O3M SAF VALIDATION REPORT

Validated products:

Identifier	Name	Acronym
O3M-01 O3M-41	Near Real Time Total Ozone from GOME-2A & B	MAG-N-03 MBG-N-O3
O3M-06 O3M-42	Offline Total ozone from GOME-2A & B	MAG-O-03 MBG-O-O3
O3M-110	Reprocessed Total ozone from GOME-2A & B	MxG-RP1-O3

Authors:

Name	Institute
Mariliza Koukouli	Aristotle University of Thessaloniki
Eirini Zyrichidou	Aristotle University of Thessaloniki
Dimitris Balis	Aristotle University of Thessaloniki
Pieter Valks	German Aerospace Center
Nan Hao	German Aerospace Center

Reporting period: for GOME2/METOP-A, January 2007 – December 2014

for GOME2 /METOP-B, January 2013 – December 2014

Input data versions: GOME-2 Level 1B version 5.3.0. until 17 Jun 2014

GOME-2 Level 1B version 6.0.0 since 17 Jun 2014

Data processor versions: GDP 4.8, UPAS version 1.3.9

Contents:

ACRONYMS AND ABBREVIATIONS

AUTH	Aristotle University of Thessaloniki
BDM	Brion, Daumont, Malicet
BUFR	Binary Universal Form for the Representation of meteorological data
CDOP	Continues Development and Operations Proposal
DLR	German Aerospace Center
DOAS	Differential Optical Absorption Spectroscopy
FM98	Flight Model
GDP	GOME Data Processor
GOME	Global Ozone Monitoring Experiment
HDF	Hierarchical Data Format
MetOp	Meteorological Operational satellite
NRT	Near-real-time
NTO/O3	Near-real-time Total Ozone Product
O3MSAF	Ozone Monitoring Satellite Application Facility
OMI	Ozone Monitoring Instrument
OTO/O3	Offline Total Ozone Product
SCIAMACHY	Scanning Imaging Absorption SpectroMeter for Atmospheric Chartography
SZA	Solar Zenith Angle
TOMS	Total Ozone Mapping Spectrometer
WOUDC	World Ozone and UV Data Center

Applicable O3MSAF Documents

[ATBD] Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, NO_2, BrO, SO_2, H_2O, HCHO and Cloud Properties (GDP 4.8 for O3M-SAF OTO and NTO), DLR/GOME-2/ATBD/01, Rev. 3/A, Valks, P., et al., March 2015.

[PUM] Product User Manual for GOME Total Column Products of Ozone, NO_2, BrO, SO_2, H_2O, HCHO and Cloud Properties, DLR/GOME/PUM/01, Rev. 3/A, Valks, P., et. al., 2015.

[PRD] Product Requirements Document, SAF/O3M/FMI/RQ/PRD/001/Rev. 1.7, D. Hovila, S. Hassinen, D. Loyola, P. Valks, J., S. Kiemle, O. Tuinder, H. Joench-Soerensen, F. Karcher, 2015.

Technical information

GOME2/Metop-A & GOME2/Metop-B products name **Total ozone column (MAG-N-O3, MBG-N-O3, MAG-O-O3, MBG-O-O3, MxG-RP1-O3)**

Validation reporting period **January 2007 –December 2014**

Level-2 processor version **GDP 4.8, UPAS version 1.3.9**

Input GOME-2/MetOp-A Level-1B data version table

Start Date	*Start Orbit*	*Level 1B Version*
Jan., 2007	1235	5.3.0
Jun. 17, 2014	39748	6.0.0

Input GOME-2/MetOp-B Level-1B data version table

Start Date	*Start Orbit*	*Level 1B Version*
Jan. 2013	1235	5.3.0
Jun. 17, 2014	9062	6.0.0

1. INTRODUCTION

The main aim of this report is to validate the GOME2/MetOp-A and GOME-2/MetOp-B near real time, offline and reprocessed total ozone products against the Dobson and Brewer spectrophotometer ground-based networks. In addition, we directly compare the current and previous algorithm (GDP4.8 vs GDP4.7), in order to study the effect of the new algorithm on the measurements.

The structure of the report follows: first we present summary global averages of the statistics from the comparisons between the GOME2/MetOp-A and GOME2/MetOp-B, hereafter GOME-2A and GOME-2B respectively, GDP4.8 total ozone products and the ground based instruments, separately performed for the Dobson and Brewer spectrophotometers.

The global summary statistics include:

- Mean difference per latitude band (10^o) between ground-based and GOME-2A and GOME-2B.
- Solar zenith angle dependence of the differences.
- Time series of the mean differences between ground-based and GOME-2A and GOME-2B instruments for the Northern Hemisphere.
- Time series of the monthly mean differences between ground-based and GOME-2A and GOME-2B instruments for the Southern Hemisphere (only for the Dobson comparisons).
- Histogram of the differences.
- Total ozone dependence of the differences.
- Dependence of the differences on cloud cover and cloud top height.

In all plots that also contain an error bar this represents the 1-σ, i.e. the standard error on the mean percentage differences. The mean values are always extracted from averaging of all individual daily measurements that fall within the bin in question. The monthly mean values were calculated when we had at least two valid measurements per month.

The statistics are presented in the following sequence: first, the global and latitudinal comparisons of GOME-2A GDP4.8 and ground-based stations, followed by the same for GOME-2B. These comparisons do not apply to the common periods of both instruments, but each one is presented in its own time line, i.e. GOME-2A from 2007 to 2014 and GOME-2B from 2013 to 2014. In the following sections, the common days between GOME-2A and GOME-2B are analysed for both GDP4.7 and GDP4.8 and the plots then contain direct comparative information. These comparison results are summarized in Tables in section 8 and finally the conclusions of the current report are presented in the final section 9.

In all the plots we use the label "GOME2A" for GOME-2/MetOp-A and "GOME2B" for GOME-2/MetOp-B.

2. DATA SOURCES

a. GOME2 Metop-A and Metop-B TOCs

The GOME-2A and GOME-2B total ozone products (both NRT and offline) and reprocessed data set have been processed with the DOAS algorithm version GDP4.8. All products have been stored locally and have been separately compared with ground-based data. Since the comparisons of NRT and offline/reprocessed GOME-2 with the ground based measurements are almost identical, as they use exactly the same algorithm and settings, in this report we only present the results for the offline/reprocessed data.

Instrument characteristics

We present, in table format, the instrument characteristics of each satellite instrument considered in the direct comparisons. Differences in the estimated total ozone can be a result of differences in the level-1 products, in the instruments and satellites themselves and therefore such differences should be taken into account when comparing two satellite datasets.

Table 2.1 Main characteristics of the GOME2/MetOp-A and GOME2/MetOp-B instruments affecting the total ozone column products.

	GOME-2/MetOp-A	**GOME-2/MetOp-B**
Principle	UV/VIS grating spectrometer	UV/VIS grating spectrometer
Detectors	Reticon linear diode array	Reticon linear diode array
Spectral resolution	0.26 nm	0.26 nm
Spatial resolution (default)	80 x 40 km^2 40 x 40 km^2 since July 15, 2013	80 x 40 km^2
Swath width	1920 km 960 km since July 15, 2013	1920 km
Eq. crossing time	09:30 LT	09:30 LT
Level-0-to-1b alg.	GOME2 PPF 5.3.0, 6.0.0	GOME2 PPF 5.3.0, 6.0.0
Level-1-to-2 alg.	GDP 4.8	GDP 4.8

* 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.

Algorithm changes in GDP 4.8

A detailed description of the GOME-2 total ozone algorithm can be find in the ATBD by Valks et al., 2015. Details on the differences between GDP4.7 and GDP4.8 are summarized below in brief:

DOAS algorithms

- Improved Kurucz Solar reference spectrum (SAO2010) for wavelength calibration
- Improved Ozone (I0 effect) and Ring Cross sections (using new high resolution solar spectra SAO2010)

Iterative AMF/VCD

- Updated scan angle correction. In GDP4.8, the correction factors were calculated using 7 years GOME-2A data for GOME-2A and 2 years GOME-2B data for GOME-2B respectively.

Cloud treatment

- Using new cloud (version 3.0) algorithms:
- OCRA: PMD degradation correction + new cloud-free map based on GOME-2A data
- ROCINN: New Tikhonov inversion + updated RTM (spectroscopy, a-priori surface albedo, etc)

These new cloud settings have resulted in a different distribution of cloud fraction especially for low cloud fraction, with a reduced number of cloud free scenes. Also the cloud top height is found to generally be lower in GDP4.8 than in GDP4.7. The distribution of the cloud albedos have changed as well, ranging from 0.25 to 1.0 in the GDP 4.7, and from 0.05 to 1.0 in the GDP 4.8.

All these algorithm changes are expected to affect the AMF calculation, the ozone ghost columns and the intra-cloud corrections hence the effect on the total ozone columns required a detailed validation activity given in the sections below.

Details on the new cloud algorithm can be found in Lutz, et al., 2015.

- ***GOME-2A O_3 Differences between GDP 4.8 and GDP 4.7***

The daily difference of collocated ozone data from GOME-2A is displayed in Figure 2.1 for two days, the 13^th^ of September 2007 on the left and the 13^th^ of September 2013 on the right. This comparison is based on daily gridded data at a 0.25 degree by 0.25 degree horizontal resolution. The criteria for co-location is the pixel numbers in a grid are larger than 5.

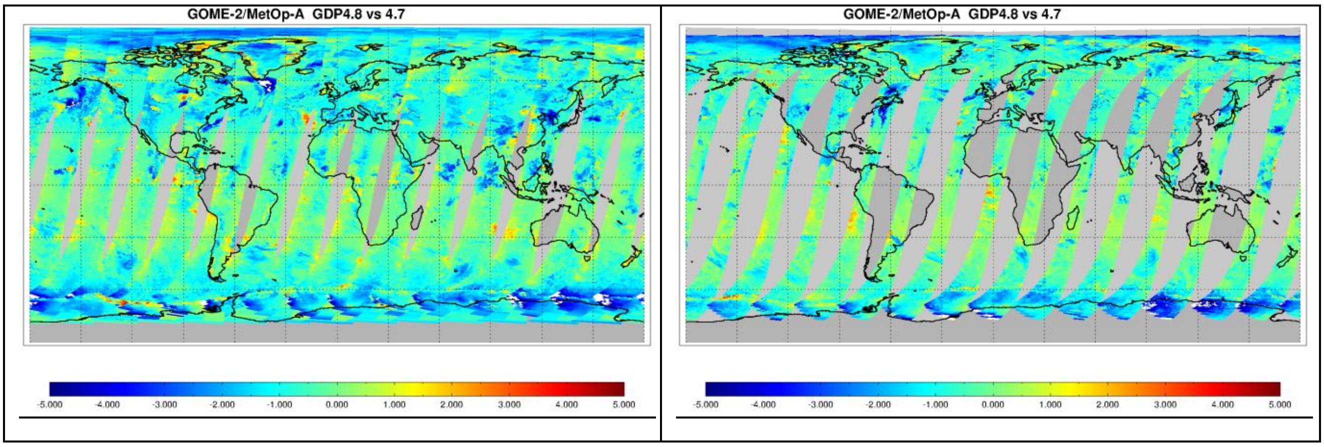


Figure 2.1. Daily relative differences (in percent) of total ozone vertical columns from GOME-2A for 13 September 2007 (left) and 13 September 2013 (right). Note that GOME-2A operated on a reduced swath of 960 km after 15 July 2013.

The differences between GDP 4.8 and GDP 4.7 are due to changes of the DOAS algorithm, the cloud treatment and the scan angle corrections. As shown in Figure 2.2, the monthly average differences of ozone columns between GDP4.8 and GDP4.7 for GOME-2A (January and October 2013 in the left and right panels respectively) are in general within ±1%. The larger differences in high latitudes, reaching the +2% level, are related to the changes of the cloud treatment which are expected to affect the AMF calculation, the ozone ghost columns and the intra-cloud correction.

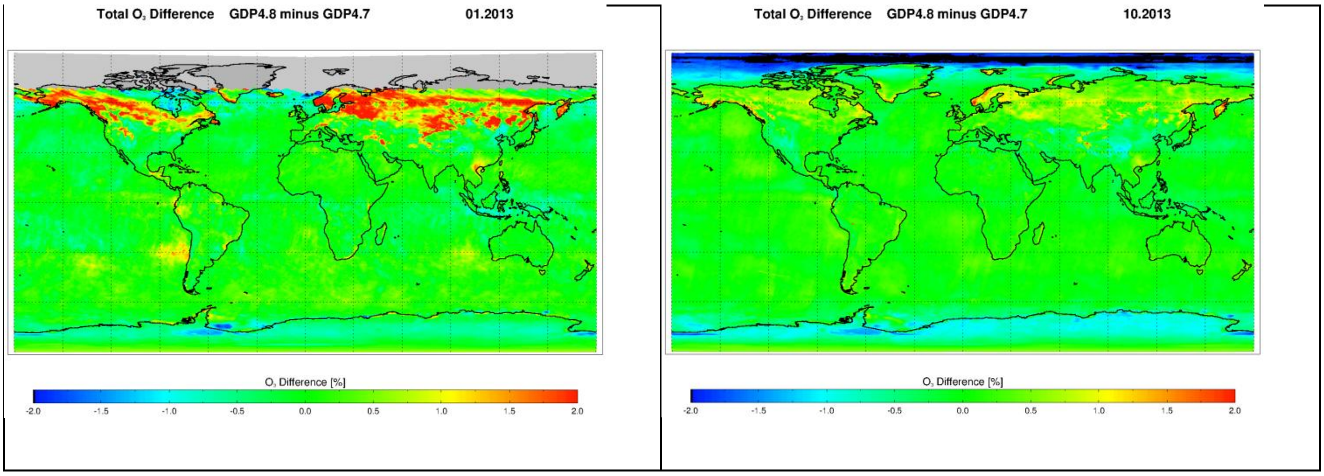
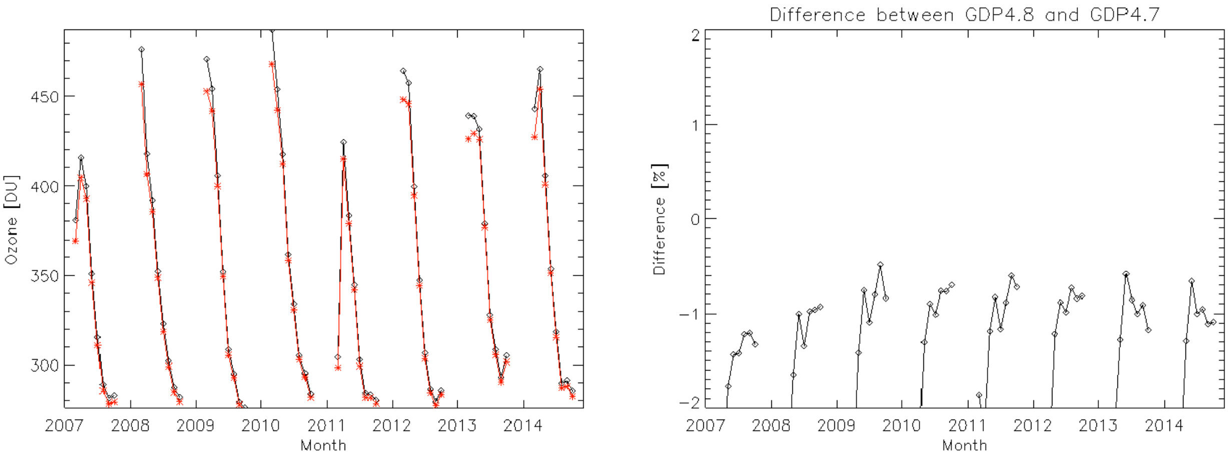
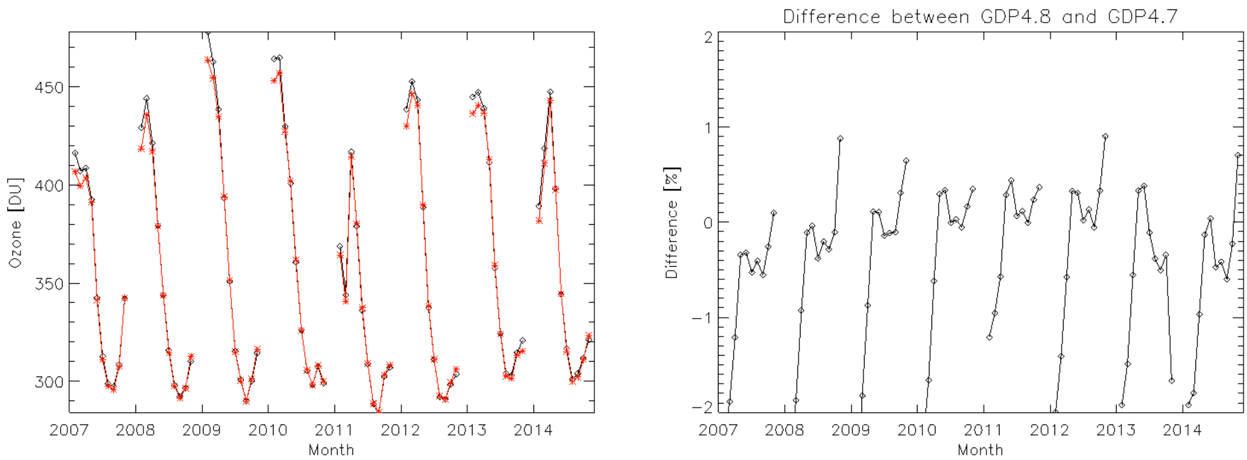


