal *et al.*, 2006] such as the improved cloud retrieval algorithms including the discrimination of sun-glint effects, the enhanced treatment for ice and snow conditions, the intra-cloud ozone correction, the updated radiative transfer modelling for large viewing angle conditions and the empirical correction to eliminate a scan angle dependency caused by an unknown bias. To be precise, the forward-scan west-east dependency of more than +1.5% in the previous GDP 4.x data has been largely eliminated in the GDP 4.4 record. The GDP 4.4 algorithm is able enough to deliver real-time data for operational needs while being truly robust in performance. As far as the validation was concerned, for middle latitude comparisons, it has already been shown that the reprocessed GOME-2 GDP 4.4 dataset under-estimates ground-based Dobson ozone by 0.5% in the Northern Hemisphere whereas it over-estimates in the Southern Hemisphere by the same amount. For northern high latitude comparisons, a good comparison relative to the Dobson measurement is found while for the southern high latitude comparisons, an under-estimation of less than 1% is observed. For the tropical stations GOME-2 under-estimates on average by 0 to 2% against the Dobson network. The GOME-2 versus Brewer comparisons over the Northern Hemisphere follow closely the GOME-2 vs. Dobson comparisons and illustrate an underestimation of 1% which tends to be slightly higher (1-2%) over the Arctic.

Equally important to the stand-alone validation and quality assurance of new atmospheric composition products is to ensure the continuity in the record between current and past missions. Even though the processors analysing the atmospheric signal might be the same, even slight differences in spectral and spatial resolution, detector technology, swath width, and other characteristics, might bring notable differences between the resulting total ozone products from different sensors. In the following, we will delve into the comparisons between GOME-2 total ozone for the years 2007 to 2010 against other three satellite instruments active during these time period, using as reference a background of ground-based total ozone measurements. In Section 2 we discuss the different datasets used with some details on both instruments and algorithms given as well as known issues with the total ozone

records, as these were recorded in literature. In Section 3 we show the results and their discussion with the summary and conclusions given in Section 4.

2 Data Sets and methodology.

2.1 The Four Satellite Instruments

A very brief description of the four different instruments and the relevant algorithms is given in this following section. For quick reference to the information, the main features of the four instruments, satellite platforms and data versions used in this study are further summarized in Table 1.

2.1.1 GOME on board ERS-2

The Global Ozone Monitoring Experiment, (GOME), is an across-track nadir-viewing spectrometer on board ERS-2 which is a Sun-synchronous polar orbiting satellite with a period of about 100 min and a local equator crossing time of 10.30h. In normal viewing mode, GOME performs three forward scans followed by a backward scan. Each forward scan has a footprint size of $320 \text{ km} \times 40 \text{ km}$ for a 1.5-s detector readout integration time. The maximum swath is 960 km, with nominal scan angle $\pm 31^\circ$ at the spacecraft, global coverage is achieved at the equator within three days. GOME has 3584 spectral channels distributed over four serial readout detectors; the wavelength range is 240 to 793 nm, with a moderate spectral resolution of 0.2 to 0.4 nm. More details on the GOME instrument are given by *Burrows et al.*, 1999a. The instrument was switched off on July 5th, 2011 when the ERS-2 satellite was decommissioned. The GOME total ozone columns presented in this work here are the operational GDP 4.1 products [*Van Roozendaal et al*, 2006]. Known issues with the GOME total ozone columns reported in [*Balis et al.*, 2007a] include a small mean seasonal dependence remaining 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. Also, there is an overestimation of GOME for SZA between 60° and

70°, and a reversal of the SZA trend at 75° for the GDP 4.1/Dobson comparisons. 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 4.1 ozone data record enables it to be used confidently for ozone trend monitoring.

2.1.2 SCIAMACHY on board ENVISAT

The Scanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY) was launched in March 2002 aboard the European platform ENVISAT and has been operational for more than ten years providing global coverage in approximately six days [Bovensmann *et al.*, 1999]. ENVISAT is in a Sun-synchronous orbit with an inclination of 98.5°, a mean altitude of 796 km and has a period of 100 min, performing about 14 to 15 orbits per day. SCIAMACHY is an eight-channel spectrometer covering the spectral range from 240 nm to 2380 nm and uses different viewing geometries for retrieving total trace gas columns (nadir) and profiles (limb and solar/lunar occultation). The nominal swath is 960 km with a typical footprint size of 60 km x 30 km for ozone observations, global coverage is achieved at the equator within six days. The total ozone column data used in this study are from the SCIAMACHY Differential Optical Absorption Spectroscopy (SDOAS) algorithm, which is the prototype algorithm for GOME(-2) and SCIAMACHY operational products GDP4.x and SGP5.0, with details found in Lerot *et al.*, 2009. With respect to ground-based data on the whole, there is no appreciable systematic bias and more than 75% of the measurements used in this paper agree within 5% for the set of ground stations used in Lerot *et al.*, 2009. It is also important to note that the SCIAMACHY total O₃ columns suffer from a small but statistically significant decreasing trend, ranging between from between -0.20 and -0.50% per annum. The issue is related to the well-known degradation effect in SCIAMACHY which generates time-dependent spectral features in the measured reflectances and introduces an artefact trend in the ozone total columns.

2.1.3 OMI on board Aura

The Ozone Monitoring Instrument (OMI) [Levelt *et al.*, 2006] is one of four instruments aboard the NASA EOS-Aura satellite, launched on 15 July 2004 [Schoeberl *et al.*, 2006]. OMI is a compact nadir viewing, wide swath (daily global coverage), ultraviolet-visible (270 nm to 500 nm) imaging spectrometer that was contributed to the Aura mission by Netherlands and Finland. The foot pixel size at nadir is 13 km x 25 km. In contrast to GOME and SCIMACHY, the foot pixel size is not constant but increases for the off-nadir positions. Two total ozone column data products are available: the OMI-TOMS data product which is based on the long-standing TOMS V8 retrieval algorithm [Bhartia *et al.*, 2004] and the OMI-DOAS data product which is a DOAS type algorithm [Veefkind *et al.*, 2006] developed by the Royal Netherlands Meteorological Institute (KNMI). Balis *et al.*, 2007b, have shown that, although both algorithms infer total ozone column data for OMI ground pixels, they differ in many aspects of their algorithmic approach and this is reflected in the validation and comparison campaigns. Balis *et al.*, 2007b, showed a globally averaged agreement of better than 1% for OMI-TOMS data and better than 2% for OMI-DOAS data with the ground-based observations. The OMI-TOMS data product was found to be of high overall quality with no significant dependence on solar zenith angle or latitude. The OMI-DOAS data product had no significant dependence on latitude except for the high latitudes of the Southern Hemisphere where it systematically overestimated the total ozone value. In addition a significant dependence on solar zenith angle is found between OMI-DOAS and ground-based data. In this work, both products will be considered and were extracted from the Aura Validation Data Centre [http://avdc.gsfc.nasa.gov/]. The OMI_TOMS and OMI_DOAS level-2 total ozone data are based on collection 3 level 1b data and have been processed with TOMS v8.5 and OMDOAO3 v1.0.1 algorithms respectively (see ATBD documents at <http://toms.gsfc.nasa.gov> and <http://www.temis.nl>). The well-known “OMI row anomaly” issue, which affects particular viewing directions which correspond to rows on the CCD detector, has been dealt with by discarding observations with the equivalent quality flag as discussed in detail at <http://www.knmi.nl/omi/research/product/index.php>.

2.1.4 GOME-2 on board MetOp-A

The Global Ozone Monitoring Experiment-2 (GOME-2) instrument is mounted on the flight-direction side of the MetOp-A satellite. GOME-2 is a nadir-viewing scanning spectrometer, with an across-track scan time of 6 seconds and a swath width of 1920 km. Global coverage of the sunlit part of the atmosphere can be achieved almost within one day. GOME-2 ground pixels have a footprint size of 80 km x 40 km, four times smaller than those for GOME (320 km x 40 km), and also improved polarisation monitoring and calibration capabilities [Munro *et al.*, 2006]. In the framework of EUMETSAT's Satellite Application Facility on Ozone and Atmospheric Chemistry Monitoring (O3M-SAF), GOME-2 total ozone data are processed at DLR operationally, both in near real time and offline, using the GDP 4.x algorithm [Valks *et al.*, 2011]. However, operational data were based on different versions of level 1b data, and in addition, the satellite measurements showed a scan angle dependency. In order to provide a homogeneous data set the complete GOME-2 dataset was reprocessed with GDP 4.4 at the end of 2009 using the most recent level 1b data (version 4), plus an additional empirical correction for the east-west scan dependencies as described in Loyola *et al.*, 2011. Refer to this paper for further details on the algorithm and the validation of the total ozone column against ground-based instruments.

Within the framework of the O3M-SAF project, the Laboratory of Atmospheric Physics (LAP) at Aristotle University of Thessaloniki, in collaboration with the Hellenic National Meteorological Service, has developed a total ozone validation facility for GOME-2 data, be found at <http://lap.physics.auth.gr/eumetsat/totalozone>.

2.2 The ground based measurements

Archived total ozone column measurements from the WMO/GAW network that are routinely deposited at the WOUDC in Toronto, Canada (<http://www.woudc.org>) were used as ground reference. The WOUDC archive contains total ozone column data mainly from Dobson and Brewer UV spectrophotometers and from M-124 UV filter radiometers. A well maintained and calibrated Dobson spectrophotometer measures the ozone column with an estimated accuracy of 1% for direct Sun observations and 2–3% for zenith sky or zenith cloud observations for Sun elevation higher than 15°. The Dobson spectrophotometer is a large and manually controlled two-beam instrument based on the differential absorption method in the ultraviolet Huggins

band where ozone exhibits strong absorption features. The measurement principle relies on the ratio of the direct sunlight intensities at two standard wavelengths. Since 1957, Dobson spectrophotometers have been deployed operationally in a worldwide network. The Brewer grating spectrophotometer is in principle similar to the Dobson, however, it has an improved optical design and is fully automated. The ozone column abundance is determined from a combination of four wavelengths between 306 nm and 320 nm. Since the 1980s, Brewer instruments are part of the ground-based networks as well. Most Brewers are single monochromators, but a small number of systems are double monochromators with improved stray light performance.

The WOUDC stations considered and the reasoning behind the particular selection of stations have already been presented in a series of validation papers, such as for the validation of ten years of GOME observations [Balis *et al.*, 2007a], the OMI-TOMS and OMI-DOAS dataset [Balis *et al.*, 2007b], a new version of the OMI-TOMS algorithm [Antón *et al.*, 2010] and the GOME-2 validation [Loyola *et al.*, 2011].

The selection investigation and criteria have been discussed in detail in both Balis *et al.*, 2007a and 2007b. We offer here a brief summary for completion purposes. For each ground-based station a series of statistics and plots were performed, separately for direct sun measurements and zenith-sky observations. Daily coincidences with the satellite central latitude and longitude of the footprint falling within a 150km radius of the ground station were found and used for the creation of monthly, seasonal and yearly time series and scatter plots. Since the WOUDC data are daily means, there is no temporal treatment of the satellite observations. The percentage of the relative differences between ground and satellite TOC is used as the comparative tool for the validation. The statistics are then performed on a per latitude belt basis, per Hemisphere basis and per the totality of the data basis, always keeping the two types of ground-based instruments separate and using only direct sun observations, as the most reliable. In the following, per Hemisphere and per the totality of the data comparisons will be shown and discussed upon, for space reasons.

3 Results and Discussion

In the following section, the long term performance of GOME-2 total ozone columns will firstly be examined against that of the three instruments discussed above over the

background of the Dobson and Brewer total ozone columns. In order to compare possible trends, features and regions of special interest, contour representations of the latitudinal variability of the differences against time are presented first. This will enable the pictorial identification of regions and times where each of the instruments might have faced issues, as well as the long term stability of all total ozone column records. Secondly, the time series of the differences averaged over the Northern Hemisphere enables the study of the seasonality effects and possible trends. Thirdly, a contour representation of the differences as a function of SZA and season permits the identification of lingering SZA and seasonal dependencies and how these compare among the different algorithms and instruments.

A point to be made at this point is a possible source for the differences between the Dobson and the Brewer comparisons shown later on in this paper. There exists an inherent deviation between the treatment of the ground-based measurements and the analysis of the satellite observations and that is the choice of ozone absorption cross-sections; the ground-based data are analysed using the Bass & Paur, 1985, ozone cross-sections whereas with the GDP 4.x family of algorithms, cross-sections measured with each instrument are used, namely the GOME measured flight model cross sections for O_3 , known as the GOME FM98 data [Burrows et al., 1999b] for the GOME and GOME-2 instruments and the SCIAMACHY flight model data (SFM) (Bogumil et al., 2003) for the SCIAMACHY instrument. The exception being TOMS V8 which uses, to the best of our knowledge, the Bass & Paur, 1985, cross-sections. It certainly falls within the plans for the next generation of algorithms, namely the GDP5.0 Direct Fitting algorithm, to harmonise the use of cross-sections for the satellite observations, however this will come at a later stage. On the impact of the temperature, the point as we understand it is more whether some of the reported differences could come from the fact that ground-based data sets use cross-sections at a fixed temperature (while temperature dependences are accounted for in satellite algorithms). We postulate that the answer is most probably yes, and it probably explains part of the different behaviour between Dobson and Brewer instruments as the Dobson measurements are more sensitive to temperature variations.