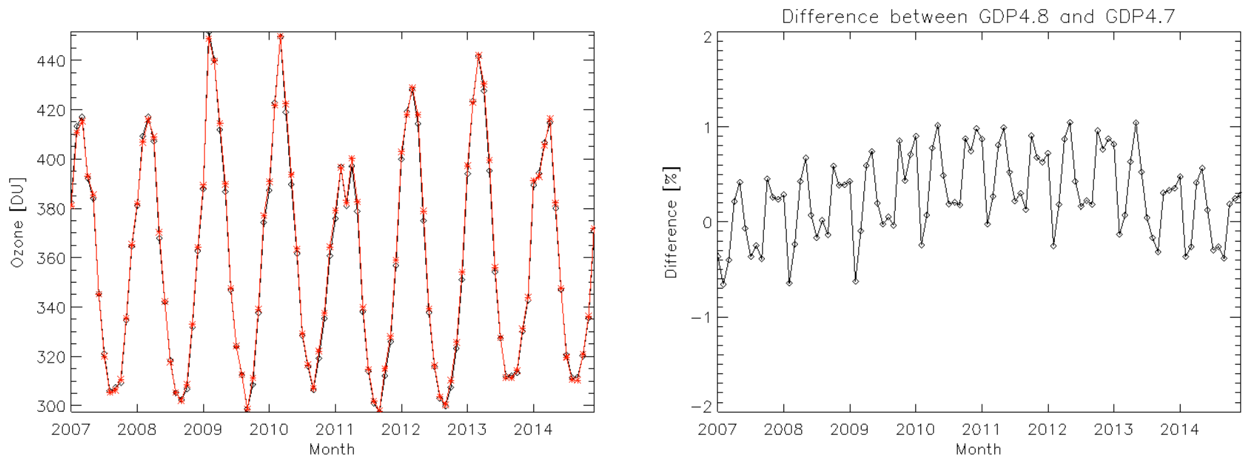
Figure 2.2. Monthly average relative differences (in percent) between total ozone columns from GOME-2A for January 2013 (left) and October 2013 (right).

a) Latitude : 80N-90N



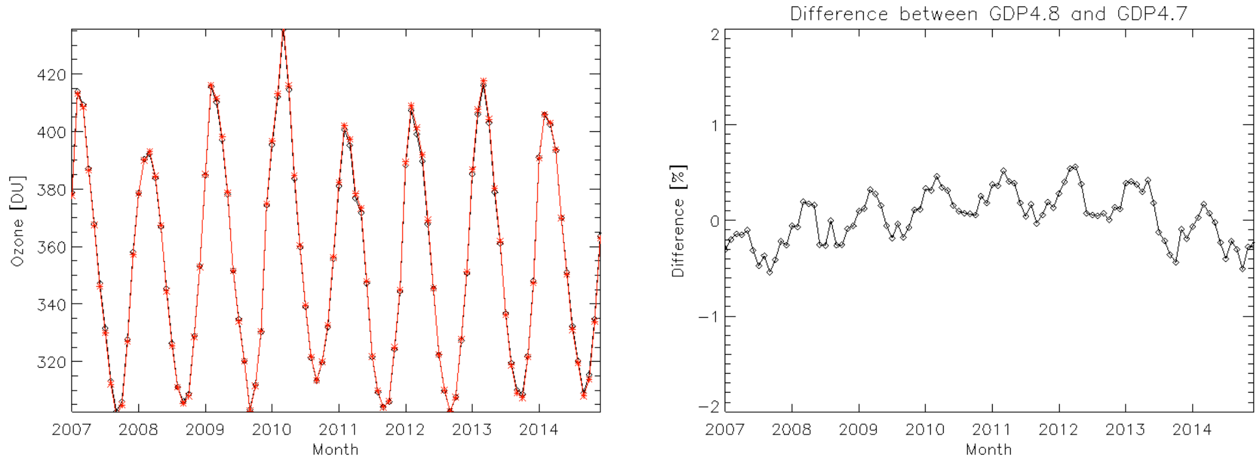
b) Latitude : 70N-80N



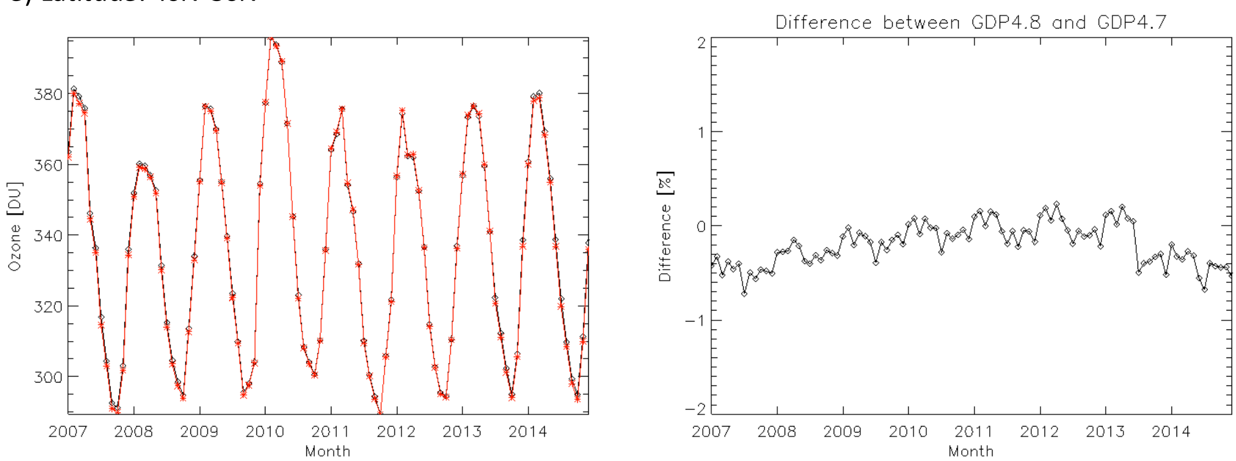


c) Latitude : 60N-70N

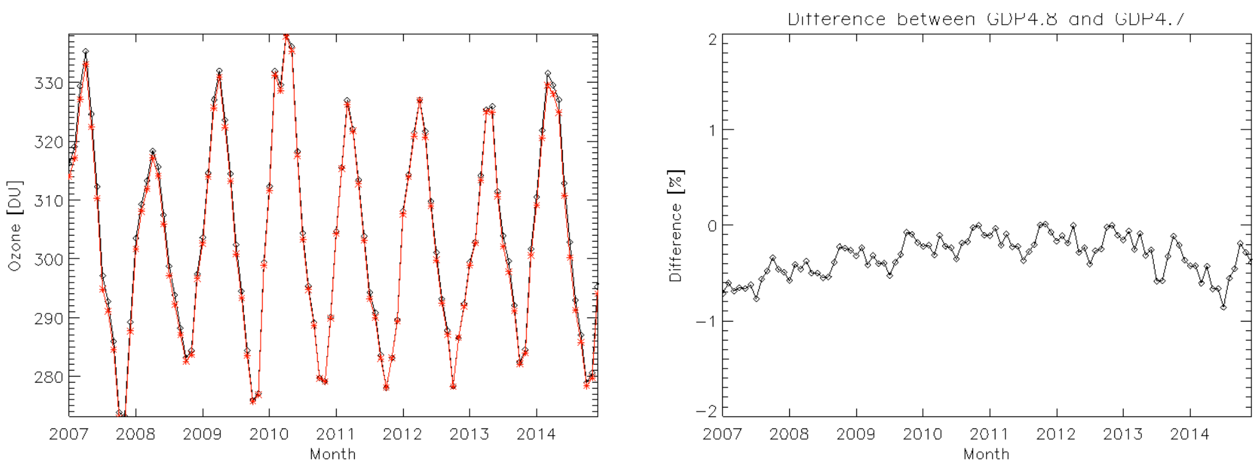
d) Latitude : 50N-60N



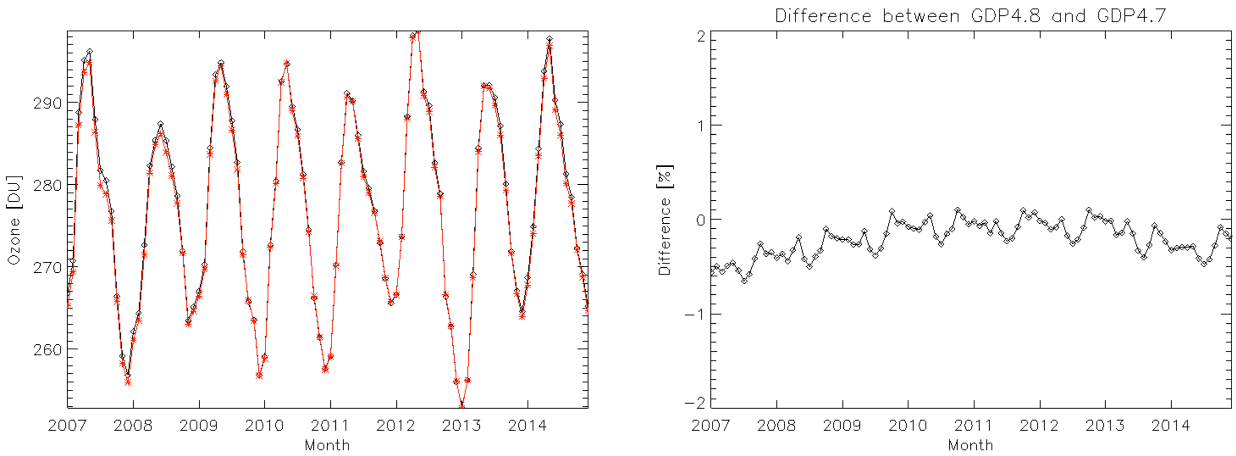
e) Latitude: 40N-50N



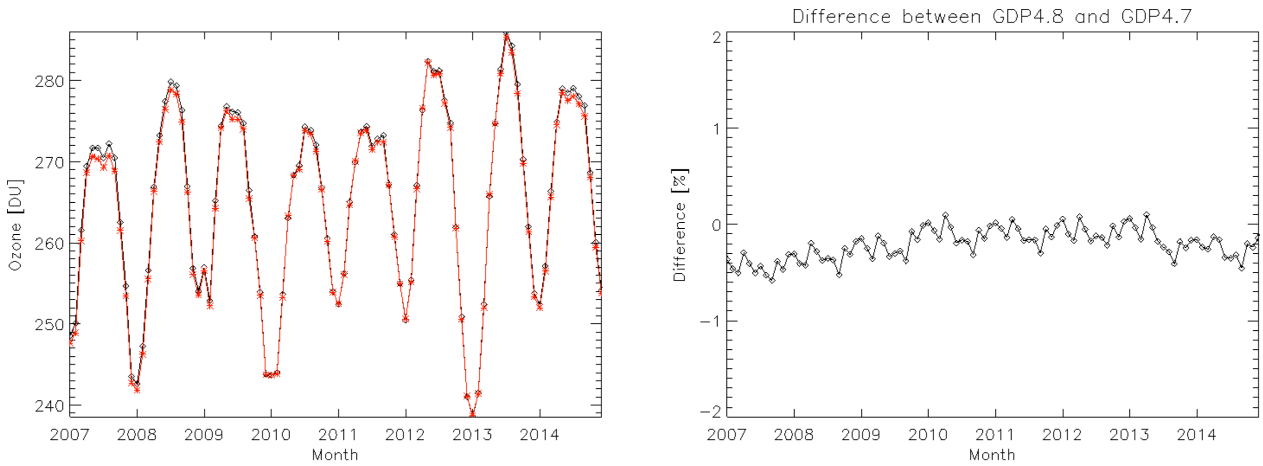
f) Latitude : 30N-40N



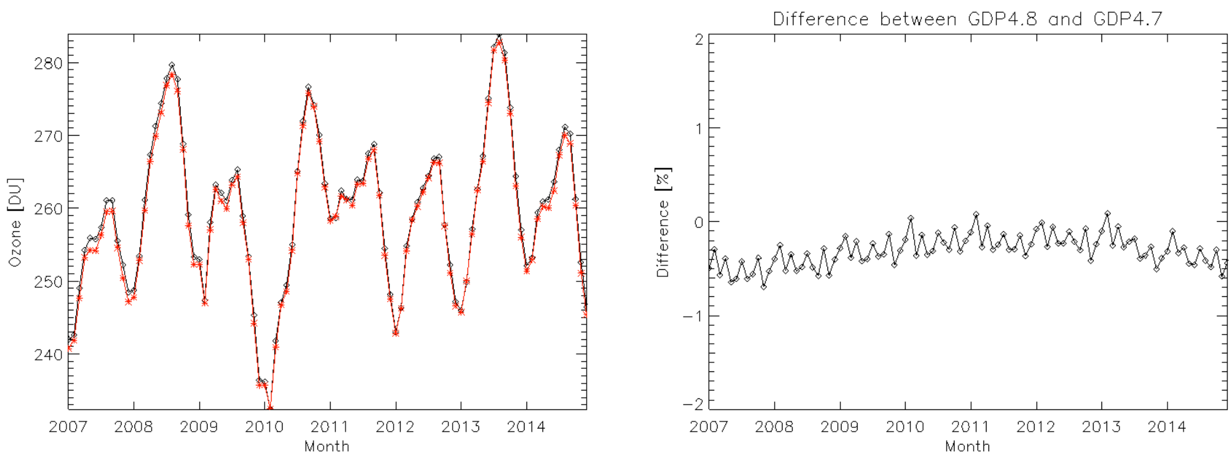
g) Latitude: 20N-30N:



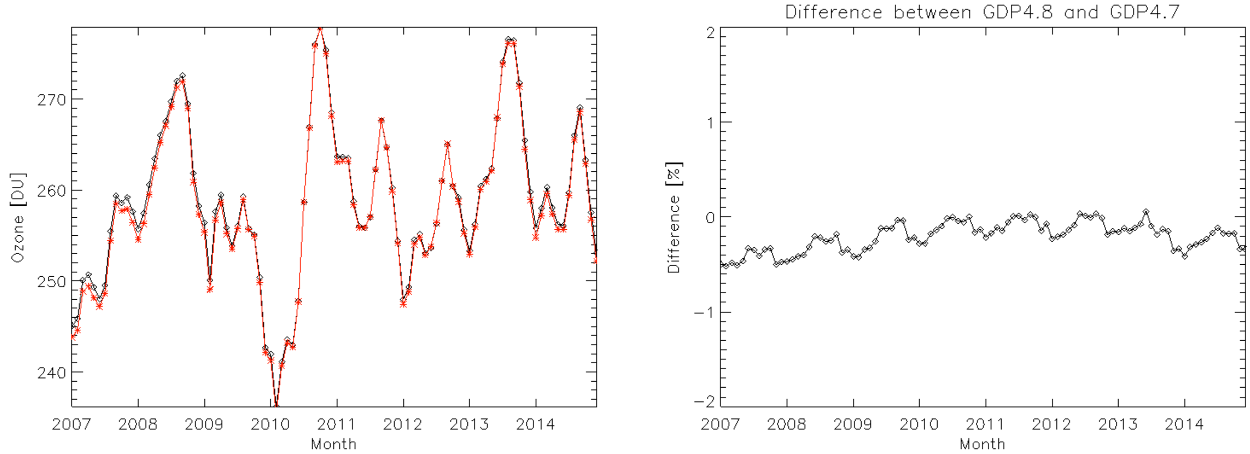
h) Latitude: 10N-20N



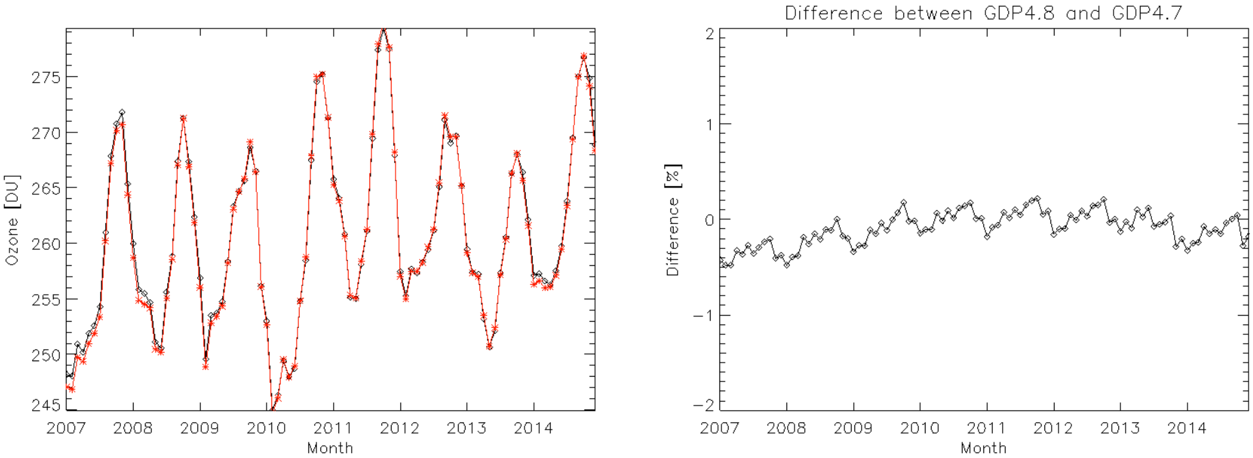
i) Latitude: 0-10N



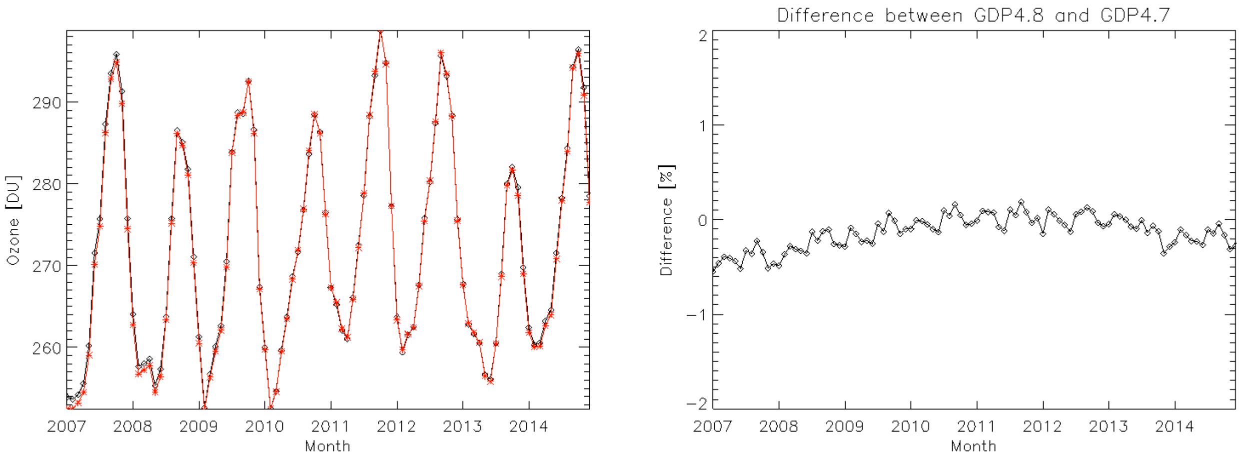
j)Latitude: 10S-0



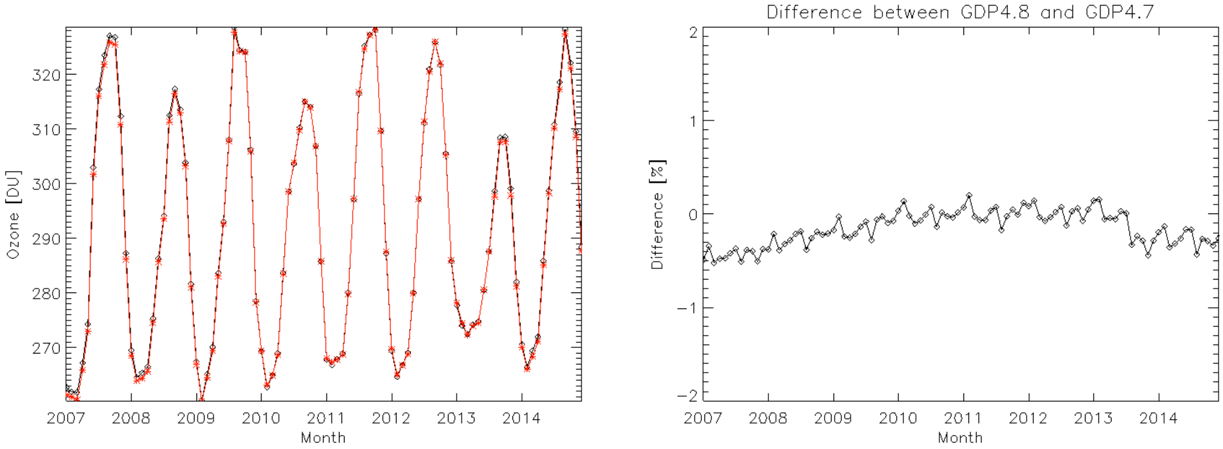
k) Latitude: 20S-10S



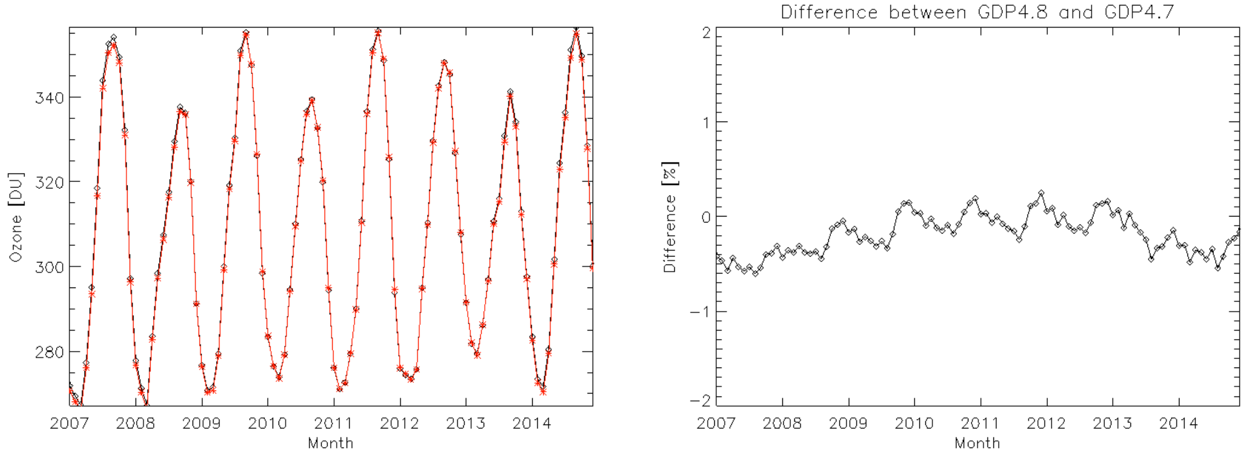
l) Latitude: 30S-20S



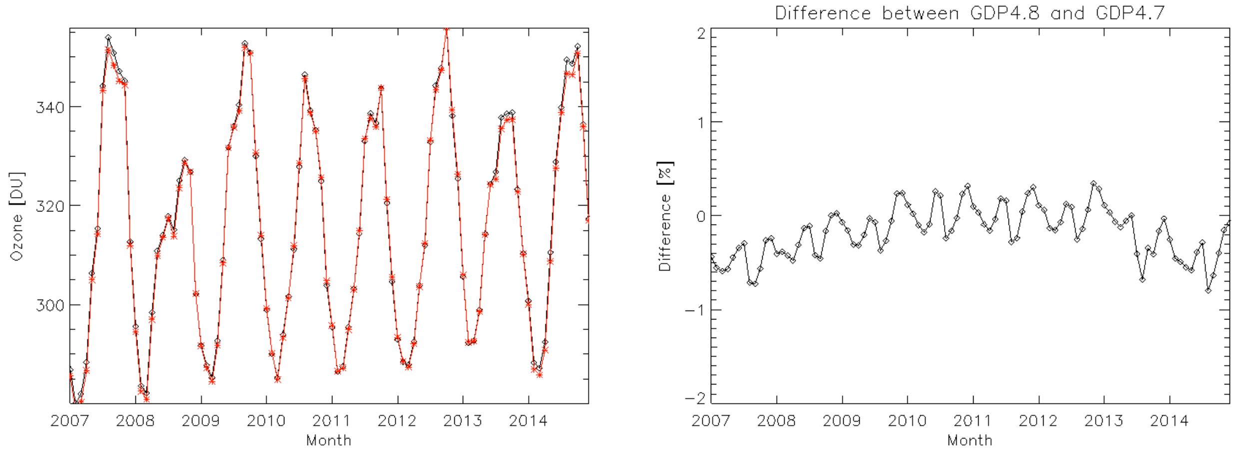
m) Latitude: 40S-30S



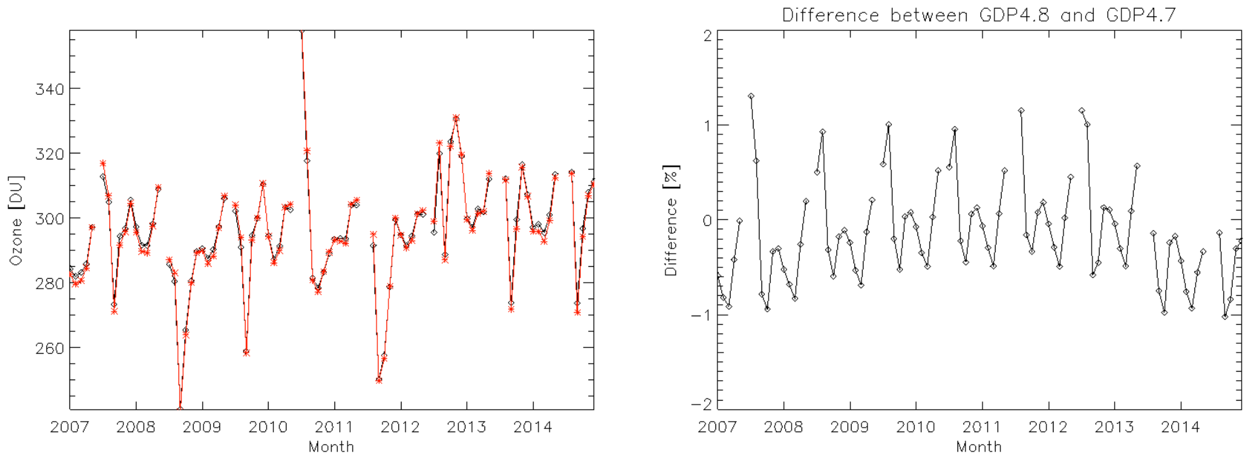
n) Latitude: 50S-40S



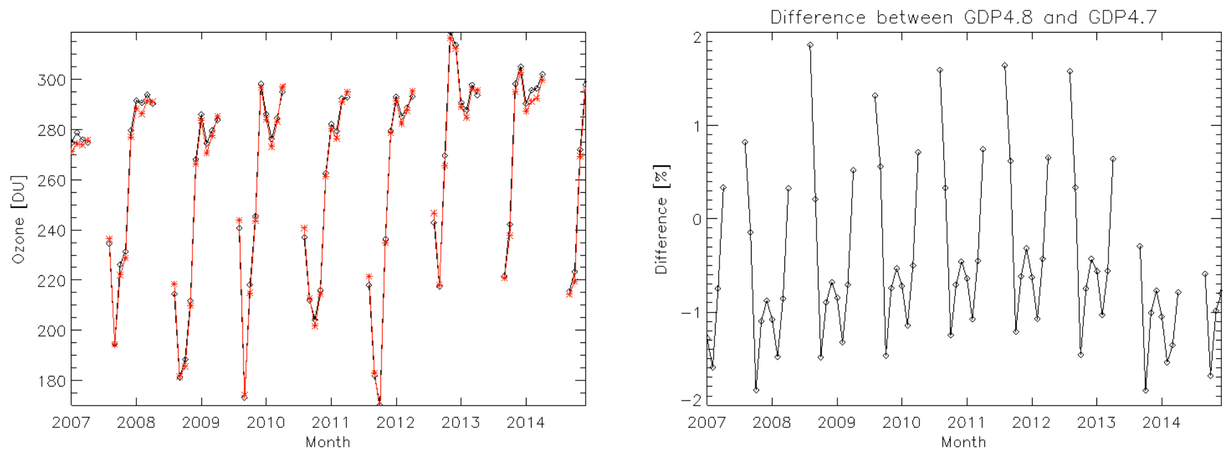
o) Latitude: 60S-50S



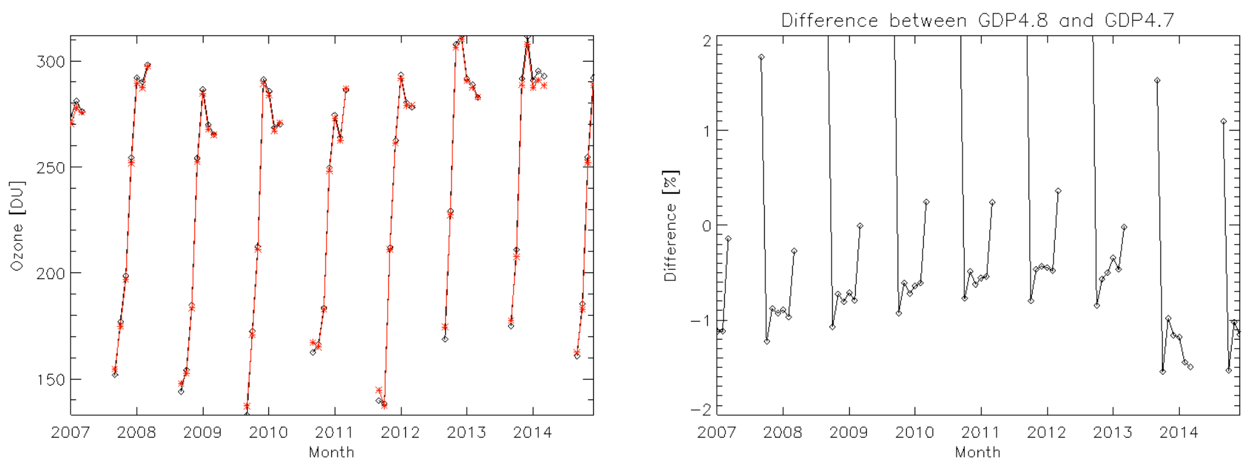
p) Latitude: 70S-60S



q) Latitude: 80S-70S



r) Latitude: 90S-80S



s) All Latitude

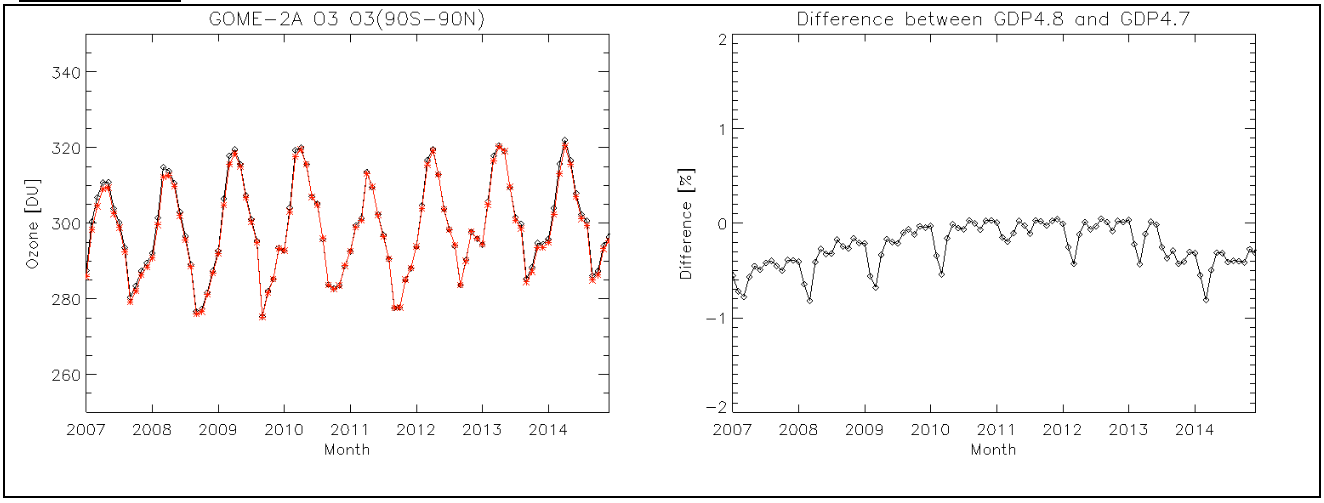


Figure 2.3. (**Left**) Total O_3 from GOME-2A GDP 4.8 (red) and 4.7 (black) and (**right**) relative difference for 10° latitude bands (a-r) and for all latitudes (s).

In **Figure 2.3** the TOC from GOME-2A and associated differences between GDP4.8 and GDP4.7 are shown for different latitude bands and on a global scale for the duration of the mission. On average, i.e. on a global scale [final panel], GOME-2A GDP4.8 produces about 0.24±0.22% lower TOCs than GOME-2A GDP4.7, with larger difference at high latitudes and distinct behavior at different latitude bands. Of note are the NH middle-latitude bands, i.e. panels d), e) and f). In the beginning of the mission, GDP4.8 produces smaller TOCs up to around year 2010 where the differences approach zero and, depending on the latitude band, turn slightly positive, and then the differences again move towards the negative after year 2013. This is an important