3.1 Long term stability

3.1.1 Latitudinal behaviour

In Figure 1 a contour representation of the difference between each of the four satellite instruments and the Dobson total ozone columns is shown in order to visualise the long-term picture of each instrument for the totality of the data available and not only the coincidences to GOME-2. Hence the stability of each instrument and algorithm behaviour may be examined. The comparative graphs are ordered from top to bottom in launch date from the one furthest back in time to the present. In order to create as homogeneous and hence comparable graphs possible, these contour representations were created with a spread of 0.2 of the year [i.e. one and a half months approximately] for the x-axis and 30 degrees in latitude for the y-axis. The original data were averaged with a one month spacing in time and 15 degrees in latitude. A common feature for all graphs which should be discussed here is the high negative values [more than -2.5%] around 15 degrees North which are due to the station of Bangkok, the sole station of the belt, which presents a near-constant ground-based over-estimation from year 2005 onwards. This feature may appear more pronounced in some figures than others depending on the averaging/binning performed as needed. The station is not excluded from any comparisons.

In the upper left of Fig. 1, the near 15 years of GOME total ozone monitoring is presented with the large gaps in the southern hemisphere after year 2003 due to the ERS-2 tape recorder failure. The gaps at the high latitudes, also noticeable for the other instruments are due to the lack of ground-based stations, and not a satellite effect. A definite seasonal effect persists through the years with lows [underestimation] during the summer months for the Northern Hemisphere. The general picture appears to be that the GOME measurements oscillate between over- and under-estimating the ground based estimates between -0.5% and 1% depending on the latitude belt and the season. This seasonal dependency northwards and southwards of 40° has also been reported in [Balis *et al.*, 2007a] with an amplitude that does not exceed 1%–1.5% for the Dobson comparisons. In the upper right, a similar structure can be observed for the SCIAMACHY comparisons which start at the end of year 2002 with some instrumental issues still lingering appearing as the

strong overestimations in all latitude bands for those months. From mid-2003 onwards the northern middle latitudes demonstrate a seasonally affected underestimation of around -1.5% whose amplitude also shows interannual variability whereas the southern middle and high latitudes over-estimate by around 1%. These findings agree very well with the validation analysis presented in Lerot *et al.*, 2009 even though the set of ground-based stations used there is not entirely the same as the one used in this study. In the middle left panel, the OMI_TOMS comparisons display the most homogeneous character of all instruments, with values that revolve around the zero difference [light green colours denote $\pm 0.5\%$], no spikes at the high latitudes and a persistent slight underestimation above $30^\circ - 60^\circ$ N. A different picture in the middle right panel for the OMI_DOAS comparisons presenting the largest overestimates by 2-4% compared to the ground truth for the high latitudes of both Hemispheres for the summer months, underestimates for the winter months and in general shows a variable structure with problems for high SZA cases as well. There appears also to be a general change in the comparative behaviour between ground and OMI_DOAS from year 2009 which shown more prominently in the time series plots presented below, a feature also observed for the Brewer comparisons [not shown here.] These findings agree well with Balis *et al.*, 2007b, who found the OMI-TOMS data product of high overall quality with no significant dependence on solar zenith angle or latitude. They showed as well that the OMI-DOAS data product systematically overestimated the total ozone values in the high latitudes of the Southern Hemisphere as well as a significant dependence on solar zenith angle. Finally, in the lower panel, GOME-2 exhibits a similar behavior as GOME and SCIAMACHY with the seasonal alteration between over- and under-estimating in the mid- to high latitudes. As shown also in Loyola *et al.*, 2011, the GOME-2 total ozone data obtained with GDP 4.4 slightly underestimates ground-based ozone by about 0.5% to 1% over the middle latitudes of the Northern Hemisphere and slightly overestimates by around 0.5% over the middle latitudes in the Southern Hemisphere.

3.1.2 Long term stability

Focusing on the time series of the differences as a mean over the NH stations, in Figure 2 the time evolution of the behaviour of each instrument is shown, alongside

the standard deviation of the daily points which are included in the monthly mean value. In the same order as before, the GOME differences is shown on the upper left with the obvious seasonality with peak-to-peak that revolves around the zero line. SCIAMACHY in the upper right, shows exactly the same patterns as GOME whereas the OMI_TOMS differences in the middle left show a more or less flat behaviour as the years progress revolving around a mean of -1% and no pronounced seasonal effects. OMI_DOAS follows the GOME and SCIAMACHY seasonality peaks until mid-2009 where the differences show a noticeable change from revolving around the zero line, to an over-estimation of between 1 and 2%. The GOME-2 picture is very similar to that of the rest of the total ozone retrievals using the DOAS algorithm apart from the fact that the seasonal peak-to-peak differences go from -2% to 0%. The seasonality effect apparent in all DOAS algorithms [i.e. in all datasets shown apart from OMI_TOMS] can, for most part, be attributed to the fact that the satellite algorithms account for differences in effective temperature whereas the ground-based instruments use a steady temperature for the ozone cross-sections. A secondary effect arises from the fact that the DOAS algorithms cannot fit well the observed features at high SZAs, which appears to be both a seasonal and geometrical issue. Hence, the common features seen in Figure 2 between all algorithms apart from the OMI_TOMS one.

3.1.3 Seasonal variability

Another way to discern possible issues is to investigate the contour representation of the differences on a seasonal basis, as presented in Figure 3. The layout of this Figure is as per the previous two figures. For the visualization needs of this Figure, a 1/12 per annum radius in time was used as well as a 10 degrees radius in latitude. In the upper left, the sixteen years of GOME create a smooth and clear image which shows that for the Northern middle and high latitudes the satellite overestimates by around 1% for the winter and autumn months and underestimates by the same amount for the summer months. The Southern tropical belt around 20-30° appears a zone of underestimation of 1% whereas for the higher Southern latitudes the situation again reverses with a progressive increase in over-estimation as the Antarctic

stations are encountered which reaches the levels of 3 to 4%. This GOME overestimation in polar latitudes during winter can probably be reduced by applying GDP 4.4 (no addition of ghost column for snow/ice conditions). In the upper right, for SCIAMACHY, a similar picture is detected for the Northern Hemisphere albeit with more pronounced features: in the northern spring and summer-time mid-latitudes the underestimation reaches -2% whereas the winter months also show a higher overestimations. For the Southern Hemisphere the pattern changes somewhat with the tropical zone overshooting by around 1% for all seasons and the middle and high latitudes overestimating up to 2-3% for the summer months and underestimating by 1% for the winter ones. In the middle left, the OMI_TOMS shows no seasonal or SZA issues with a constant underestimation of -0.5% for all months and latitudes. Conversely, in the middle right, OMI_DOAS shows a similar over- to under-estimating pattern as SCIAMACHY with different amplitudes with, interestingly, quite high values reaching 3-4% for the spring Antarctic region. Finally, in the lower left, GOME-2 shows a strong, around -2%, underestimation for the tropical belt around $\pm 30^\circ$ for all seasons which persists for higher Southern latitudes for the winter months. The Equatorial belt missing from this plot is due to the fact that the stations that belong to that belt i.e. Natal/Brazil, and Nairobi/ Kenya, provide data only up to 2004 and 2000 respectively.

3.2 Joint analysis and intercomparison

In the following section the comparison between GOME-2 and the other four satellite datasets is demonstrated through a composite Figure with eight panels which follow the same colour-coding format. Common days of data between the GOME record and each of the other instruments are used. In the first row of Figure 4 the percentage differences between satellite and co-located ground-based measurements have been grouped in 10° bins in latitude. In the second row, the differences have been grouped in 5° bins in solar zenith angle. In the third row, the cloud top pressure, CTP, dependency for each of the satellites is examined for bins of 50 mbars and in the fourth and final row, the cloud fraction in bins of 5%. These plots will be hereafter referred to as *Latitudinal variability*, *SZA variability*, *CTP variability* and *cloud fraction variability* respectively.

For the case of the *Latitudinal variability* and the Dobson comparisons [Fig. 4 top left]: a very close agreement for the NH and the tropics is observed in the behaviour of all five algorithms up to 50°N. From then and northwards OMI_DOAS presents a steady 1-2% over-estimation and GOME deviates at the Arctic stations with an over-estimation of more than 2%. The other three algorithms present a very similar picture. In the SH the comparisons are more diverse: GOME is under-estimating near -0.5% around the Equator and tropics and reaching -2% in the Antarctic. A very similar behaviour is observed by SCIAMACHY and OMI_TOMS. OMI_DOAS is constantly over-estimating by between 1 and 3% whereas GOME-2 shows a more constant behaviour around the zero line with the exception of the Antarctic where the differences reach 3%. For the Brewer comparisons [Fig. 4 top right], the situation is more homogeneous with no latitudinal dependency seen in any of the algorithms and a mean of -2% for GOME-2 and OMI_TOMS, a mean around -1% for GOME and SCIAMACHY and a mean of around zero for OMI_DOAS. This can be attributed to the fact that most Brewer stations fall around the middle latitude regions where the observational geometries and TOC ranges are far more favourable for constant, unbiased ground-based time series.

For the case of the *SZA variability* and the Dobson comparisons [second row left]: all algorithms show a similar picture with under-estimation for the low SZAs, over-estimation for the high SZAs and an obvious SZA dependency. What differentiates one instrument from the other is the magnitude of the dependency; GOME-2 and OMI_TOMS show the least amount of dependency with values ranging between -0.5% for the low and 0.5% for the high SZAs. OMI_DOAS seems to show the worse behaviour with a dependency that starts around 0% in difference and rises to 3% for SZAs higher than 80°. GOME and OMI_DOAS demonstrate a dependency in between these two extremes. For the Brewer comparisons [second row right], all algorithms show the same mild positive dependency in SZA with differences that start around -2% for the low SZA and rise to 0-2% for the high SZA apart from OMI_TOMS which is showing a more stable albeit negative dependency between -0.5% and -2% for all SZAs.

For the case of the *CTP variability* and the Dobson comparisons [third row left]: GOME-2 and OMI_TOMS show no dependency whatsoever and average values around 0%. SCIAMACHY shows a mild negative dependency, moving from a 2% over-estimation for low CTP to a -1% under-estimation for high CTP. GOME and

OMI_DOAS show the strongest dependencies with averages ranging between 4% and -2%. For the Brewer comparisons [third row right] a near-exact situation can be seen. A point to consider when examining the very low [< 200 mbars] and very high [> 800 mbars] CTP cases is the amount of data points that have been used in these graphs where for example for the case of the GOME measurements only 25 points make up the 200mbar bin, 60 the 300 mbar bin whereas there exist around 600 points for the 800 mbars bin.

For the case of the *cloud fraction variability* and the Dobson comparisons [bottom left]: no dependency and near-zero comparative values for all satellites apart from OMI_DOAS which shows a constant 1% positive offset for all cloud fractions. Near-similar picture for the Brewer comparisons with all algorithms around the -1% negative offset.

The comparisons as function of *CTP variability* and *cloud fraction variability* agree well with the results from *Antón and Loyola. 2011*.

In Tables 2 and 3, some statistics for the differences between satellite and ground-based measurements are given for distinct cases of latitudinal and SZA belts. In particular, three latitude bands are shown; the tropical zone, from 0° to 30° N, a middle latitude one, from 30° to 60° N and a polar one, from 60° to 90° N, as well as three SZA bands; low SZA, from 0° to 25° , middle SZA from 25° to 70° and high SZA from 70° to 90° . Table 2 shows the statistics for the Brewers and Table 3 the statistics for the Dobson instruments. The seemingly high standard deviation is due to the fact that raw daily measurements were used for these statistics and not already binned, and hence smoothed out, data.

Delving deeper into the comparative behaviour of the GOME-2 ozone record compared to the other algorithms, the direct scatter plots of the common data points between the satellites for the Dobson locations only are given in Figure 5. With GOME-2 always on the y-axis, the scatter against GOME is shown in the upper left, OMI_TOMS in the upper right, OMI_DOAS in the lower left and SCIAMACHY in the lower right. The different colours denote the SZA values associated with the total ozone columns. The first point to note is the excellent correlation level which in all cases has an R^2 of 0.98 and the data points in all cases fall nicely in the $y=x$ line. No obvious dependency of the total ozone columns on the SZA can be seen. For the high SZAs [orange colour] associated with the low Antarctic total ozone values the GOME-2 data under-estimates the GOME and OMI_DOAS data, while agreeing well

with OMI_TOMS and SCIAMACHY. The least deviation from the $y=x$ line can be seen in the comparisons with GOME [upper left] and the highest with OMI_DOAS [lower left.] In the OMI_DOAS comparisons numerous outlier points are observed irrespective of the associated SZA which points to the fact that the reason behind the disagreement is not due to the SZA treatment of the algorithms.

4 Summary and Conclusions

In the current paper we have assessed the stability and compatibility of five years of Global Ozone Monitoring Experiment-2/Metop-A total ozone columns against that of GOME/ERS, OMI/Aura and SCIAMACHY/Envisat through an extensive inter-comparison and validation exercise using as reference Brewer and Dobson ground-based measurements. Over the background truth of the ground-based measurements, the total ozone columns are inter-evaluated using a suite of established validation techniques and the main findings follow;

- i. On average, GOME-2 data underestimate GOME data by about 0.80%, and underestimate SCIAMACHY data by 0.40%. There is no seasonal dependence of the differences between GOME-2, GOME and SCIAMACHY. The latter is expected since the three data sets are based on similar algorithms (GDP 4.x). This underestimation of GOME-2 is within the uncertainty of reference ground-based data used in the comparisons.
- ii. On average, GOME-2 data show no dependency with respect to the cloud top pressure and the cloud fraction, with discrepancies between the -1% and zero levels. The same picture is observed for the OMI_TOMS data. The rest of the algorithms show varying levels of dependency, especially for high clouds, bearing in mind the sparsity of data at the low cloud top pressure values.
- iii. On average, GOME-2 data show very little SZA dependency and only for angles larger than 75° , maintaining a constant bias compared to the ground-based data. Marked is the improvement over the GOME dataset.
- iv. On average, GOME-2 data underestimate OMI_DOAS (collection 3) data by 1.28%, without any significant seasonal dependence of the differences between them. The lack of seasonality might be expected since both GDP 4.4 and

OMI_DOAS are DOAS-type algorithms and both consider the variability of the stratospheric temperatures in their retrievals.

v. On average, GOME-2 data and OMI_TOMS (collection 3) data have almost no bias (GOME-2 underestimates by 0.10%).

vi. Overall, GOME-2 total ozone data agree at the $\pm 1\%$ level with the standard ground-based measurements as well as other satellite instrument datasets and therefore are well suited to be incorporated and hence continue the satellite long-term global total ozone record needed among others for climate monitoring studies.

Acknowledgments

Development of the GOME-2 total ozone products and their validation has been funded by the O3M-SAF project with EUMETSAT and national contributions. The GOME/ERS-2 and SCIAMACHY/ENVISAT products were generated at DLR under the auspices of the D-PAF/D-PAC project funded by ESA, the OMI/Aura products were provided by NASA and KNMI. Ground-based data used in this work were taken from the World Ozone and Ultraviolet Data Centre (WOUDC, <http://www.woudc.org>).

References

- Antón, M., M. E. Koukouli, M. Kroon, R. D. McPeters, G. J. Labow, D. Balis, and A. Serrano, Global validation of empirically corrected EP-Total Ozone Mapping Spectrometer (TOMS) total ozone columns using Brewer and Dobson ground-based measurements, *J. Geophys. Res.*, 115, D19305, doi:10.1029/2010JD014178, 2010.
- Antón, M., and D. Loyola, Influence of cloud properties on satellite total ozone observations, *J. Geophys. Res.*, 116, D03208, doi:10.1029/2010JD014780, 2011.

558 Balis, D., M. Kroon, M. E. Koukouli, E. J. Brinksma, G. Labow, J. P. Veefkind, and
559 R. D. McPeters, Validation of Ozone Monitoring Instrument total ozone column
560 measurements using Brewer and Dobson spectrophotometer ground-based
561 observations, *J. Geophys. Res.*, 112, D24S46, doi:10.1029/2007JD008796,
562 2007a.

563 Balis, D., J-C. Lambert , Van Roozendaal M., D. Loyola, R. Spurr, Y. Livschitz, P.
564 Valks, T. Ruppert, P. Gerard, J Granville and V. Amiridis, Ten years of
565 GOME/ERS2 total ozone data—The new GOME data processor (GDP) version
566 4: 2. Ground-based validation and comparisons with TOMS V7/V8, *J. Geophys.*
567 *Res.*, 112, D07307, doi:10.1029/2005JD006376, 2007b.

568 Bass A.M., and R.J. Paur, The ultraviolet cross-sections of ozone: I. The
569 measurements in Atmospheric ozone (Ed. C.S. Zerefos and A. Ghazi), Reidel,
570 Dordrecht, Boston, Lancaster, pp. 606-610, 1985.

571 Bhartia, P.K., C.G. Wellemeyer, S.L. Taylor, N. Nath, and A. Gopalan, Solar
572 backscatter ultraviolet (SBUV) version 8 profile algorithm, in Proceedings of
573 the Quadrennial Ozone Symposium, Kos, Greece, 1-8 June 2004, 295-296,
574 2004.

575 Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel,
576 A., Hartmann, M., Bovensmann, H., Frerick, J., and Burrows, J. P.:
577 Measurements of molecular absorption spectra with the SCIAMACHY pre-
578 flight model: instrument characterization and reference data for atmospheric
579 remote sensing in the 230–2380 nm region, *J. Photochem. Photobiol. A.*, 157,
580 167–184, 2003.

581 Bovensmann, H., Burrows, J., Buchwitz, M., J. Frerick, S. Noël, V. V. Rozanov, K.
582 V. Chance, A. P. H. Goede, SCIAMACHY: Mission objectives and
583 measurement modes, *J. Atmos. Sci.*, 56, 2, 127–150, 1999.

584 Burrows, J., M. Weber, M. Buchwitz, V. Rozanov, A. Ladstätter-Weissenmayer, A.
585 Richter, R. DeBeek, R. Hoogen, K. Bramstedt, K.-U. Eichmann, M. Eisinger
586 and D. Perner: The Global Ozone Monitoring Experiment (GOME): Mission
587 concept and first scientific results, *J. Atmos. Sci.*, 56, 151–175, 1999.

588 Burrows, J., A. Richter, A. Dehn, B. Deters, S. Himmelmann, S. Voigt, and J. Orphal,
 589 Atmospheric remote sensing reference data from GOME: Part 2. Temperature-
 590 dependent absorption cross-sections of O₃ in the 231 – 794 nm range, *J. Quant.*
 591 *Spectrosc. Radiat. Transfer*, 61, 509– 517, 1999b
 592 Lerot C, Van Roozendaal, M., van Geffen, J., et al., Six years of total ozone column
 593 measurements from SCIAMACHY nadir observations, *Atmos. Meas. Tech.*, 2,
 594 87-98, 2009.
 595 Levelt P.F., G.H.J. van den Oord, M.R. Dobber, A. Malkki, H. Visser, J. de Vries, P.
 596 Stammes, J.O.V. Lundell and H. Saari, “The Ozone Monitoring Instrument”,
 597 *IEEE Trans. Geosc. Rem. Sens.*, 44 (5), 1093-1101, 2006.
 598 Loyola, D. G., M. E. Koukouli, P. Valks, D. S. Balis, N. Hao, M. Van Roozendaal, R.
 599 J. D. Spurr, W. Zimmer, S. Kiemle, C. Lerot and J-C. Lambert, The GOME-2
 600 total column ozone product: Retrieval algorithm and ground-based validation, *J.*
 601 *Geophys. Res.*, 116, D07302, doi:10.1029/2010JD014675, 2011
 602 Munro, R., M. Eisinger, C. Anderson, J. Callies, E. Corpaccioli, R. Lang, A.
 603 Lefebvre, Y. Livschitz, and A. Perez Albinana: GOME-2 on METOP: From in-
 604 orbit verification to routine operations, paper presented at EUMETSAT
 605 Meteorological Satellite Conference, Helsinki, Finland, 12–16 June 2006.
 606 Schoeberl M., A. Douglass, E. Hilsenrath, P.K. Bhartia, R. Beer, J. W. Waters, M.
 607 Gunson, L. Froidevaux, J. Gille, J. Barnett, P. Levelt, and P. DeCola, Overview
 608 of the EOS Aura Mission, *IEEE Trans. Geo. Rem. Sens.*, 44, 1066–1074,
 609 doi:10.1109/TGRS.2005.861950, 2006.
 610 Valks P., D. Loyola, N. Hao, M.Rix and S. Slijkhuis, Algorithm Theoretical Basis
 611 Document for GOME-2 Total Column Products of Ozone, NO₂, tropospheric
 612 NO₂, BrO, SO₂, H₂O, HCHO, OClO and Cloud Properties, DLR/GOME-
 613 2/ATBD/01, Iss/Rev: 2/E, [http://atmos.caf.dlr.de/gome2/docs/DLR_GOME-](http://atmos.caf.dlr.de/gome2/docs/DLR_GOME-2_ATBD.pdf)
 614 [2_ATBD.pdf](http://atmos.caf.dlr.de/gome2/docs/DLR_GOME-2_ATBD.pdf), accessed 15/12/2011, 2011
 615 Van Roozendaal, M., P.Peters, P.C.Simon, H.K.Roscoe, A.E.Jones, L.Bartlett,
 616 G.Vaughan, F.Goutail, J-P.Pommereau, E.Kyro, C.Whalstrom, G.Braathen and
 617 P.C.Simon, Validation of ground-based visible measurements of total ozone by

618 comparison with Dobson and Brewer spectrophotometers, *J. Atmos. Chem.*, 29,
619 55-83, 1998.

620 Van Roozendaal M., D. Loyola, R. Spurr, D. Balis, J-C. Lambert, Y. Livschitz, P.
621 Valks, T. Ruppert, P. Kenter, C. Fayt and C. Zehner (2006), Ten years of
622 GOME/ERS-2 total ozone data—The new GOME data processor (GDP) version
623 4: 1. Algorithm description, *J. Geophys. Res.*, 111, D14311,
624 doi:10.1029/2005JD006375

625 Veefkind, J.P., J.F. de Haan, E.J. Brinksma, M. Kroon and P.F. Levelt, Total Ozone
626 from the Ozone Monitoring Instrument (OMI) Using the DOAS technique,
627 *IEEE Trans. Geo. Rem. Sens.*, 44, 5, 1239-1244,
628 doi:10.1109/TGRS.2006.871204, 2006.

629

TABLES

Table 1. Main characteristics of the GOME-2/MetOp GOME/ERS-2, SCIAMACHY and OMI instruments relevant to the total ozone column products.

	GOME/ERS-2	SCIAMACHY	OMI	GOME-2/MetOp
Principle	UV/VIS grating spectrometer	UV/VIS/NIR grating spectrometer	UV/VIS grating spectrometer	UV/VIS grating spectrometer
Detectors	Reticon linear diode array	Reticon linear diode array	2-dimensional CCD	Reticon linear diode array
Spectral resolution	0.20 nm	0.26 nm	0.45 nm	0.26 nm
Spatial resolution (default)	320 x 40 km²	60 x 30 km²	Up to 13 x 24 km²	80 x 40 km²
Swath width	960 km	960 km	2600 km	1920 km
Dataset span	1995-2011	2002-2011	2004-2011	2007-2011
Eq. crossing time	10:30 LT	10:00 LT	13:30 LT	09:30 LT
Level-0-to-1b alg.	GDP L01 4.0	IPF 6.03	OML1BRUG (v003)	GOME-2 PPF 4.x
Level-1-to-2 alg.	GDP 4.1	SDOAS (SGP 5.0)	OMDOAO3 v1.0.5 & OMT03 v1.1.0	GDP 4.4

* In addition to the parameters listed here, the differential signal-to-noise characteristics of the instruments can have an impact on the total ozone column retrieval as well.

Table 2. The average differences between satellite and ground-based instrument for the five datasets shown in this paper. The mean and standard deviation for the latitudinal and the SZA variability are given for the Brewer instruments for low, middle and high latitudes and angles respectively.

	BREWER	
	LATITUDINAL VARIABILITY	SZA VARIABILITY

	Tropical [0-30°N]	Mid- Latitudes [30-60°N]	Polar [60-90°N]	Low [0-25°]	Middle [25-70°]	High [70-90°]
GOME-2	-2.0±2.4%	-1.9±2.4%	-1.5±3.0%	-2.1±2.4%	-1.8±2.4%	-0.4±3.2%
GOME	-1.7±2.6%	-1.0±2.8%	-0.02±3.4%	-1.8±2.7%	-0.6±2.8%	1.1±3.5%
OMI_TOMS	-0.8±2.2%	-1.0±2.1%	-1.5±2.4%	-0.8±2.1%	-1.0±2.4%	-1.1±4.3%
OMI_DOAS	-0.2±3.2%	-0.02±2.9%	-0.5±10.8%	0.03±3.3%	0.3±3.0%	1.1±9.7%
SCIAMACHY	-1.3±2.8%	-1.0±2.6%	-0.6±3.0%	-1.6±2.7%	-0.8±2.6%	0.3±3.5%

Table 3. The average differences between satellite and ground-based instrument for the five datasets shown in this paper. The mean and standard deviation for the latitudinal and the SZA variability are given for the Dobson instruments for low, middle and high latitudes and angles respectively.

	DOBSON					
	LATITUDINAL VARIABILITY			SZA VARIABILITY		
	Tropical [0-30°N]	Mid- Latitudes [30-60°N]	Polar [60-90°N]	Low [0-25°]	Middle [25-70°]	High [70-90°]
GOME-2	-1.3± 3.2%	-0.9±2.7%	-0.6±3.2%	-1.7±3.2%	-0.8±3.0%	0.7±5.8%
GOME	-1.3± 15.1%	-0.4±10.3%	0.6±4.0%	-1.4±2.2%	0.2±3.3%	2.0±6.3%
OMI_TOMS	-0.9±3.3%	-0.9±3.3%	-0.5±3.0%	-0.9±3.3%	-0.5±3.0%	0.06±3.8%
OMI_DOAS	-0.2± 4.0%	0.2±4.0%	1.3±4.7%	0.3±4.1%	0.9±3.9%	2.8±5.9%
SCIAMACHY	-1.1±3.6%	-0.5±3.1%	-0.4±3.8%	-1.0±3.6%	0.05±3.6%	1.0±5.8%

FIGURE CAPTIONS

Figure 1. Contour representation of the difference between satellite and Dobson ground-based measurements for the five datasets discussed in this paper as a function of latitude and time. **Top left:** GOME. **Top right:** SCIAMACHY. **Middle left:** OMI_TOMS. **Middle right:** OMI_DOAS. **Bottom:** GOME-2.

Figure 2. Time series of the differences between satellite and Dobson ground-based measurements for the five datasets for the Northern Hemisphere stations only. **Top left:** GOME. **Top right:** SCIAMACHY. **Middle left:** OMI_TOMS. **Middle right:** OMI_DOAS. **Bottom:** GOME-2.

Figure 3. Contour representation of the percentage differences between satellite and Dobson ground-based total ozone measurements depicted as a function of solar zenith angle and season. **Top left:** GOME. **Top right:** SCIAMACHY. **Middle left:** OMI_TOMS. **Middle right:** OMI_DOAS. **Bottom:** GOME-2.