point to keep in mind since most of the Brewer and Dobson spectrophotometers are located in those NH bands and the comparisons that follow are mostly affected by those latitudes. It is also significant to note that, when comparing the GOME-2A and GOME-2B behavior, we are discussing only years 2013 and 2014 which are not necessarily representative of the total GOME-2A behavior, as is clearly seen in this Figure.

- ***GOME-2B O_3 Differences between GDP 4.8 and GDP 4.7***

The daily differences of collocated ozone data from GOME-2B are displayed in Figure 2.4 for the 13th of September 2013. The comparison of GOME-2B is based on the daily gridded data at a 0.25 degree by 0.25 degree horizontal resolution. The criteria for co-location is the pixel numbers in a grid are larger than 5. In Figure 2.5, the monthly average differences of ozone columns between GDP4.8 and GDP4.7 for GOME-2B (January and October 2013 in the left and right panels respectively) are within ± 1%. As for GOME-2A [Figure 2.2, left panel] there are large differences in January at northern high latitudes, albeit on a smaller scale.

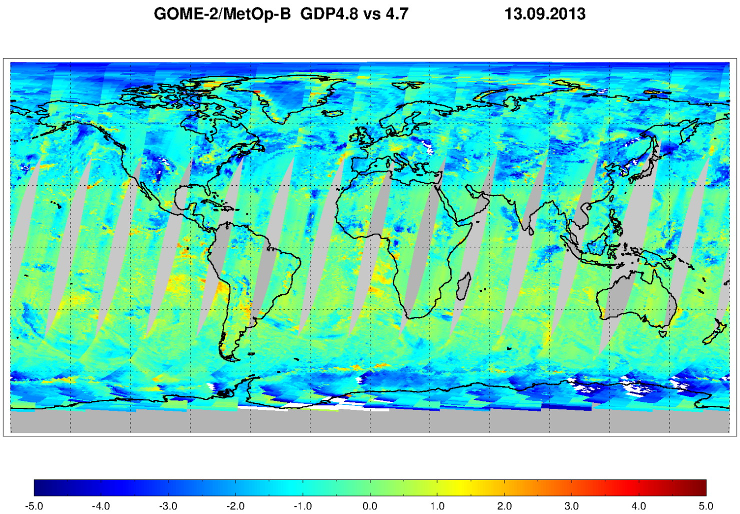


Figure 2.4. Daily relative differences (in percent) of total ozone vertical columns from GOME-2B for the 13th of September 2013.

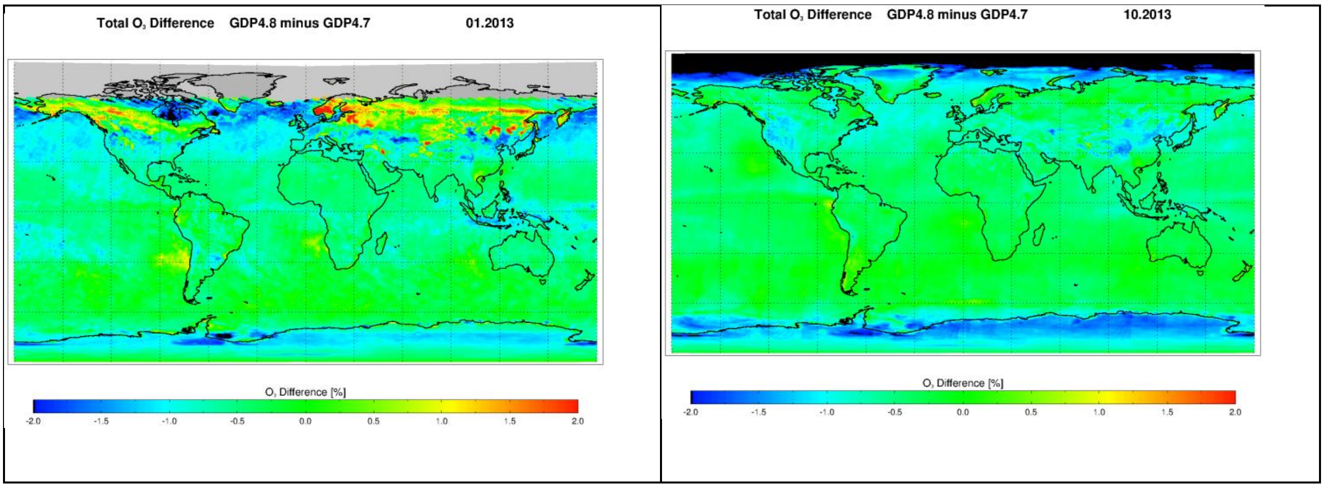
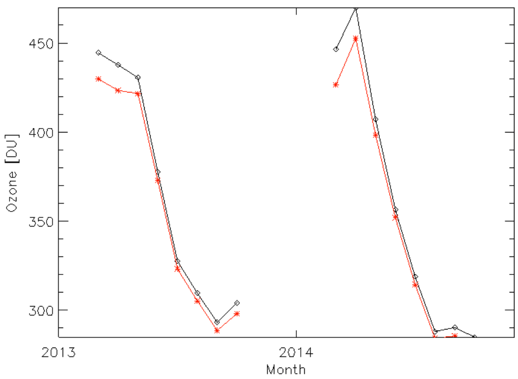
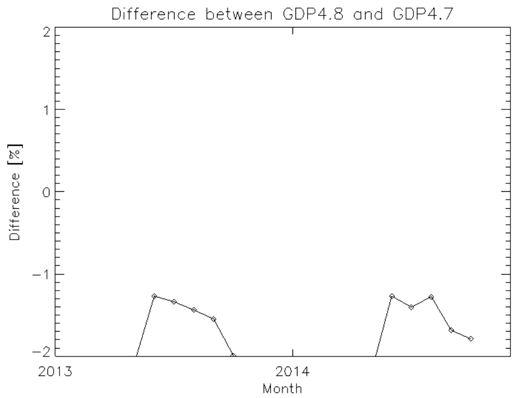


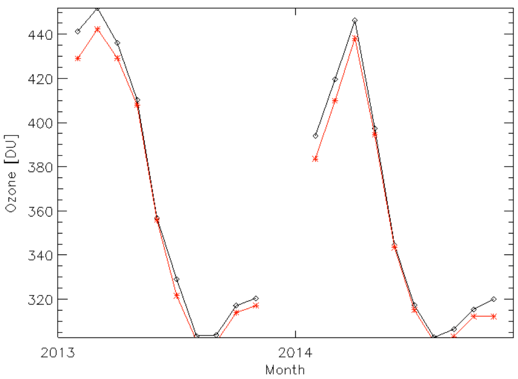
Figure 2.5. Monthly average differences (in percent) between total ozone columns from GOME-2B for January 2013 (left) and October 2013(right).