Figure 4. Comparisons of GOME-2 [purple], GOME [dark blue], OMI-TOMS [light blue], OMI-DOAS [light green] and SCIAMACHY [red] total ozone against the same ground-based measurements. Left column are the Brewer comparisons and the right column the Dobson comparisons. From top to bottom the dependence of the percentage differences as function of latitude, solar zenith angle, cloud top pressure and cloud fraction.

Figure 5. Direct scatter plot comparisons of the satellite total column at the overpass sites for the Dobson ground-based stations. Top row: GOME-2 [on the y-axis] to GOME [left plot] and OMI_TOMS [right plot.] Bottom row: GOME-2 [on the y-axis] to OMI_DOAS [left plot] and SCIAMACHY [right plot.] The points are colour-coded according to the SZA related to that measurement. The r-squared and y-intercept of the scatter line [red] are also given.

FIGURES

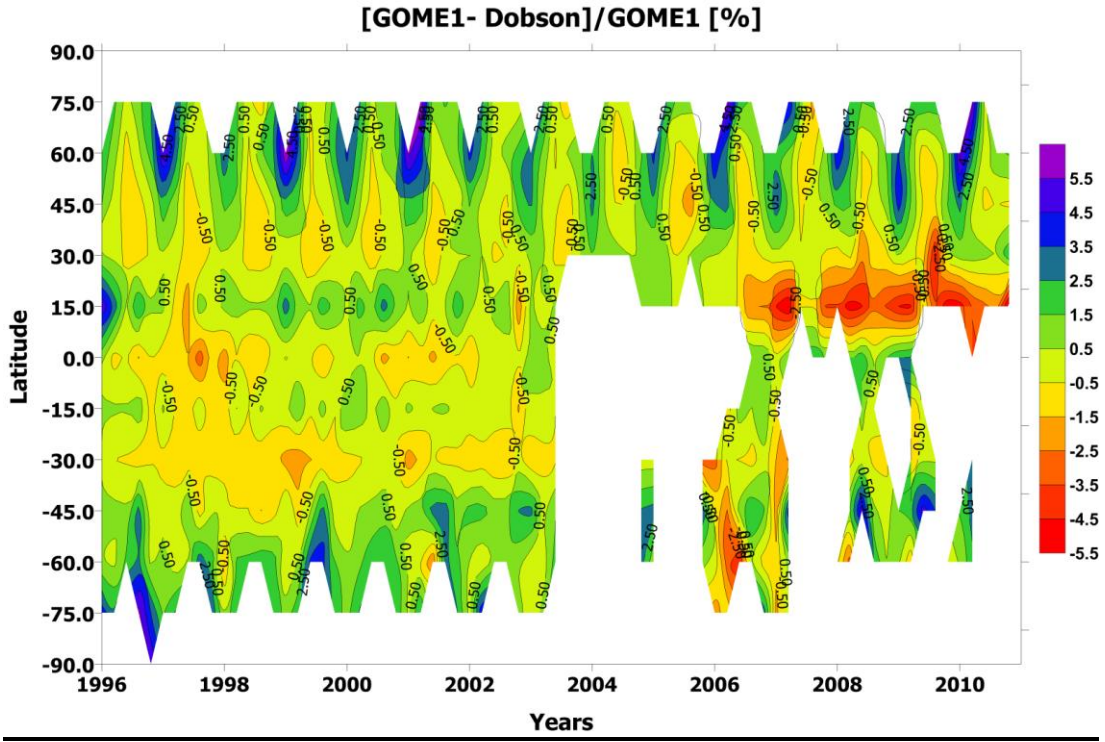


Figure 1 Top Left

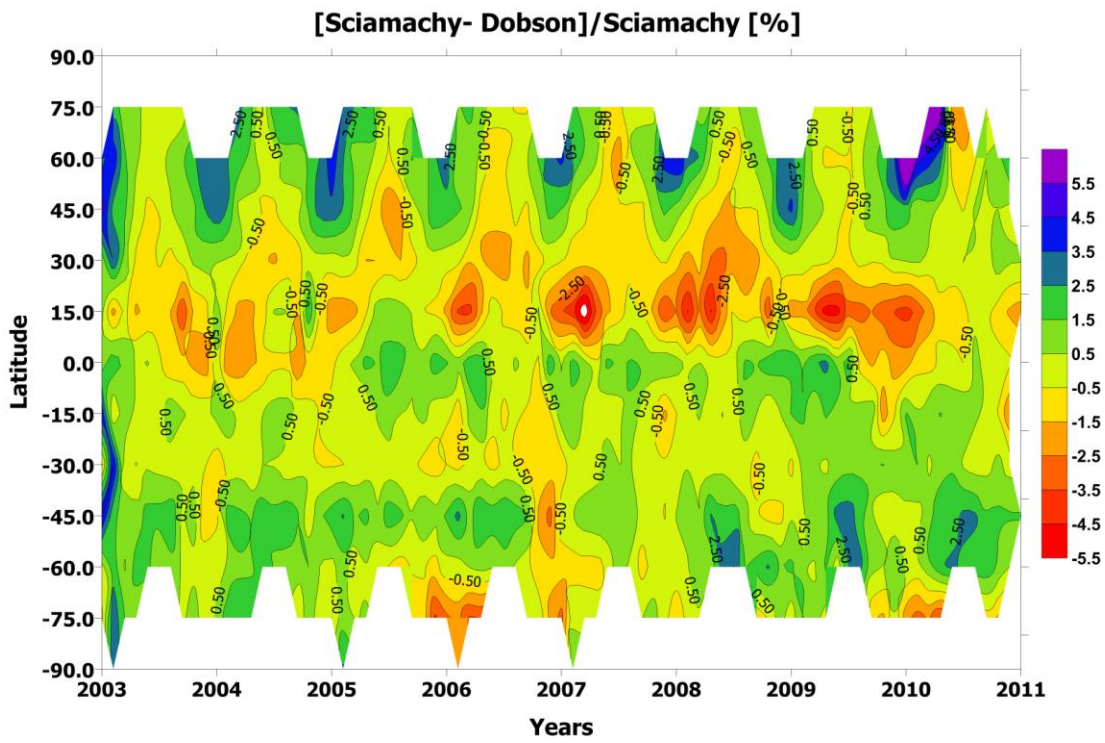


Figure 1 Top Right

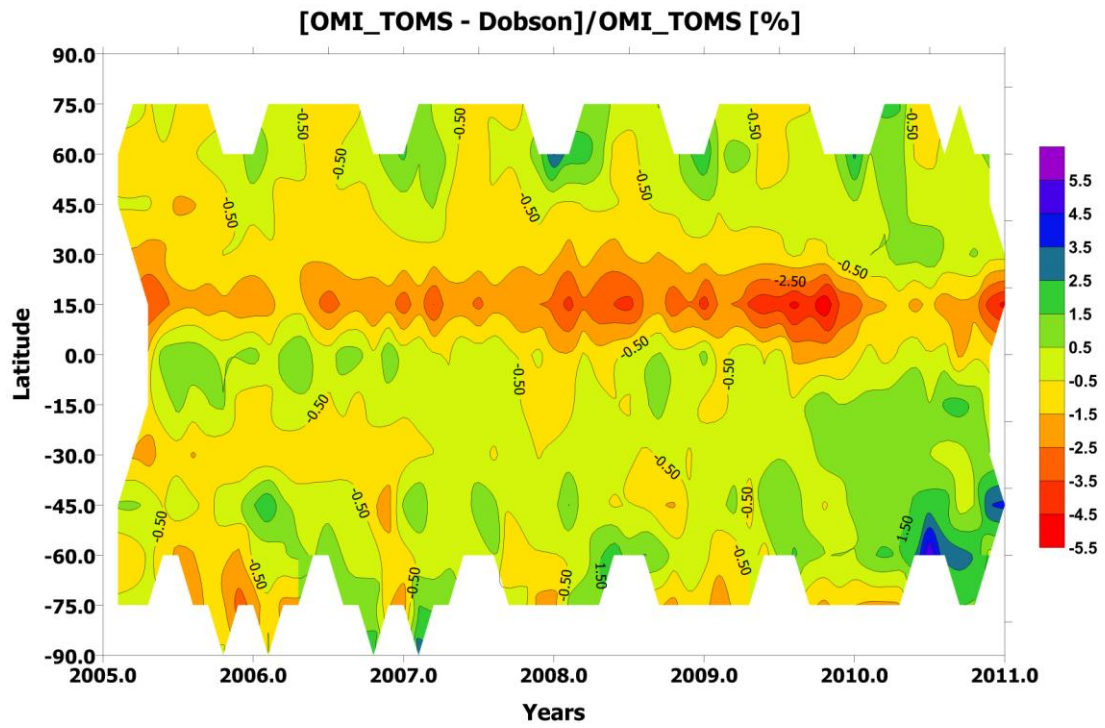


Figure 1 Middle Left

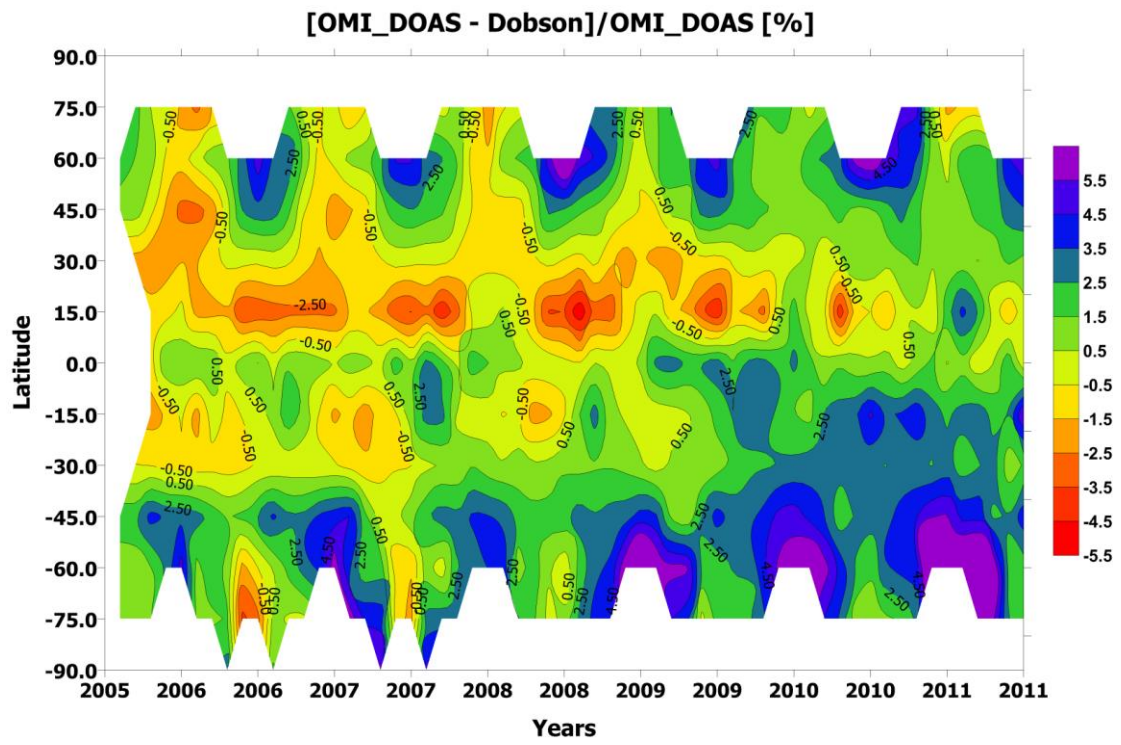


Figure 1 Middle Right

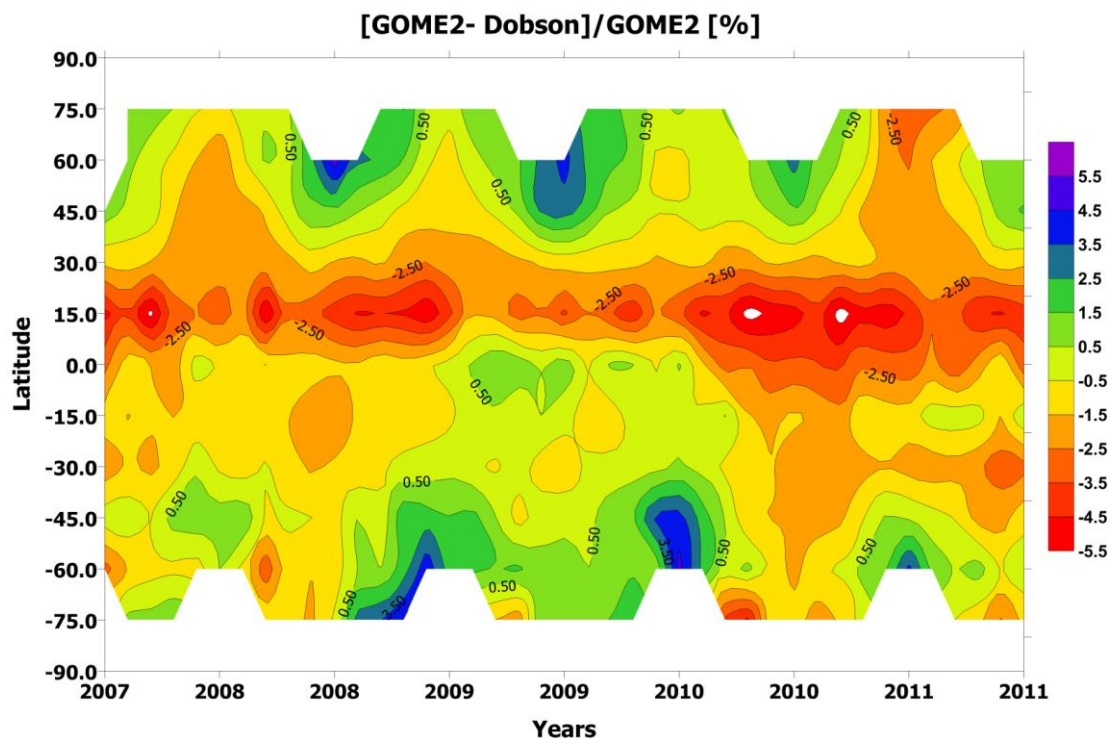


Figure 1 Bottom.

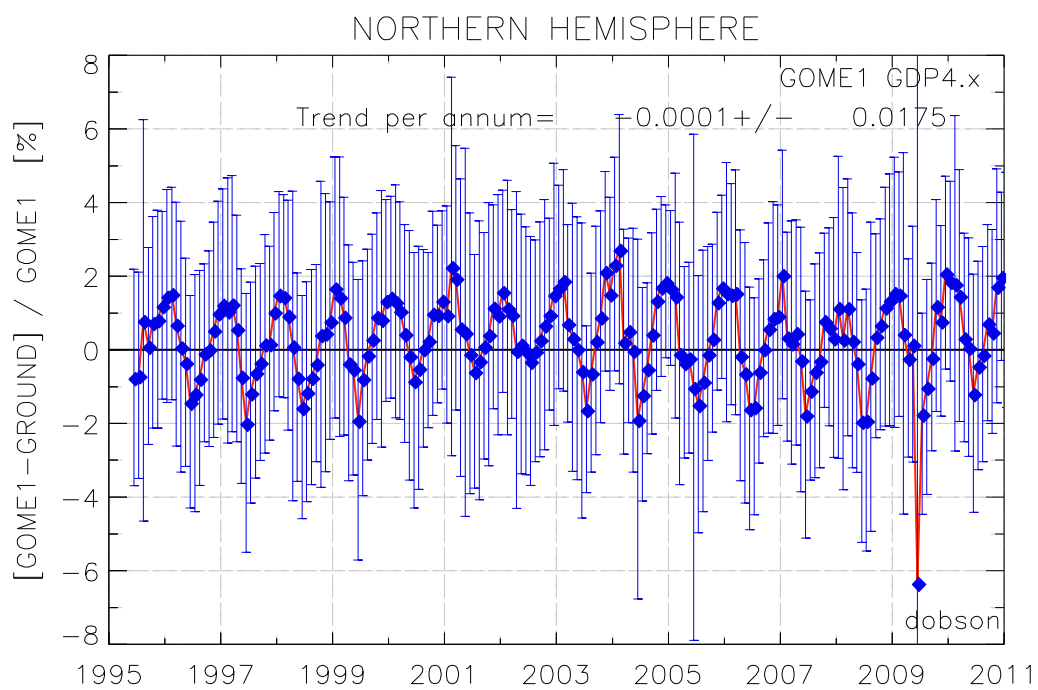
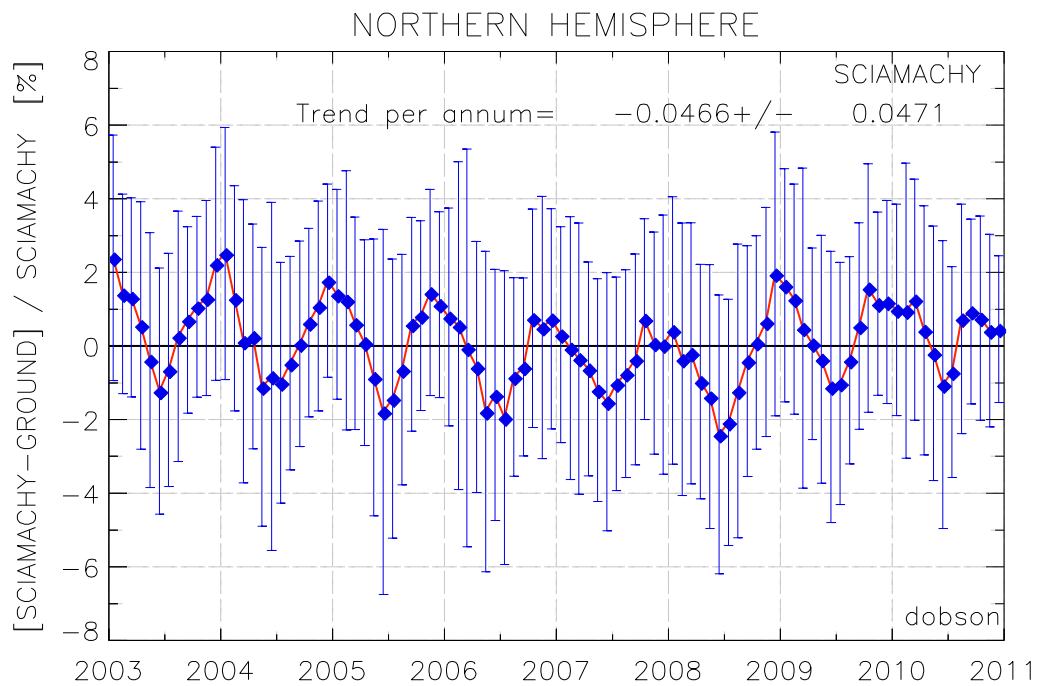