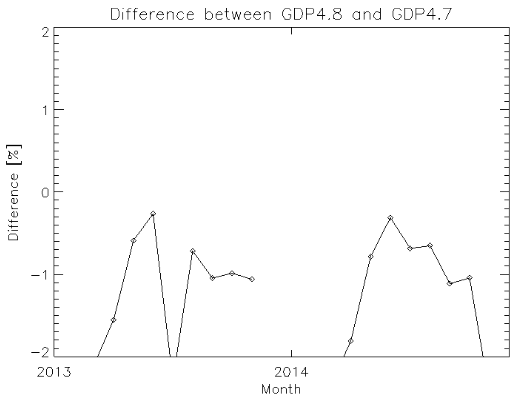
A) LATITUDE : 80N-90N



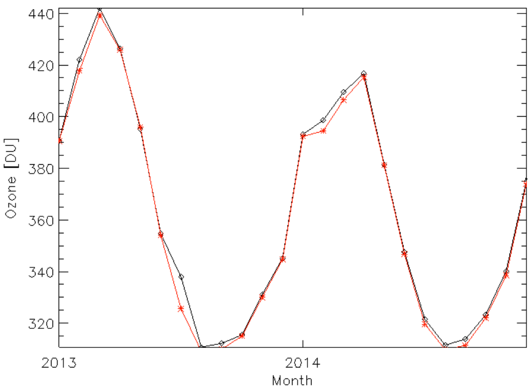


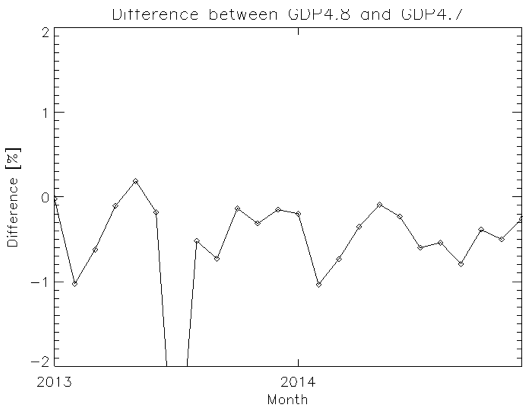
B) LATITUDE : 70N-80N



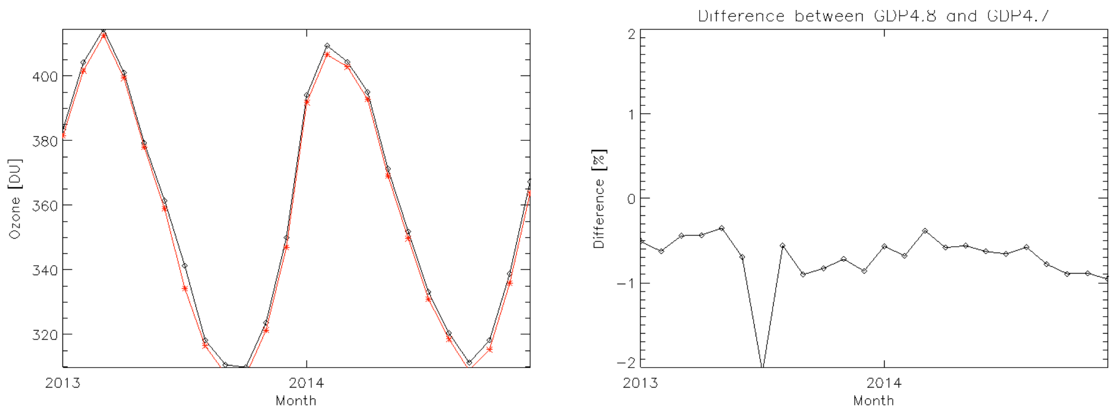


C) LATITUDE : 60N-70N

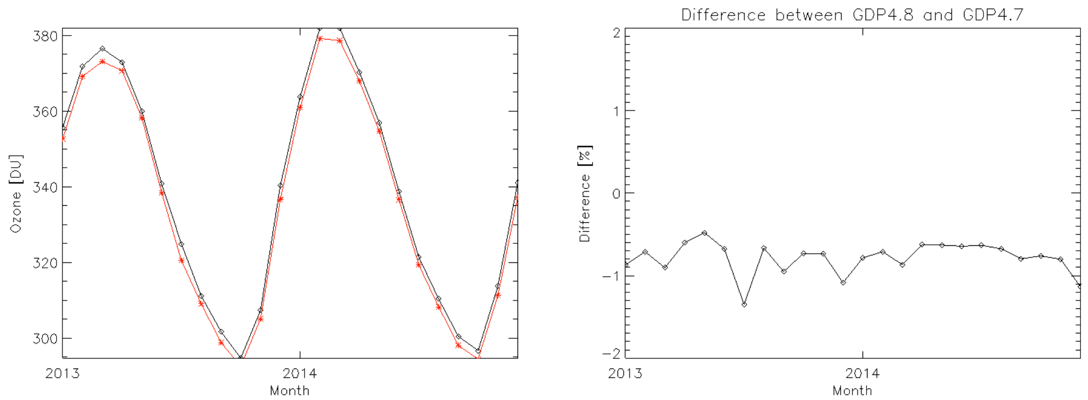




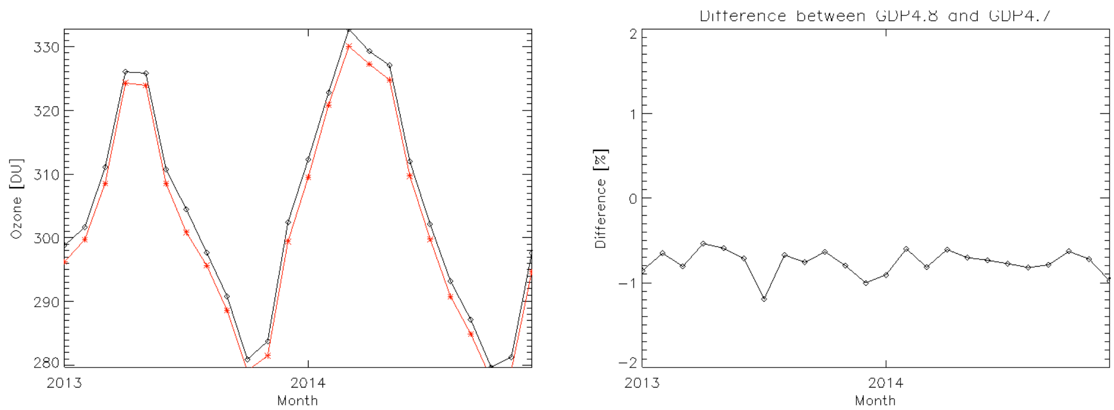
D) LATITUDE : 50N-60N



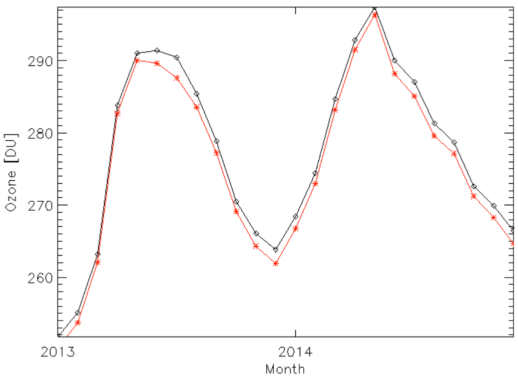
E) LATITUDE: 40N-50N

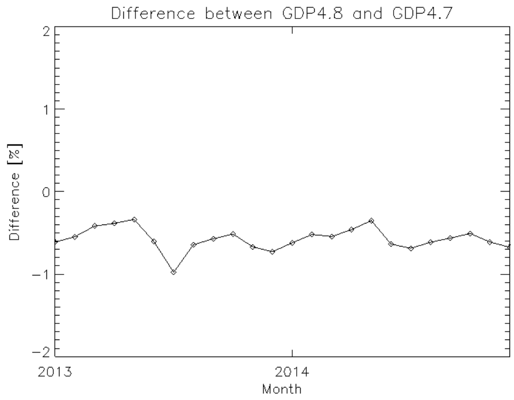


F) LATITUDE : 30N-40N

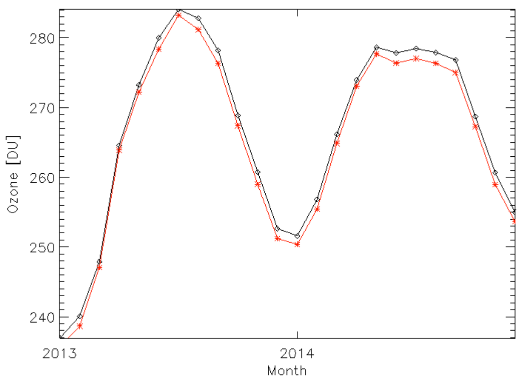


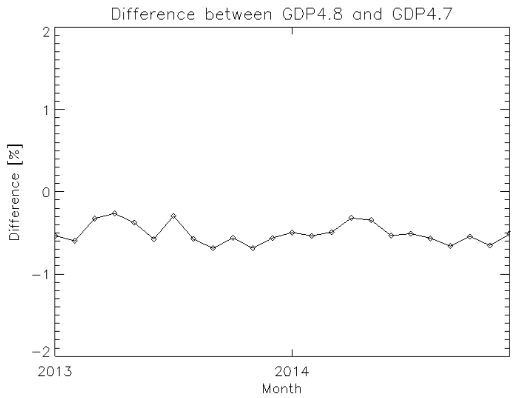
G) LATITUDE: 20N-30N:



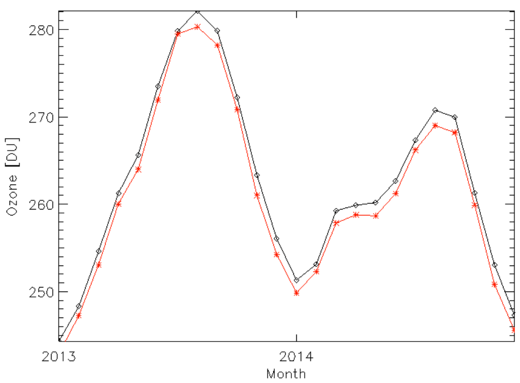


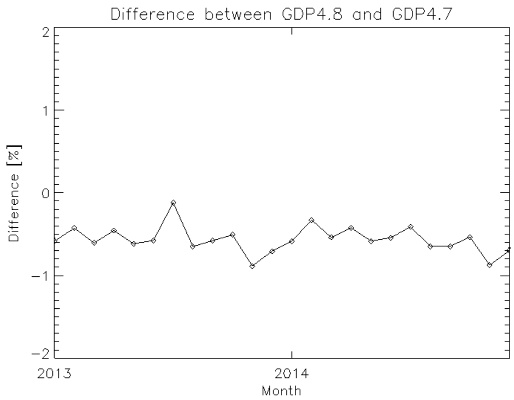
H) LATITUDE: 10N-20N



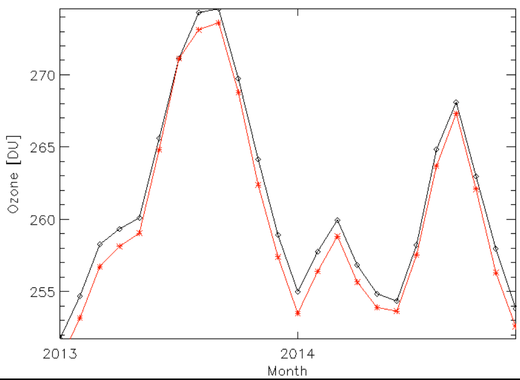


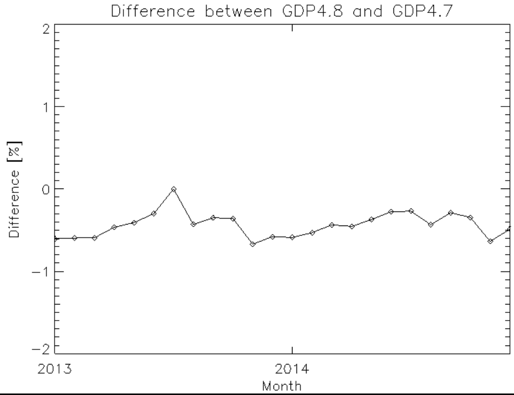
I) LATITUDE 0-10N



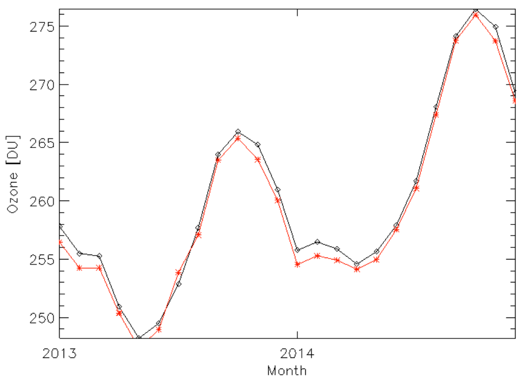


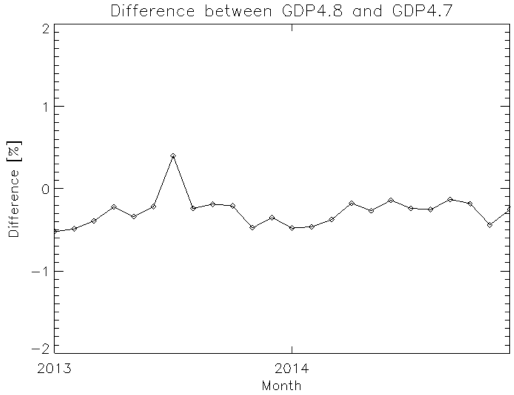
J) LATITUDE: 10S-0



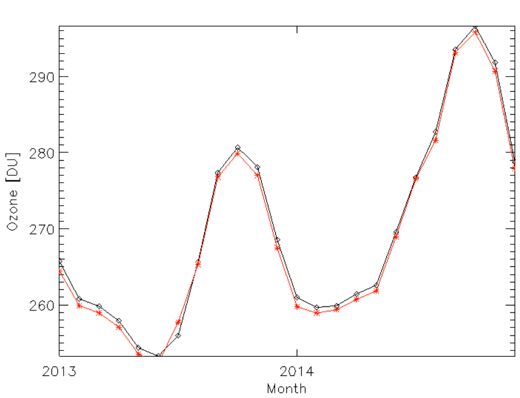


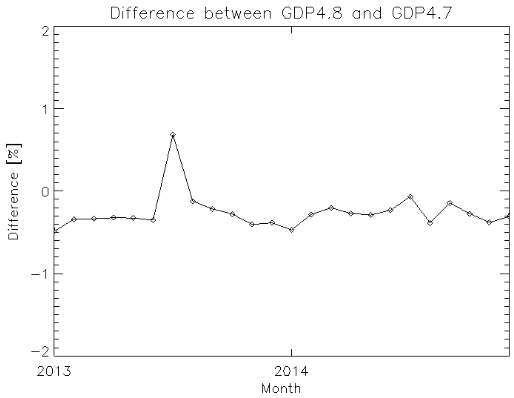
K) LATITUDE: 20S-10S



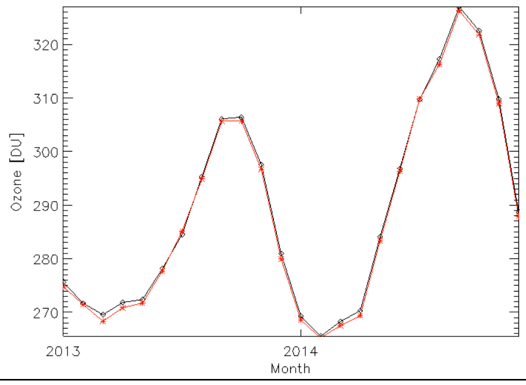


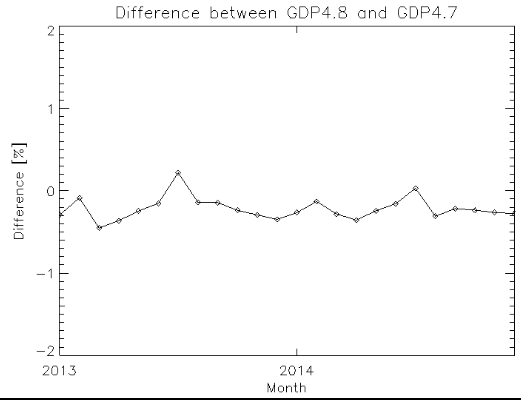
L) LATITUDE: 30S-20S



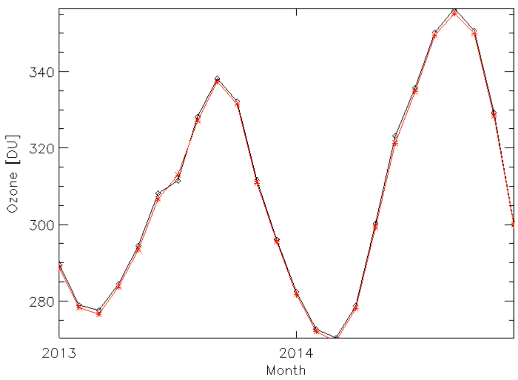


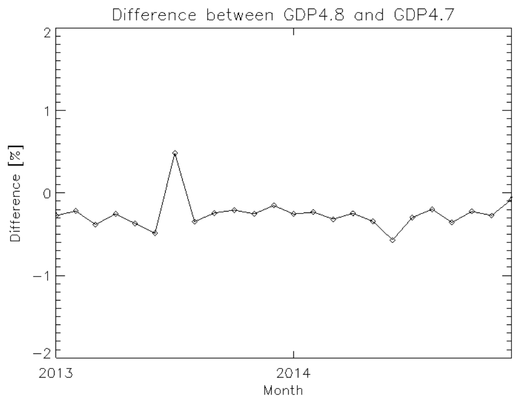
M) LATITUDE: 40S-30S



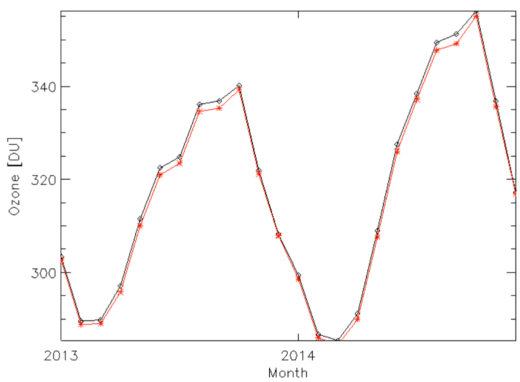


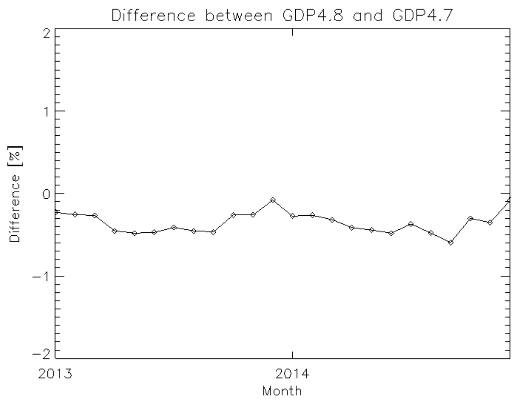
N) LATITUDE: 50S-40S



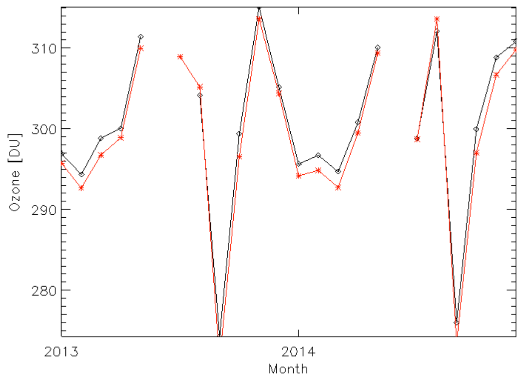


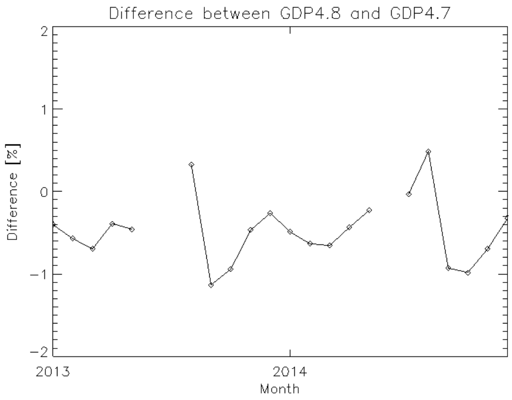
O) LATITUDE: 60S-50S



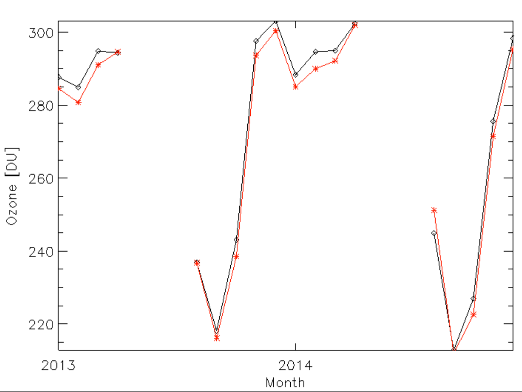


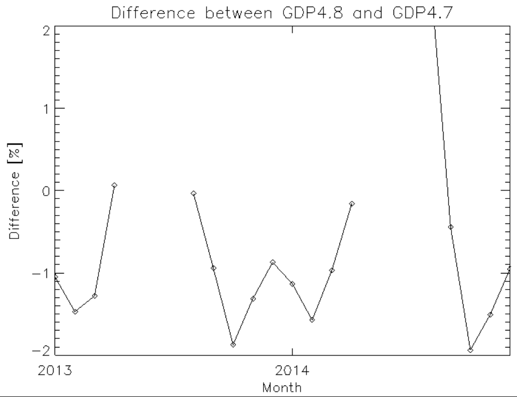
P) LATITUDE: 70S-60S



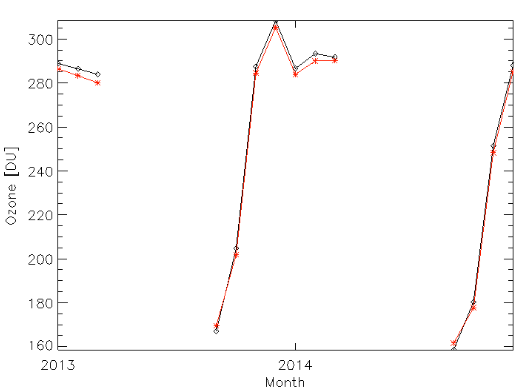


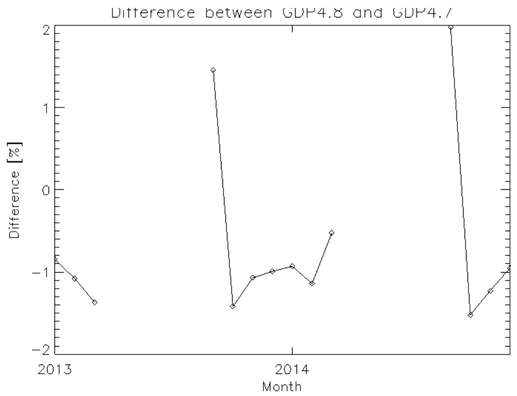
Q) LATITUDE: 80S-70S





R) LATITUDE: 90S-80S





S) ALL LATITUDES

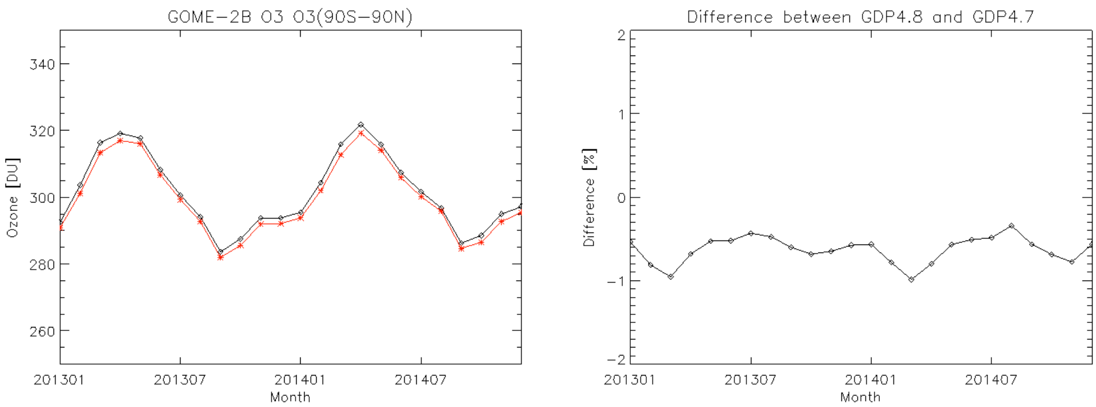


Figure 2.6 (**Left**) Total O_3 from GOME-2B GDP 4.8 (red) and 4.7 (black) and (**right**) relative difference for 10° latitude bands (a-r) and for all latitudes (s).

In **Figure 2.6** the TOC from GOME-2B and associated differences between GDP4.8 and GDP4.7 are shown for different latitude bands on a global scale for the duration of the mission. Differences range within the 1% level, between -1% and 0%, for the two full years shown. A point of interest for GOME-2B appears to be the Arctic band where differences greater than -2% are seen in the two spring times covered by the mission. Some features are also seen in the Antarctic band, however we may be facing representativeness issues here which merit finer bands in latitude to examine properly. These features are under investigation. On average, GOME-2B GDP4.8 produces about 0.61±0.52% lower value than GOME-2B GDP4.7 with larger difference at high latitudes and distinct behavior at different latitude bands.

b. Ground-based observations

Archived Brewer and Dobson total ozone data have been downloaded from the World Ozone and UV Data Centre (WOUDC) at Toronto, Canada (www.woudc.org). These data are of archived quality and are usually quality controlled by each station investigator and also by the WOUDC team. Most stations upload their data to the database two to four months after observation. In this report we use for the comparisons only archived data for the period **January 2007 to December 2014** depending on the availability of data for each individual station. For the Brewer and Dobson comparisons daily mean TOCs are used from the WOUDC datasets. For each common day of measurements, the closest satellite measurement to the ground is found by comparing the central satellite pixel lat/lon to the ground lat/lon using a 150km radius haversine formula. For each day of measurements, the closest satellite measurement is selected within a 150km radius from the station.

Total ozone column data from a large number of stations have already been used extensively both for trend studies [e.g. WMO 1998, 2002, 2006] as well as for validation of satellite total ozone data [e.g. Lambert et al., 1999; Fioletov et al., 1999; Weber et al., 2005, Balis et al., 2007a; 2007b, Antón et al., 2009, Loyola et al., 2011; Koukouli et al., 2012, Labow et al., 2013, Hao et al., 2014, and so on.] These studies have shown that Dobson and Brewer data can agree within 1% when the major sources of discrepancy are properly accounted for. Dobson measurements suffer from a temperature dependence of the ozone absorption coefficients used in the retrievals which might account for a seasonal variation in the error of ±0.9% in the middle latitudes and ±1.7% in the Arctic, and for systematic errors of up to 4% [Bernhard et al., 2005]. The error of individual total ozone measurements for a well maintained Brewer instrument is about 1% (e.g. Kerr, 1988). Despite the similar performance between the Brewer and Dobson stations, small differences within ±0.6% are introduced due to the use of different wavelengths and different temperature dependence for the ozone absorption coefficients [Staehelin et al., 2003]. 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.

The WOUDC stations considered for the comparisons are listed in Tables 2.1 and 2.2 and are spatially depicted in Figure 2.7 and Figure 2.8. In all comparison plots and statistics presented in this report, the direct sun observations provided by the Brewers and Dobsons are utilized.

Table 2.1: List of Brewer ground-based stations used for the comparisons

ID	NAME	LAT	LON	ELEV(m)	Nobs
10	NEW_DELHI	28.63	77.22	216	511
21	EDMONTON	53.57	-113.52	668	3424
24	RESOLUTE	74.72	-94.98	64	1318
35	AROSA	46.77	9.67	1860	2065
53	UCCLE	50.8	4.35	100	3895
65	TORONTO	43.78	-79.47	198	1844
76	GOOSE	53.32	-60.38	44	1763
77	CHURCHILL	58.75	-94.07	35	2511
96	HRADEC_KRALOVE	50.18	15.83	285	3902
99	HOHENPEISSENBERG	47.8	11.02	975	2124
100	BUDAPEST	47.43	19.18	140	1553
111	AMUNDSEN-SCOTT	-89.98	-24.8	2835	718
174	LINDENBERG	52.22	14.12	98	895
213	EL_ARENOSILLO	37.1	-6.73	41	1798
261	THESSALONIKI	40.52	22.97	4	1108
262	SODANKYLA	67.37	26.65	179	615
267	SONDRESTROM	67	-50.98	150	1518
279	NORKOPING	58.58	16.12	0	1999

282	KISLOVODSK	43.73	42.66	2070	1889
284	VINDELN	64.25	19.77	0	1569
290	SATURNA	48.78	-123.13	0	1942
295	MT_WALIGUAN	36.17	100.53	3816	1579
305	ROME_UNIVERSITY	41.9	12.52	0	1785
308	MADRID	40.45	-3.55	0	2512
315	EUREKA	79.89	-85.93	10	1170
316	DEBILT	52	5.18	0	1886
318	VALENTIA	51.93	-10.25	0	1807
322	PETALING_JAYA	3.1	101.65	46	1857
326	LONGFENSHAN	44.75	127.6	0	1652
331	POPRAD-GANOVCE	49.03	20.32	0	2050
332	POHANG	36.03	129.38	0	1638
346	MURCIA	38	-1.17	69	2310
348	ANKARA	39.95	32.88	891	2166
352	MANCHESTER	53.45	-2.26	61	1792
353	READING	51.42	-0.96	51	1887
376	MRSA_MTROUH	31.33	27.22	35	1777
401	SANTA_CRUZ	28.35	-16.29	50	251
405	LA_CORUNA	43.33	-8.5	62	1983
411	ZARAGOZA	41.66	-0.94	235	2145
435	PARAMARIBO	5.78	-55.2	5	1333
447	GODDARD	38.99	-76.83	100	952
454	SAN_MARTIN	-68.13	-67.1	30	703
468	CAPE_D'AGUILAR	22.18	114.23	75	771
478	ZHONGSHAN	-69.4	76.35	26	387
479	AOSTA	45.71	7.33	585	1553
481	TOMSK	56.48	84.97	170	2214
512	UNIVERSITY_OF_TORONTO	43.63	-79.43	189	1286

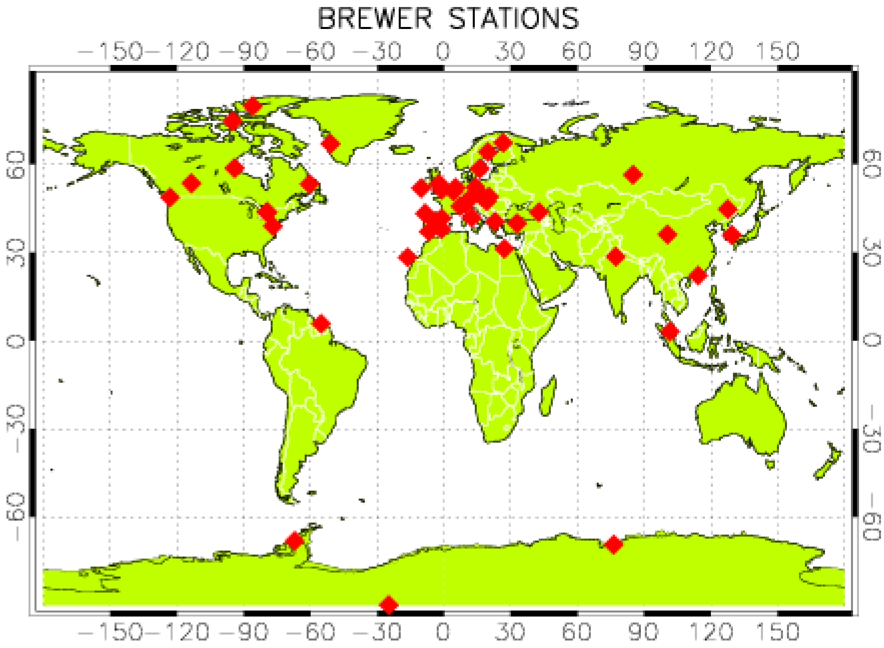


Figure 2.7. Spatial distribution of the Brewer ground-based stations used for the comparisons.