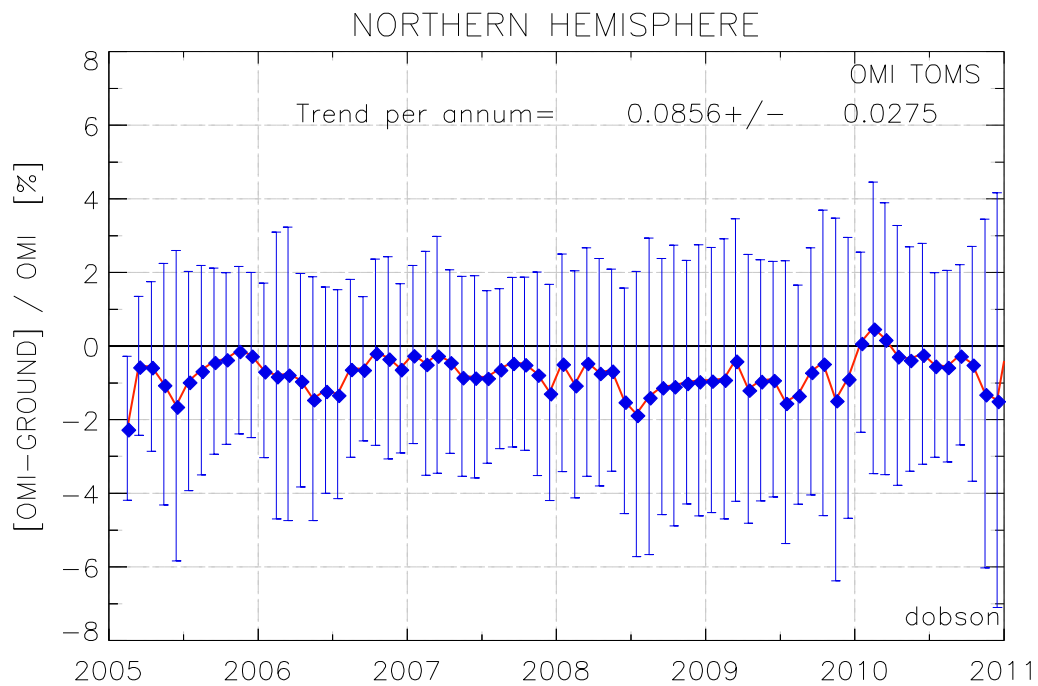
Figure 2 Upper Left



690
691

Created on Fri Jun 08 09:34:43 2012
Figure 2 Upper Right

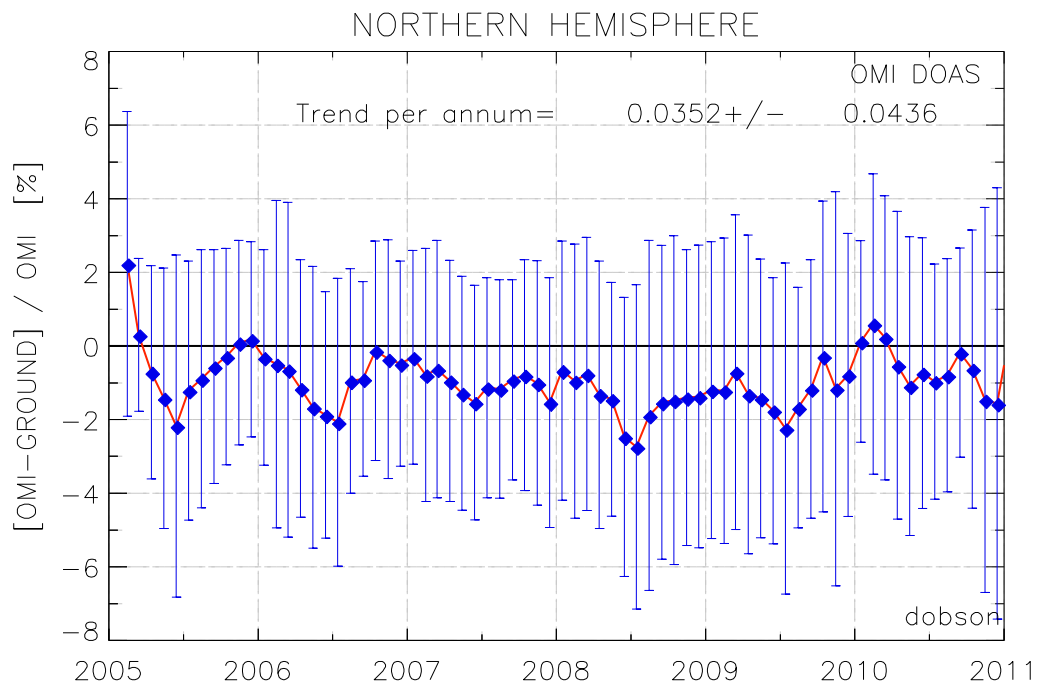
Aristotle University of Thessaloniki



692
693

Created on Fri Jun 08 09:35:45 2012
Figure 2 Middle Left

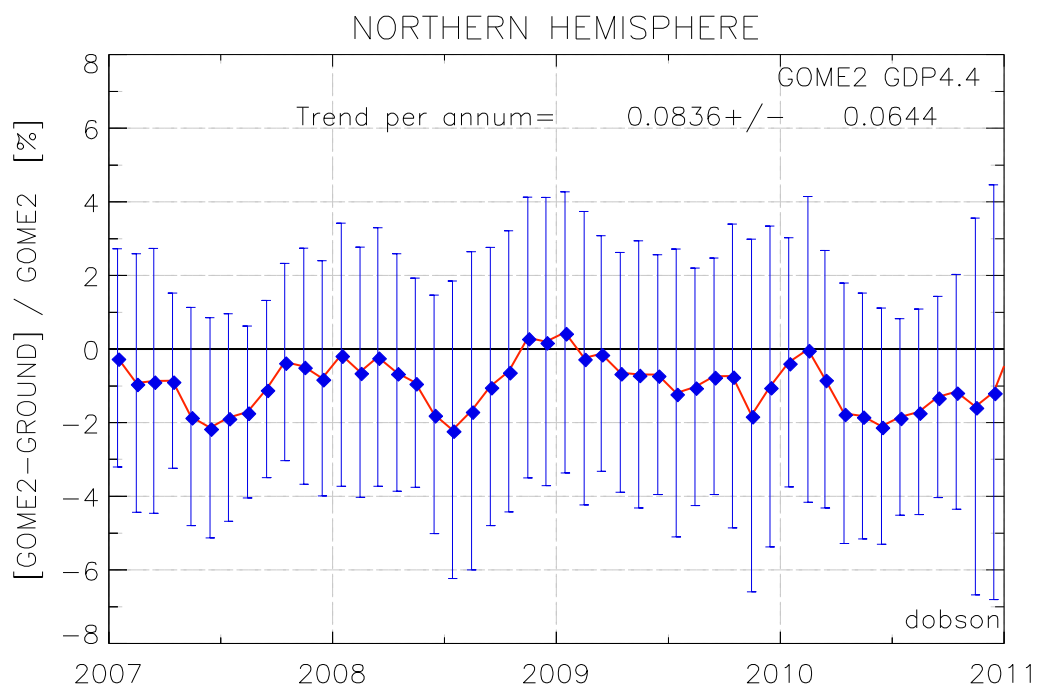
Aristotle University of Thessaloniki



694
695

Created on Fri Jun 08 09:36:05 2012
Figure 2 Middle Right

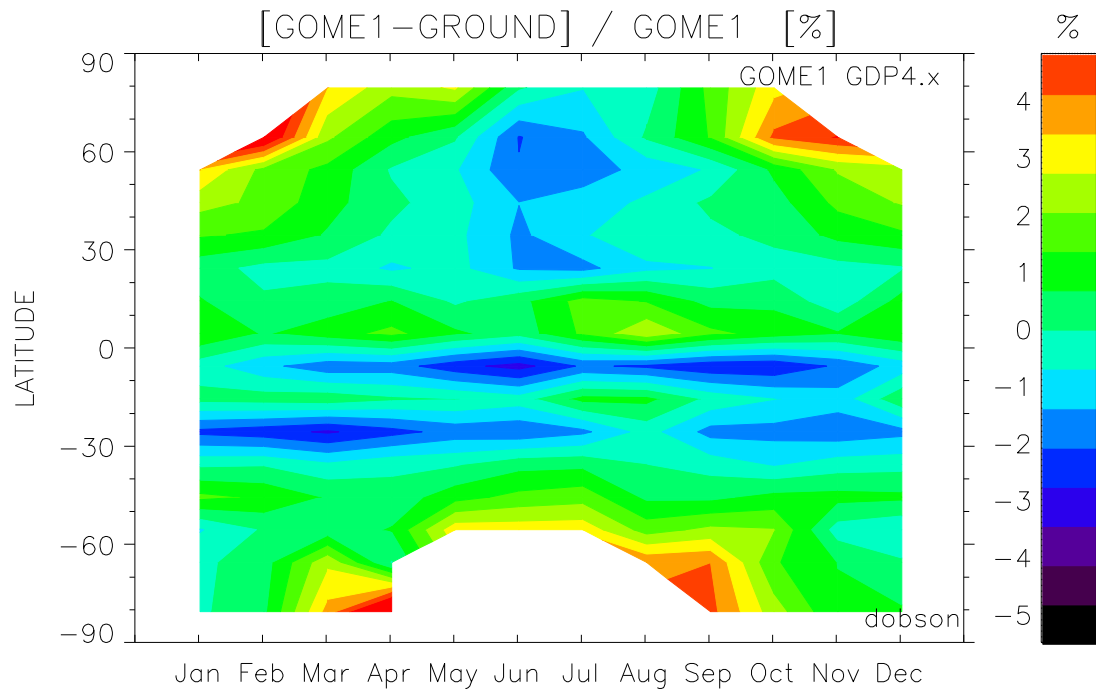
Aristotle University of Thessaloniki



696
697

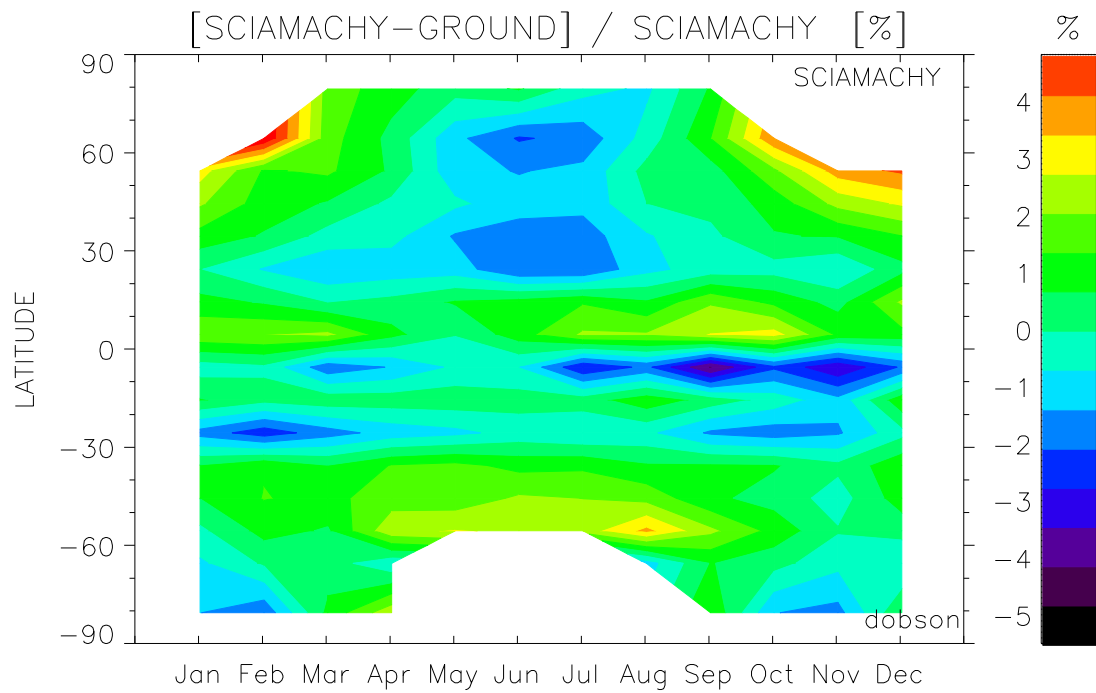
Created on Fri Jun 08 09:43:58 2012
Figure 2 Bottom

Aristotle University of Thessaloniki



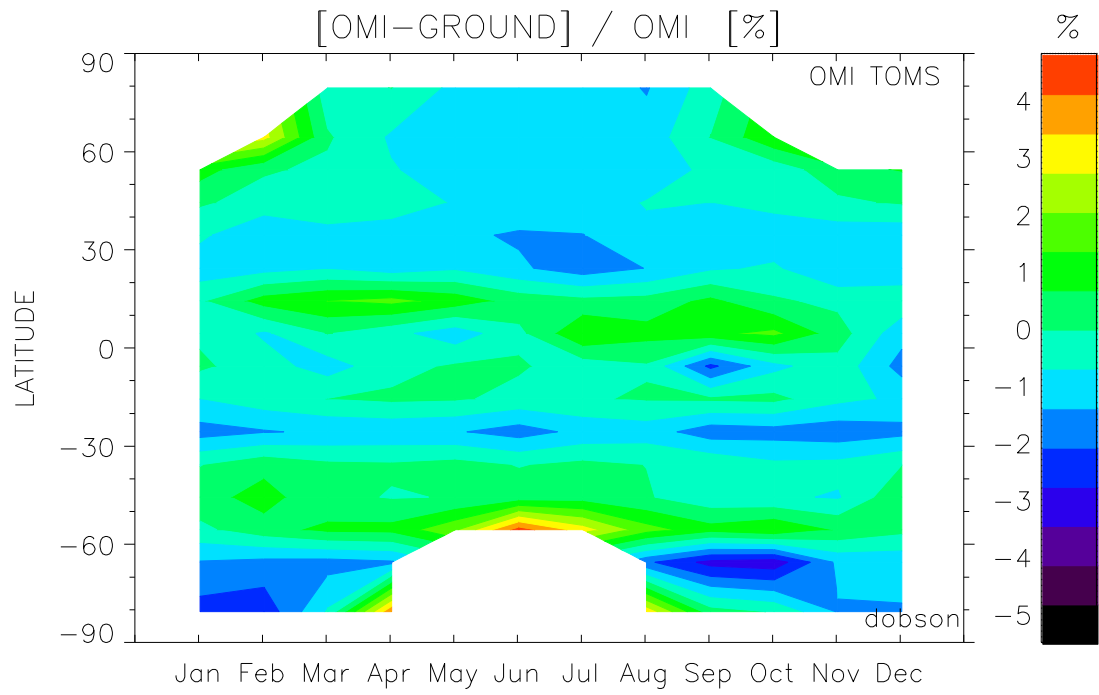
698
699

Created on Fri Jun 08 09:17:53 2012
Figure 3 Upper Left



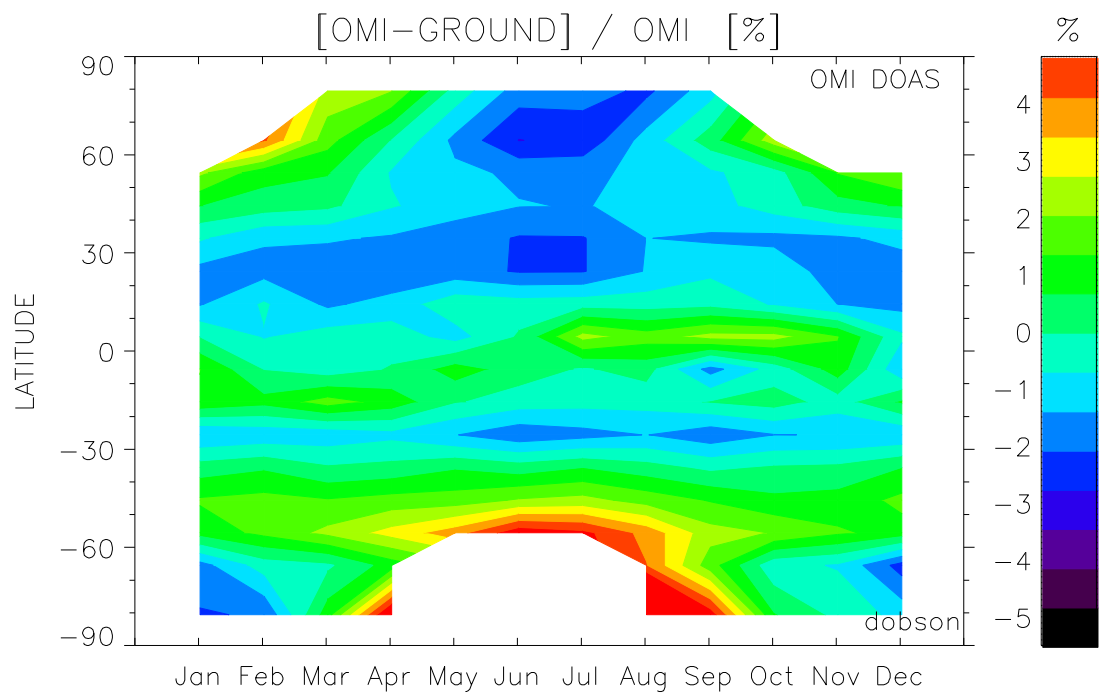
700
701

Created on Fri Jun 08 09:22:36 2012
Figure 3 Upper Right



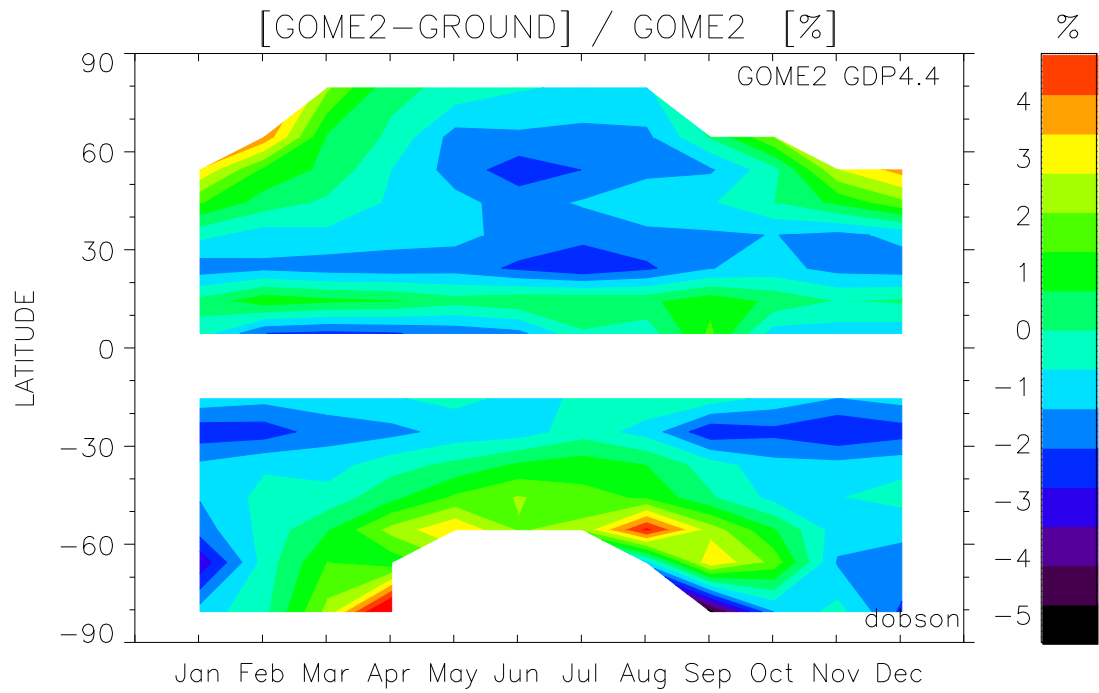
702
703

Created on Fri Jun 08 09:20:44 2012
Figure 3 Middle Left.