Table 2.2: List of Dobson ground-based stations used for the comparisons

ID	NAME	LAT	LON	ELEV(m)	Nobs
2	TAMANRASSET	22.8	5.52	1395	2062
11	QUETTA	30.18	66.95	1799	1497
12	SAPPORO	43.05	141.33	19	839
14	TATENO	36.05	140.13	31	1060
19	BISMARCK	46.77	-100.75	511	909
20	CARIBOU	46.87	-68.02	192	606
27	BRISBANE	-27.47	153.03	5	1492
29	MACQUARIE_ISLAND	-54.48	158.97	6	582
31	MAUNA_LOA	19.53	-155.58	3397	1180
35	AROSA	46.77	9.67	1860	1678
40	HAUTE_PROVINCE	43.92	5.75	580	825
43	LERWICK	60.15	-1.15	90	804
51	REYKJAVIK	64.13	-21.9	60	355
53	UCCLE	50.8	4.35	100	61
57	HALLEY_BAY	-75.52	-26.73	31	1531
67	BOULDER	40.02	-105.25	1634	1429
68	BELSK	51.83	20.78	180	1023
84	DARWIN	-12.47	130.83	0	1383
91	BUENOS-AIRES	-34.58	-58.48	25	1841
96	HRADEC_KRALOVE	50.18	15.83	285	1012
99	HOHENPEISSENBERG	47.8	11.02	975	1198
101	SYOWA	-69	39.58	21	615

105	FAIRBANKS	64.8	-147.89	138	681
106	NASHVILLE	36.25	-86.57	182	1147
107	WALLOPS_ISLAND	37.87	-75.52	4	534
111	AMUNDSEN-SCOTT	-89.98	-24.8	2835	717
152	CAIRO	30.08	31.28	35	744
159	PERTH	-31.95	115.85	2	947
175	NAIROBI	-1.27	36.8	1710	562
190	NAHA	26.2	127.67	29	769
191	SAMOA	-14.25	-170.57	82	445
199	BARROW	71.32	-156.6	11	324
208	SHIANGHER	39.77	117	13	713
209	KUNMING	25.02	102.68	1917	575
213	EL_ARENOSILLO	37.1	-6.73	41	634
214	SINGAPORE	1.33	103.88	14	804
216	BANGKOK	13.73	100.57	2	422
232	VERNADSKY-FARADAY	-65.25	-64.27	7	1912
245	ASWAN	23.97	32.45	193	919
252	SEOUL	37.57	126.95	84	895
253	MELBOURNE	-37.48	144.58	125	1676
256	LAUDER	-45.03	169.68	3701	1302
265	IRENE	-25.25	28.22	1524	995
268	ARRIVAL_HEIGHTS	-77.83	166.4	250	414
284	VINDELN	64.25	19.77	0	308
293	ATHENS	38	23.7	15	1263
311	HAVANA	23.17	-82.33	50	668
339	USHUAIA	-54.85	-68.31	7	1426
340	SPRINGBOK	-29.67	17.9	1	1511
341	HANFORD	36.32	-119.63	73	1736
342	COMODORO_RIVADAVIA	-45.78	-67.5	43	1951
343	SALTO	-31.58	-57.95	31	1222
409	HURGHADA	27.25	33.72	22	1541

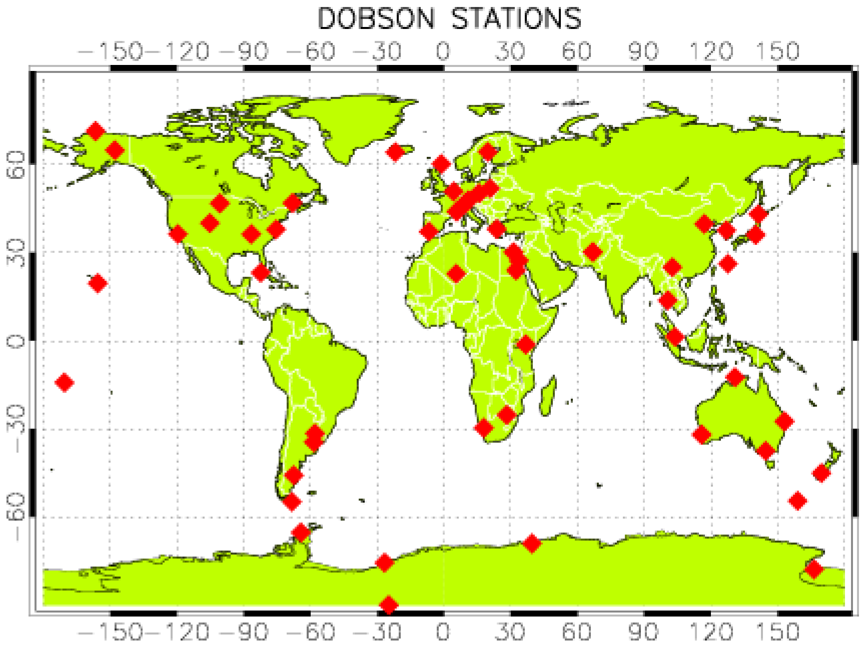
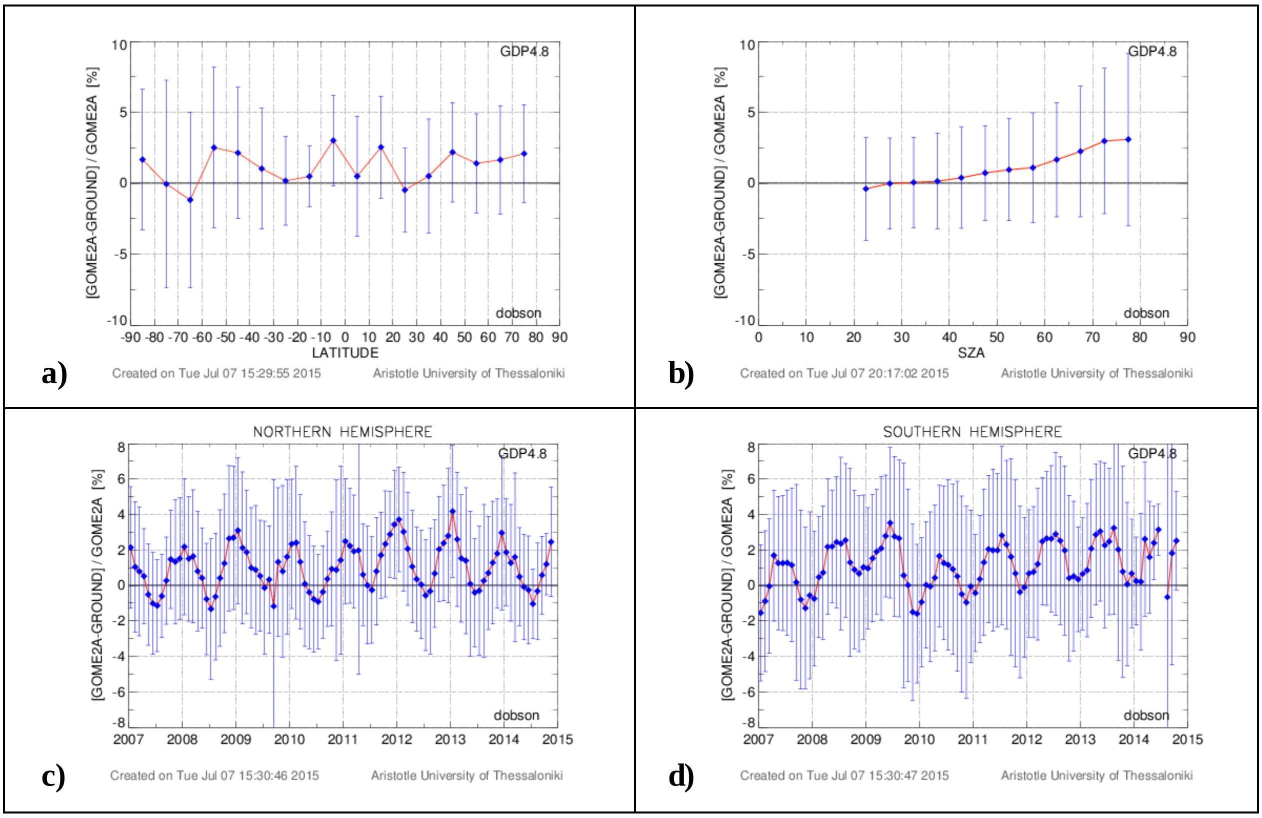


Figure 2.8 Spatial distribution of the Dobson ground-based stations used for the comparisons.

3. GLOBAL AND LATITUDINAL COMPARISONS BETWEEN GOME-2A GDP4.8 AND ARCHIVED GROUND-BASED DATA FROM WOUDC

In the following we aim to show the comparison between the new GDP4.8 total ozone product and co-located ground-based stations (Dobsons and Brewers) for GOME-2A examining possible biases, trends and algorithm effects. A note is required here as to the seaming inconsistency between **Figure 3.1**a, **Figure 3.1**i and **Figure 3.1**j. Different binning systems that have been applied for the different plots: in sub-figure a) the data are binned in 10° bins whereas in sub-figure i) the data are binned in 10° bins and also a contour filter is applied to smooth out spikes and outliers. In sub-figure j) the same binning system results in the need for a very high smoothing out in order for the contour to plot to present meaningful results. We have hence opted to keep the same smoothing between **Figure 3.1**i and **Figure 3.1**j which results in the gap in the Dobson comparisons of **Figure 3.1**j between 10°S and Equator to appear. A similar plotting effect is found in **Figure 4.1**.

3a. GOME-2A comparisons with Dobson instruments



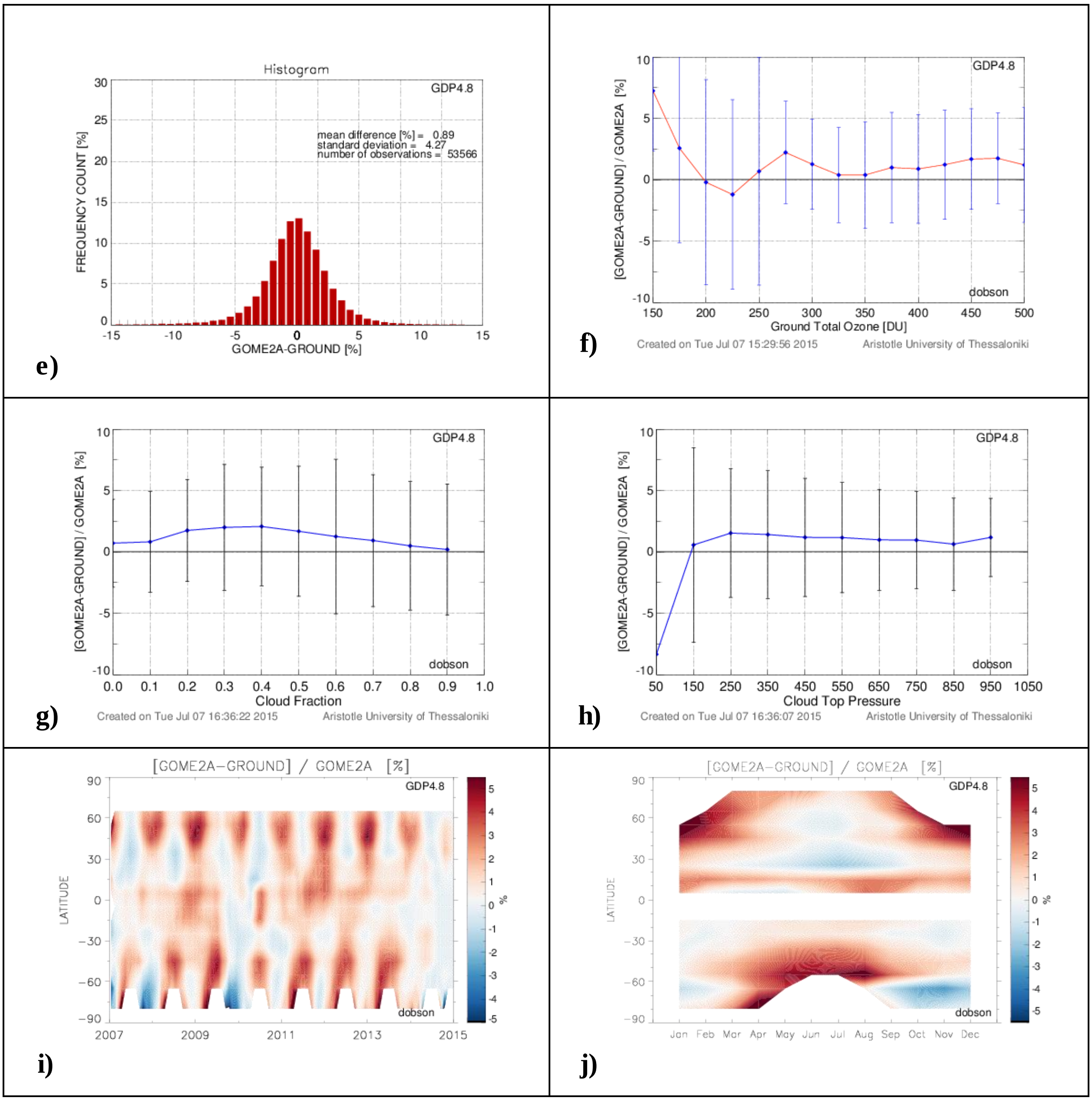


Figure 3.1. The comparison between the GOME-2A and the ground-based Dobson stations is shown in various formats. **a)** Latitudinal average differences between GOME-2A and Dobson instruments based on the period January 2007-December 2014; **b)** Solar zenith angle dependence of the differences; **c)** Time series of the difference between GOME-2A and Dobson instruments for the entire Northern Hemisphere depicted as monthly mean values; **d)** Time series of the difference between GOME-2A and Dobson instruments for the entire Southern Hemisphere depicted as monthly mean values; **e)** Histogram representation of the differences; **f)** Total ozone dependence of the differences; **g)** Cloud fraction dependence of the differences; **h)** Cloud Top Pressure dependence of the differences; **i)** Monthly variations dependence of the differences; **j)** Seasonal variations dependence of the differences.

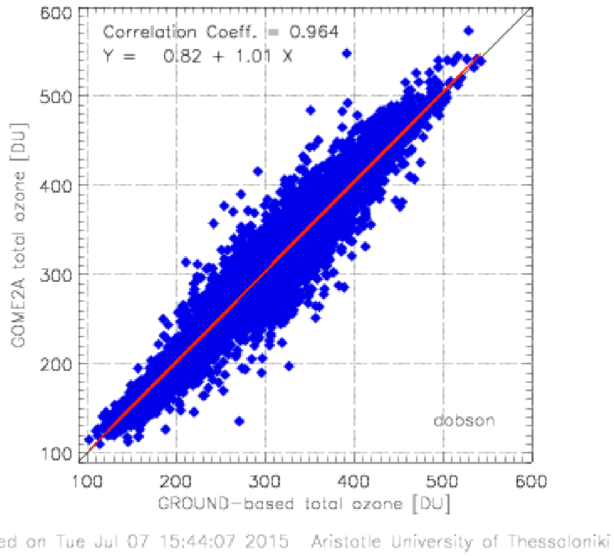


Figure 3.2. Scatter plot of the total ozone values between GOME-2A GDP 4.8 and Dobson instruments.

Discussion on the GOME-2A – Dobson comparisons

a. Based on archived Dobson data for the period January 2007 to December 2014 it is evident from Figure 3.1a (first row, left) that GOME-2A overestimates ozone by about 1.0% on average over all the latitudes of the Northern and Southern Hemisphere. Over the high latitudes of the Northern and Southern Hemisphere GOME-2A overestimates ozone by less than 2.5%. Over the tropical latitudes GOME-2A estimates on the average more ozone by 2.0% compared to Dobson measurements.

b. The time series of the differences over the Northern Hemisphere (Figure 3.1c, second row, left) show that the peak to peak variability is almost stable, about ±2%. In Southern Hemisphere a slight positive trend of the differences is observed from 2010 onwards.

c. Comparisons between GOME-2A and Dobson measurements do not show any dependence on solar zenith angle up to 50°. For SZAs greater than 60°, GOME-2A starts overestimating ozone reaching differences between 2.0 and 3.0%. (Figure 3.1b , first row, right).