704
705

Created on Fri Jun 08 09:21:26 2012
Figure 3 Middle Right



Created on Fri Jun 08 09:03:08 2012
Figure 3 Bottom

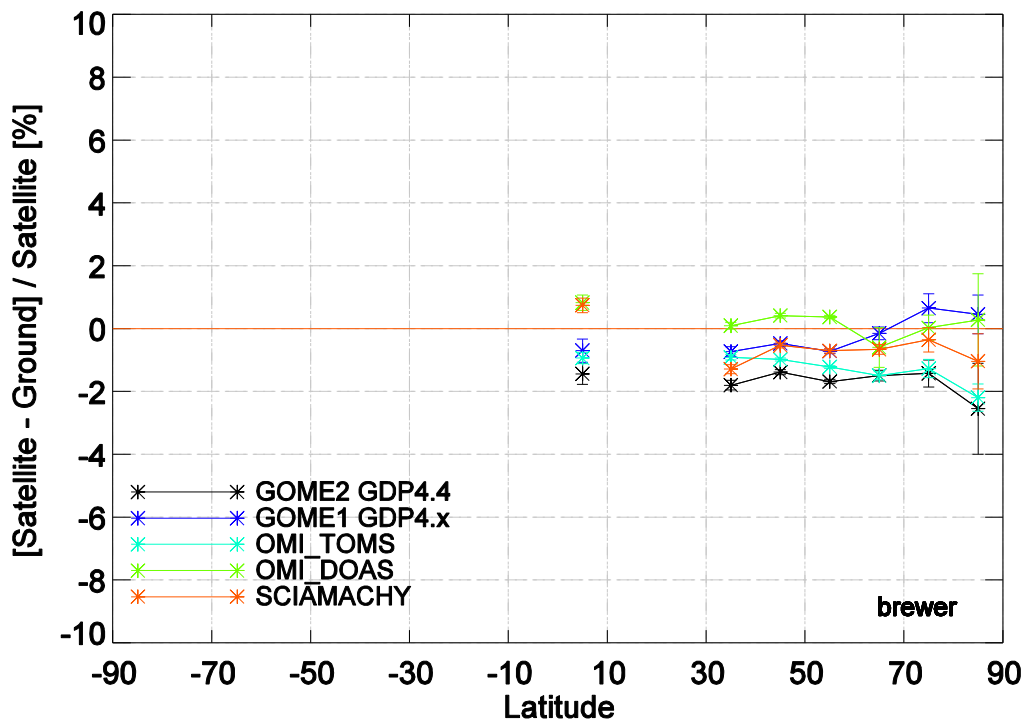
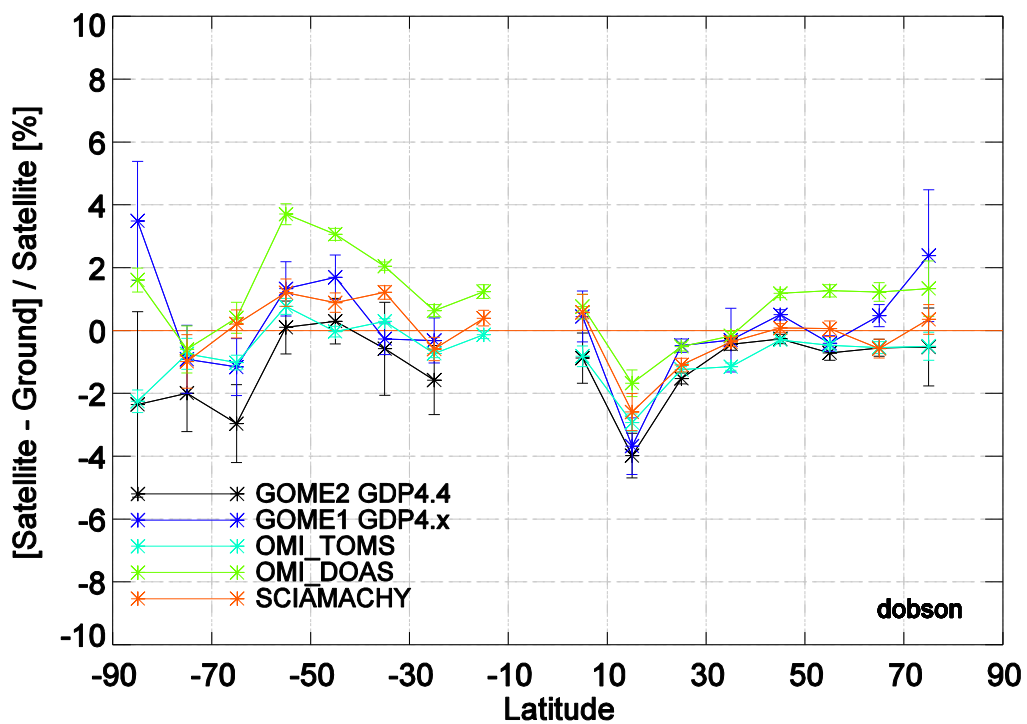
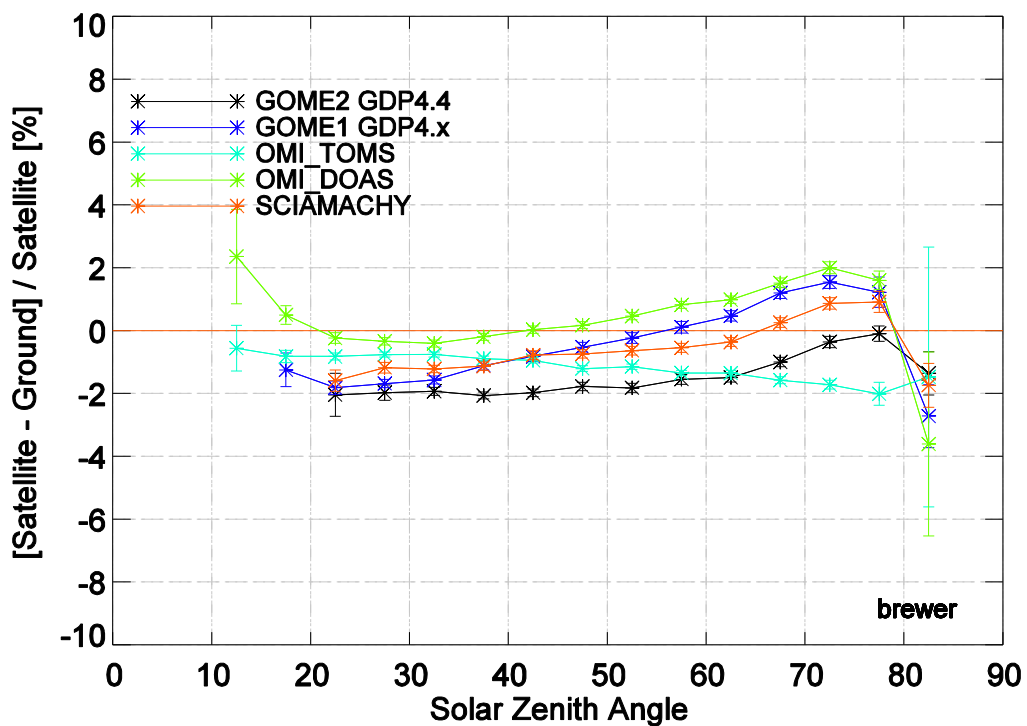


Figure 4 First Row, Left.



710
711

Figure 4 First Row, Right.



712
713

Figure 4 Second Row, Left.

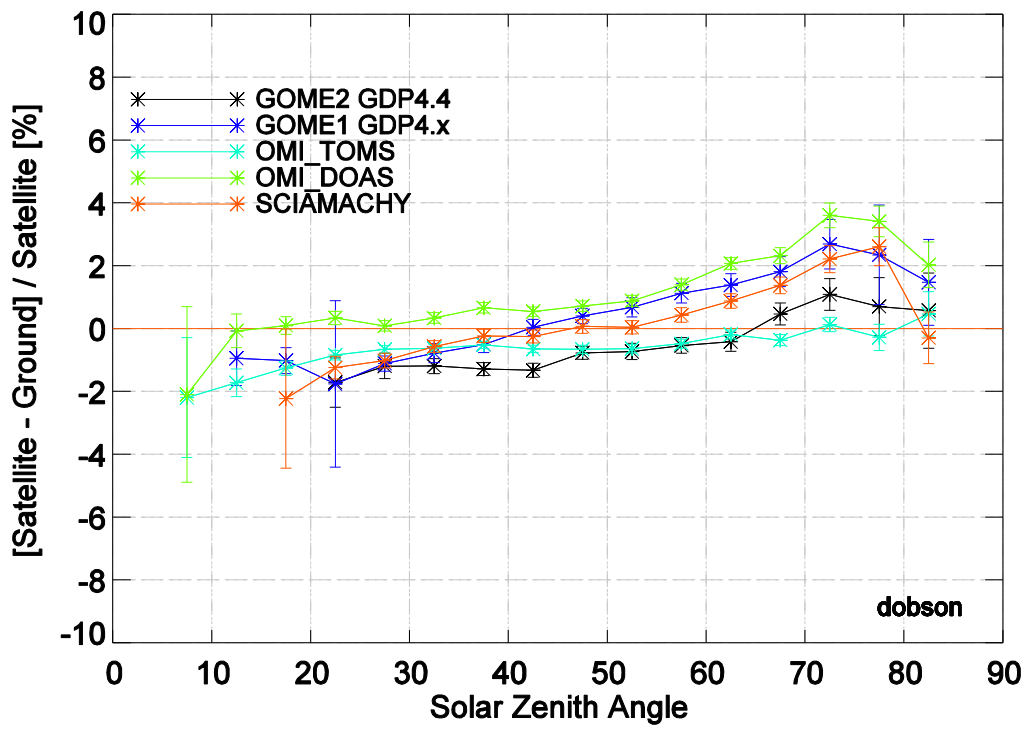


Figure 4 Second Row, Right.

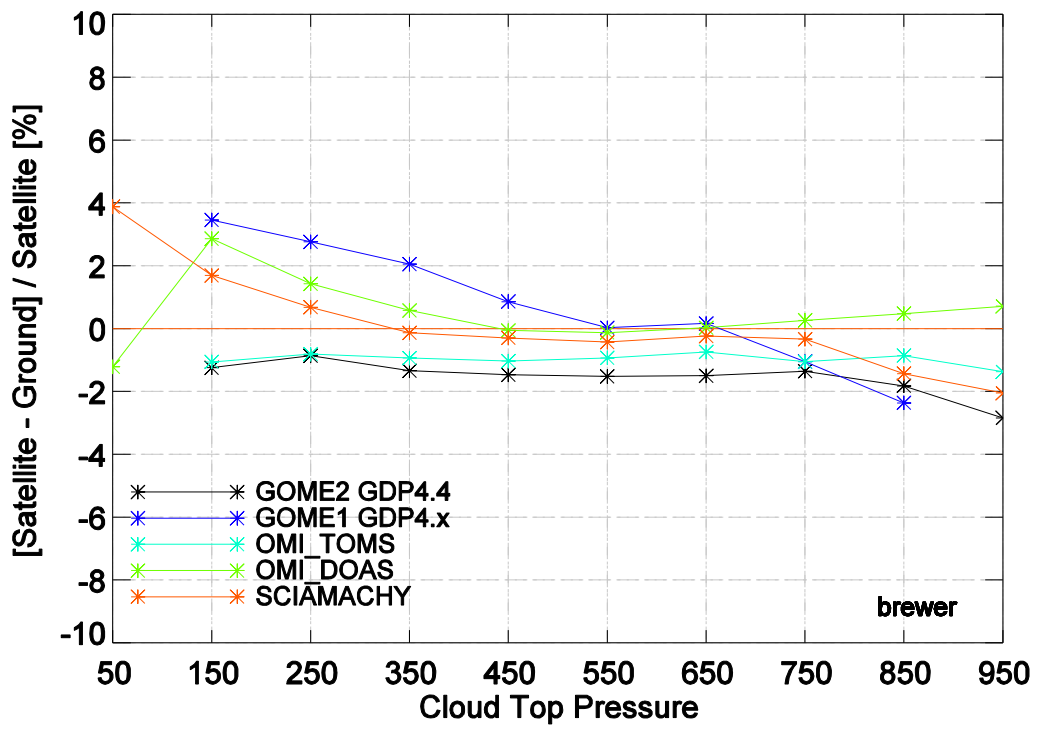


Figure 4 Third Row, Left.

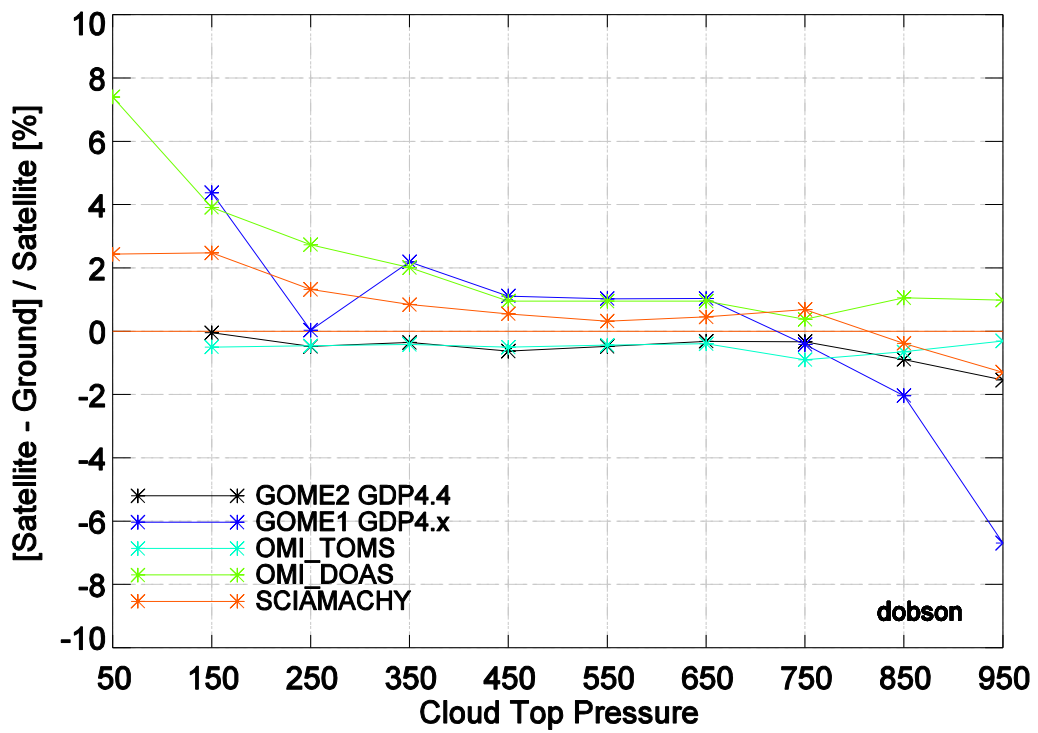


Figure 4 Third Row, Right.

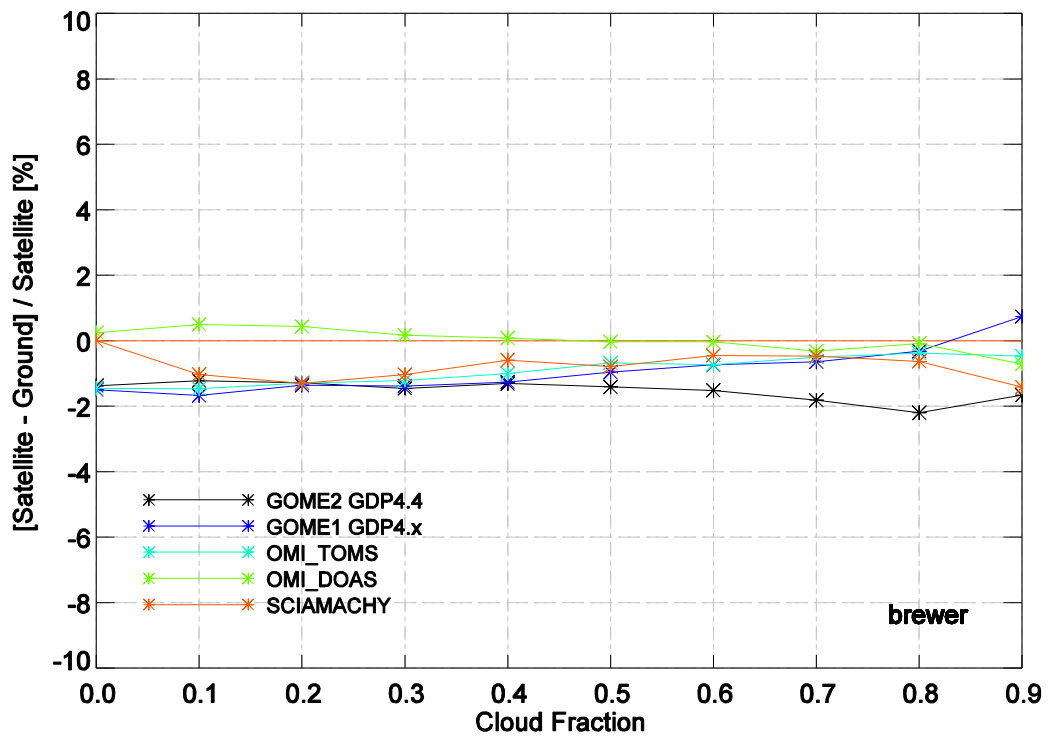
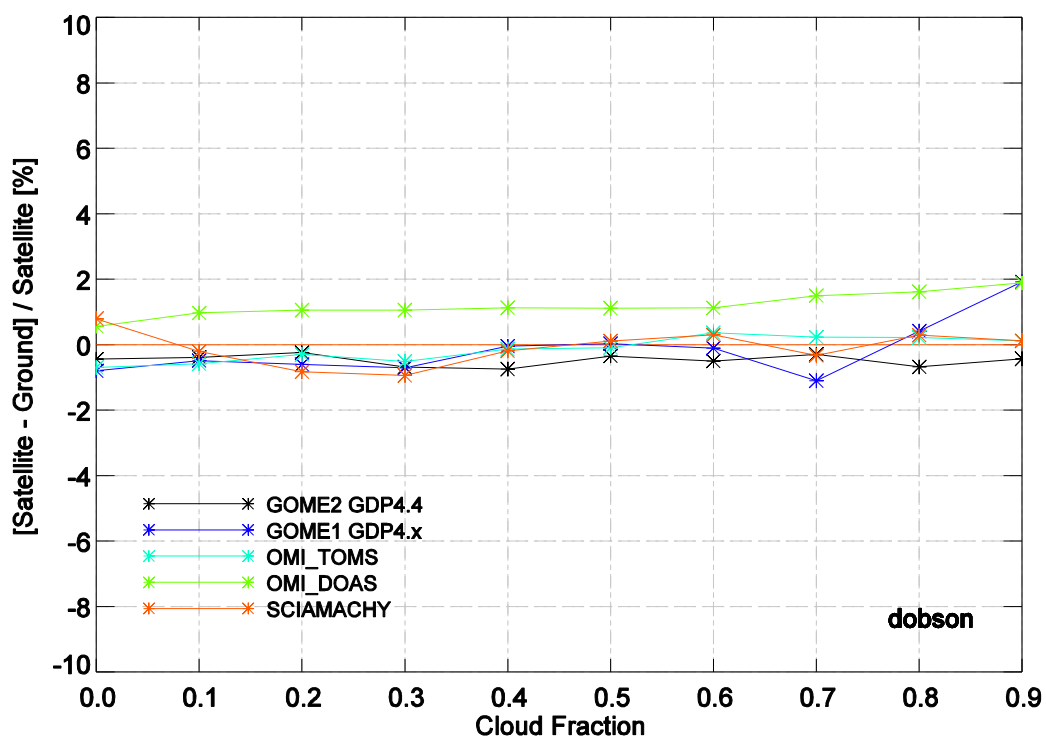


Figure 4 Fourth Row, Left.

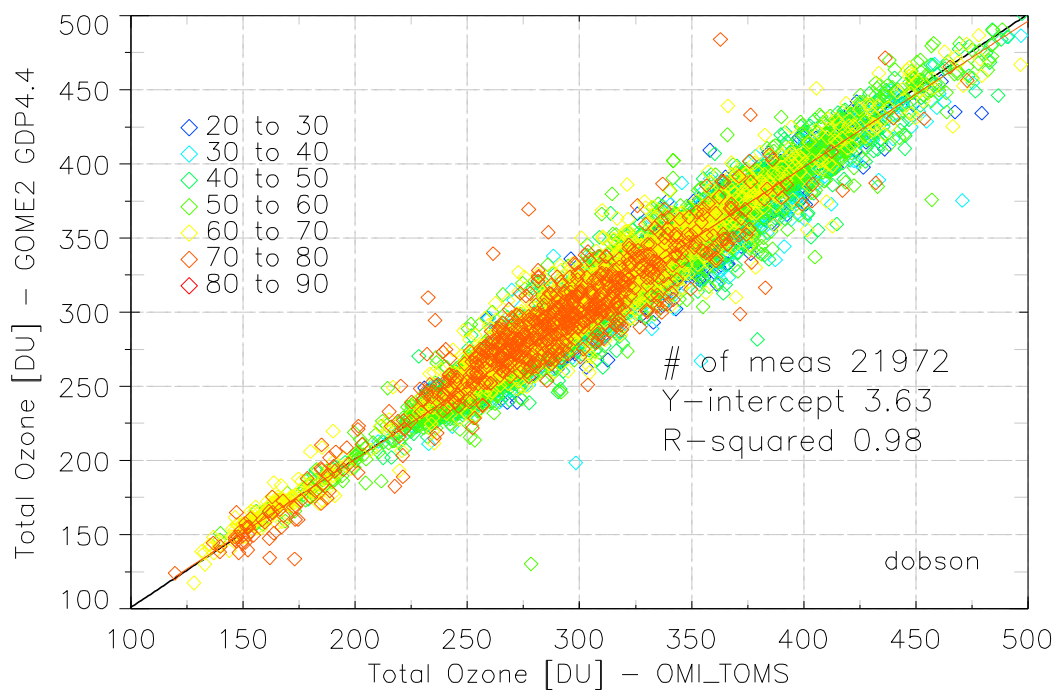
722



723

724

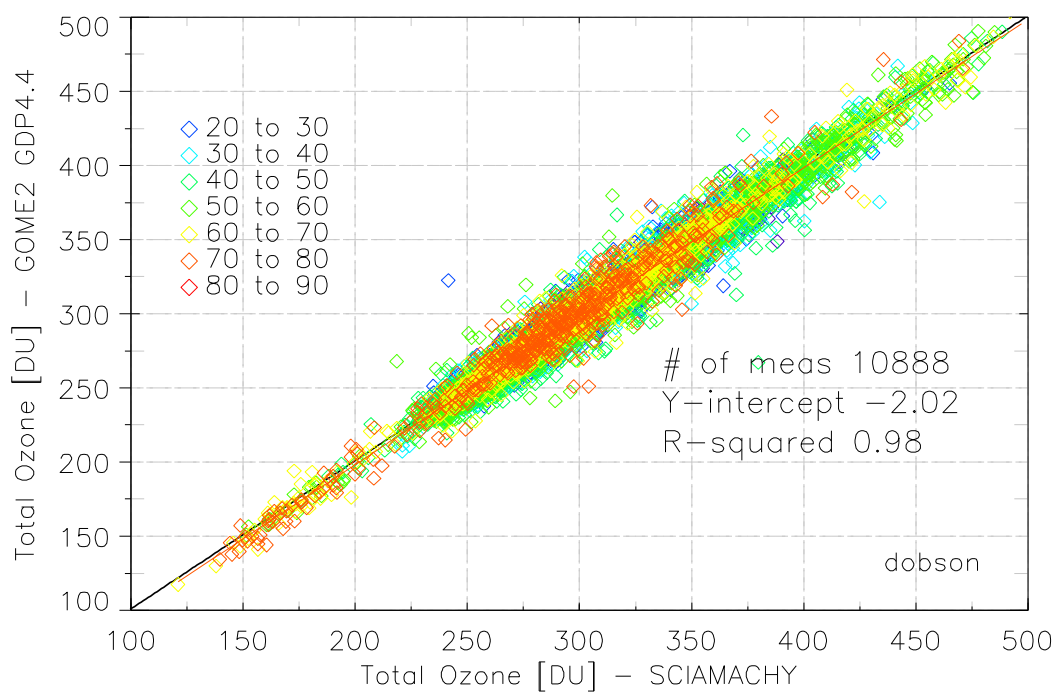
Figure 4 Fourth Row, Right.



725

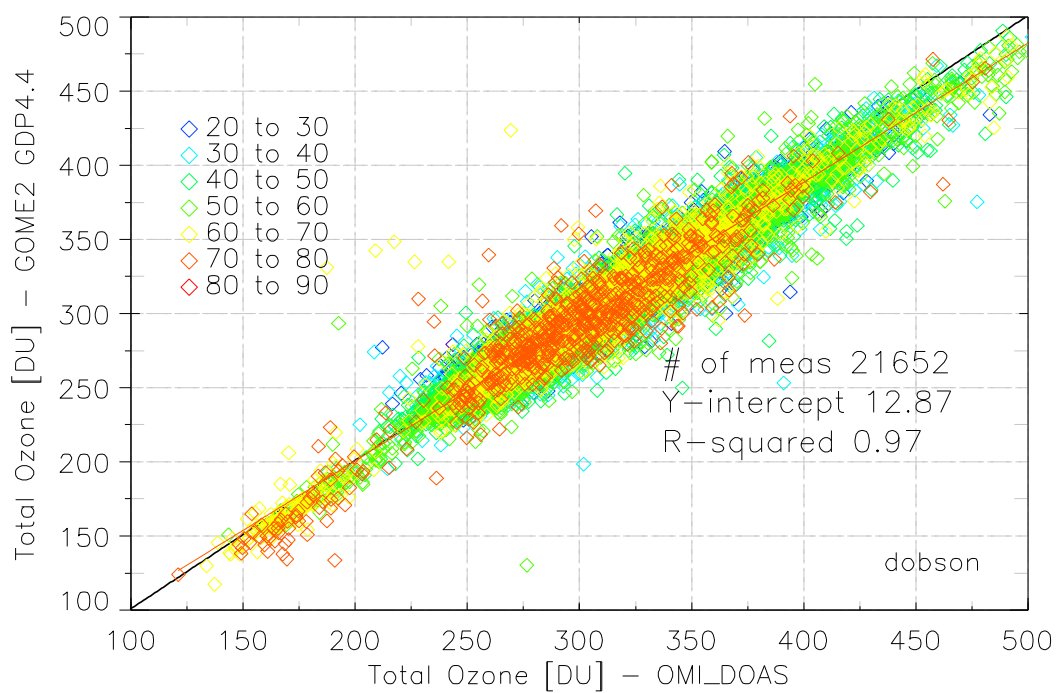
726

Figure 5 Upper, Left.



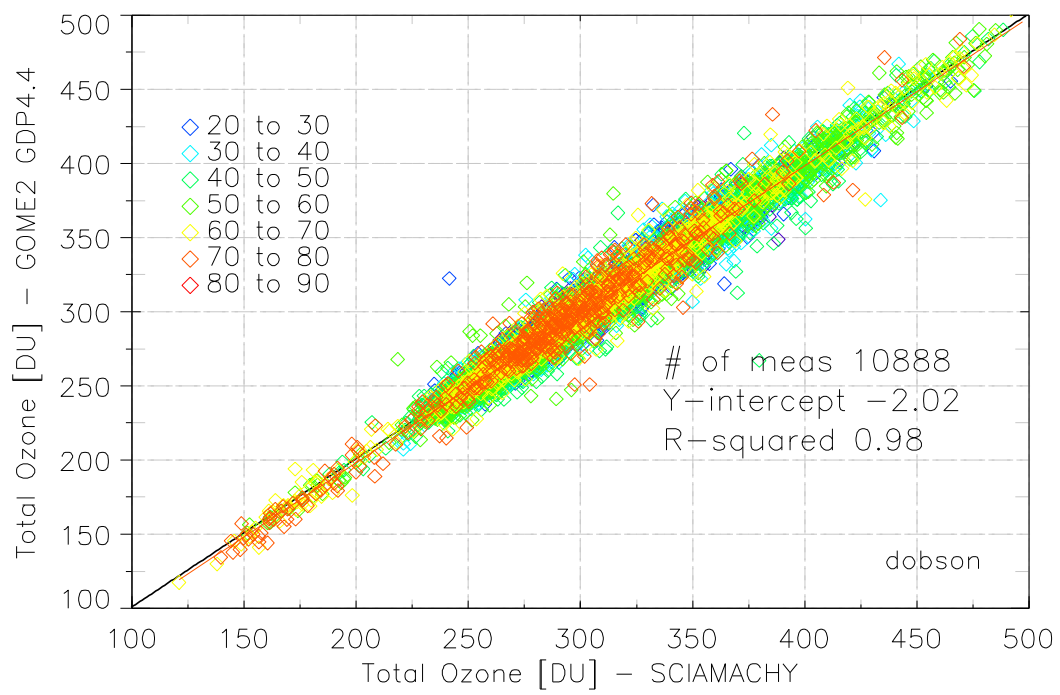
727
728

Figure 5 Upper, Right.



729
730

Figure 5 Bottom, Left.



731
732

733
734