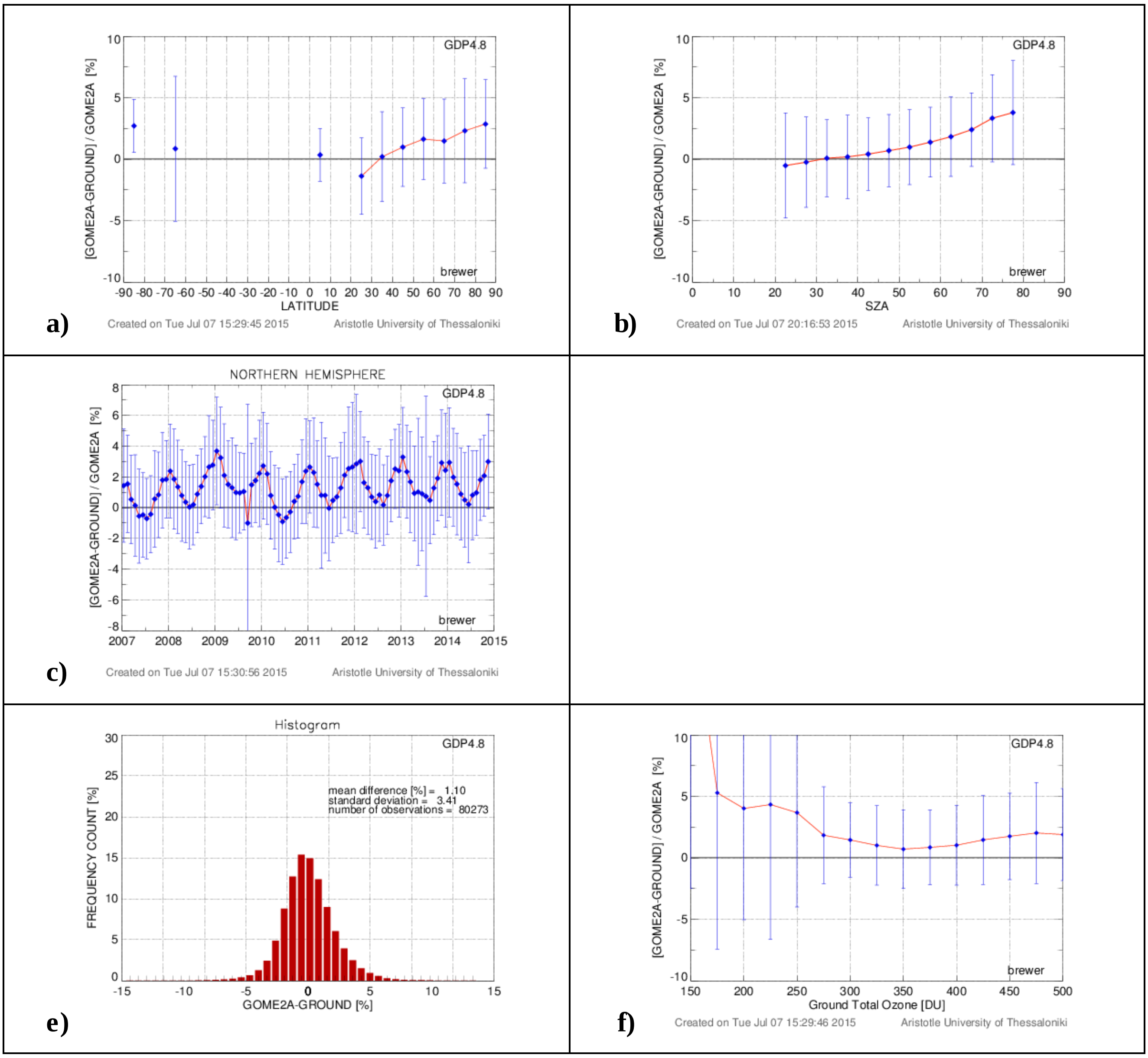
d. The global average difference between GOME-2A and Dobson observations is 0.89±4.27% based on 53566 observations (Figure 3.1e, third row, left). The two data sets show a remarkably high correlation coefficient of 0.964 as it shown in **Figure 3.2** with relatively small scatter.

e. As it is evident form Figure 3.1f, third row, right, there is no significant dependence of the differences on the total ozone content itself except for the comparisons below 200 DU where there is a small amount of co-locations.

f. No significant dependencies of the comparisons against cloud cover and cloud top height were found for this period as it evident from Figure 3.1, fourth row [g & h]. At low cloud top pressures there is small amount of co-locations (less that 50) and the representativeness of that bin is low. Further studies will be performed to understand those co-locations better.

g. There is more seasonal variability of the average difference between GOME-2A and Dobson observations over mid latitudes than over tropics (Figure 3.1i, last row, left) with higher differences during the local winter months (Figure 3.1 j, last row, right) in the NH.

3b. GOME-2A comparisons with Brewer instruments



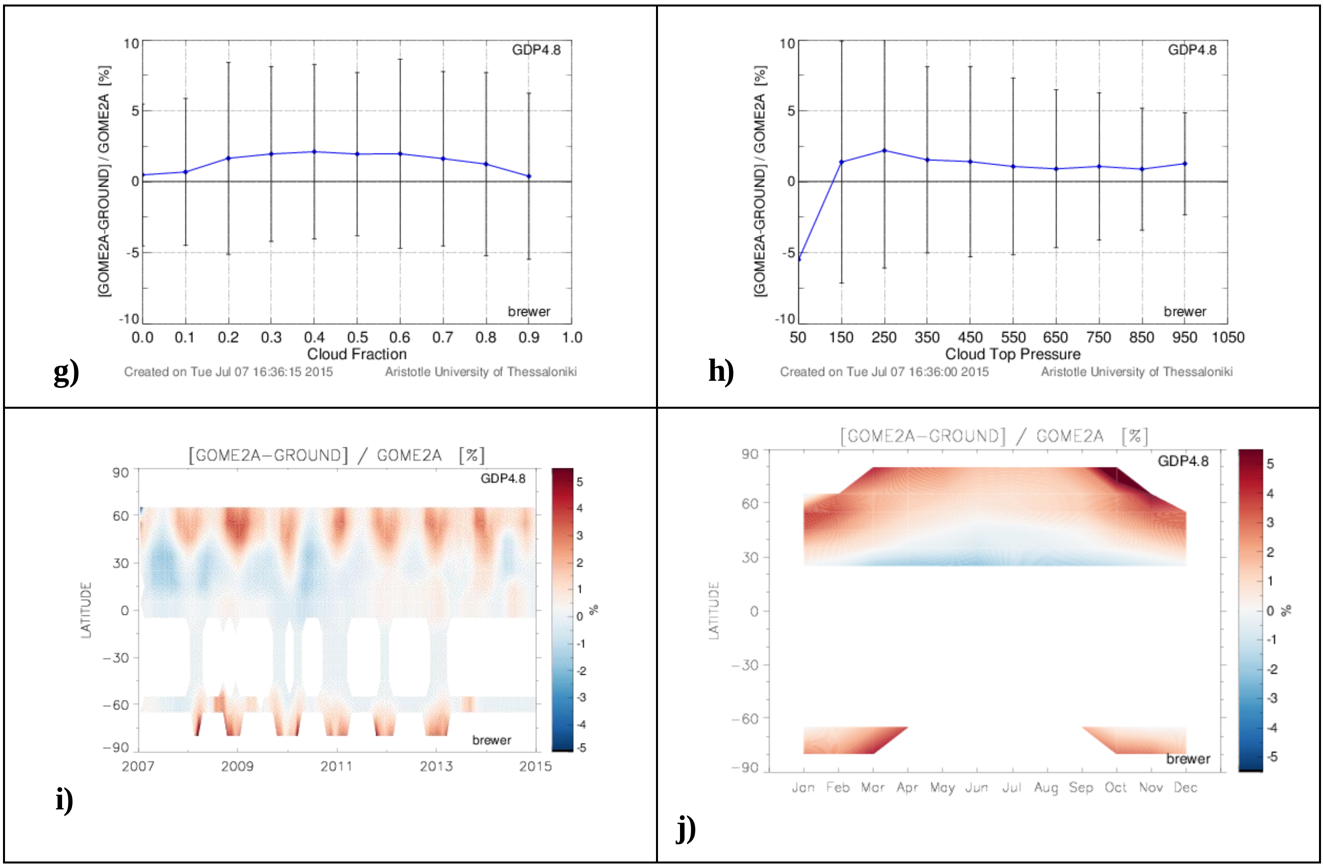


Figure 3.3. The comparison between the GOME-2A and the ground-based Brewer stations is shown in various formats. **a)** Latitudinal average differences between GOME-2A and Brewer instruments based on the period January 2007-December 2014; **b)** Solar zenith angle dependence of the differences; **c)** Time series of the difference between GOME-2A and Brewer instruments for the entire Northern Hemisphere depicted as monthly mean values; **e)** Histogram representation of the differences; **f)** Total ozone dependence of the differences; **g)** Cloud fraction dependence of the differences; **h)** Cloud Top Pressure dependence of the differences; **i)** Monthly variations dependence of the differences; **j)** Seasonal variations dependence of the differences.

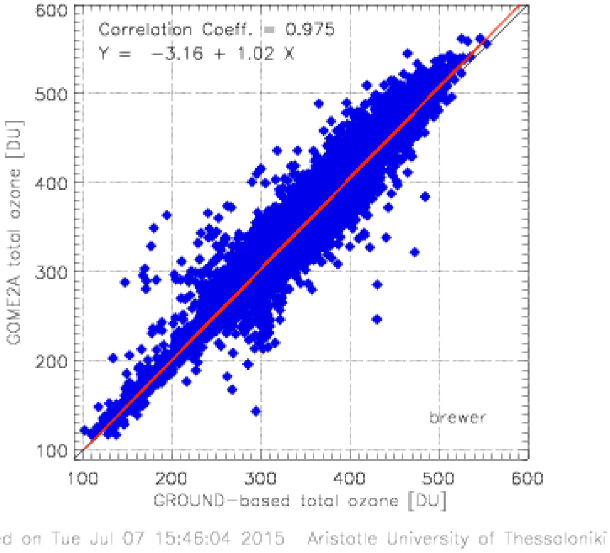


Figure 3.4 The scatter plot of the total ozone values between GOME-2A GDP 4.8 and the Brewer instruments.

Discussion on the GOME-2A – Brewer comparisons

a. Based on archived Brewer data for the period January 2007 to December 2014 it is evident from **Figure 3.3**a, (first row, left) that there is a latitudinal dependency in the NH. GOME-2A shows an overestimation of 1.5% over the Northern Hemisphere, which tends to be slightly higher (2.5-3.0%) over latitudes greater than 60°. There are no enough archived data available for this period for the Southern Hemisphere to draw conclusions from.

b. When examining the time series of the differences over the Northern Hemisphere (**Figure 3.3**c, second row, left) it is shown that the inter-annual variability seasonal dependence is small, with a constant peak-to-peak variability between 0 and 2%.

c. GOME-2A data when compared to Brewer data show a significant dependency on solar zenith angle (**Figure 3.3**b, first row, right) up to 3% that may be attributed to the satellite algorithm.

d. The average difference between GOME-2A and Brewer observations is 1.10±3.41% based however only on 80273 observations (**Figure 3.3**e, third row, left). The two data sets show a remarkably high correlation coefficient of 0.975 as it shown in Figure 3.4 with small scatter.

e. As it is evident from **Figure 3.3**f, third row, right, there is no significant dependence of the differences on the total ozone content itself except for the comparisons below 250 DU where there is a small amount of co-locations.

f. No significant dependencies of the comparisons with the Brewer against cloud cover and cloud top height were found for the certain period as it evident from **Figure 3.3**[g & h], in the fourth row. At low cloud top pressures there is small amount of co-locations (less that 50).

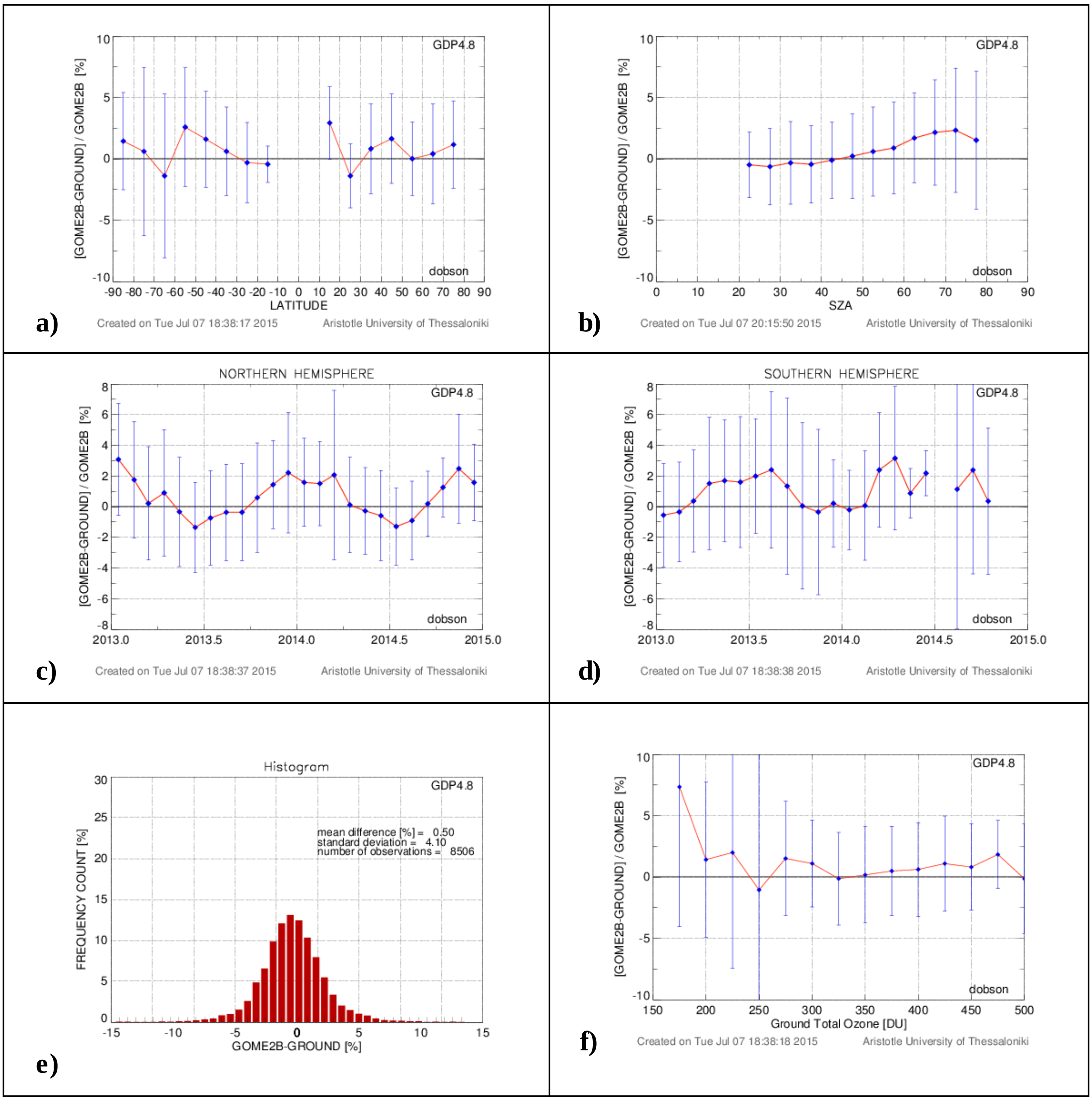
g. There is more seasonal variability of the average difference between GOME-2A and Brewer observations over mid latitudes than over tropics (**Figure 3.3**i, last row, left) with higher differences during the local winter months (**Figure 3.3**j, last row, right) of the NH.

The summary comparisons for the GOME-2A GDP4.8 instrument, both for Dobson and Brewer ground-based co-locations, as well as station and belt statistics, will be available at http://lap3.physics.auth.gr/eumetsat and will cover the period presented in the current report, as soon as this report has been accepted. The Total Ozone Validation Site in the framework of CDOP-2 has been re-designed in a novel dynamic environment with many new interactive and user-defined features.

4. GLOBAL AND LATITUDINAL COMPARISONS BETWEEN GOME-2B GDP4.8 AND ARCHIVED GROUND-BASED DATA FROM WOUDC

In the following we aim to show the comparison between the new GDP4.8 total ozone products and co-located ground-based stations (Dobsons and Brewers) for GOME-2B examining possible biases, trends and algorithm effects.

4a. GOME-2B comparisons with Dobson instruments



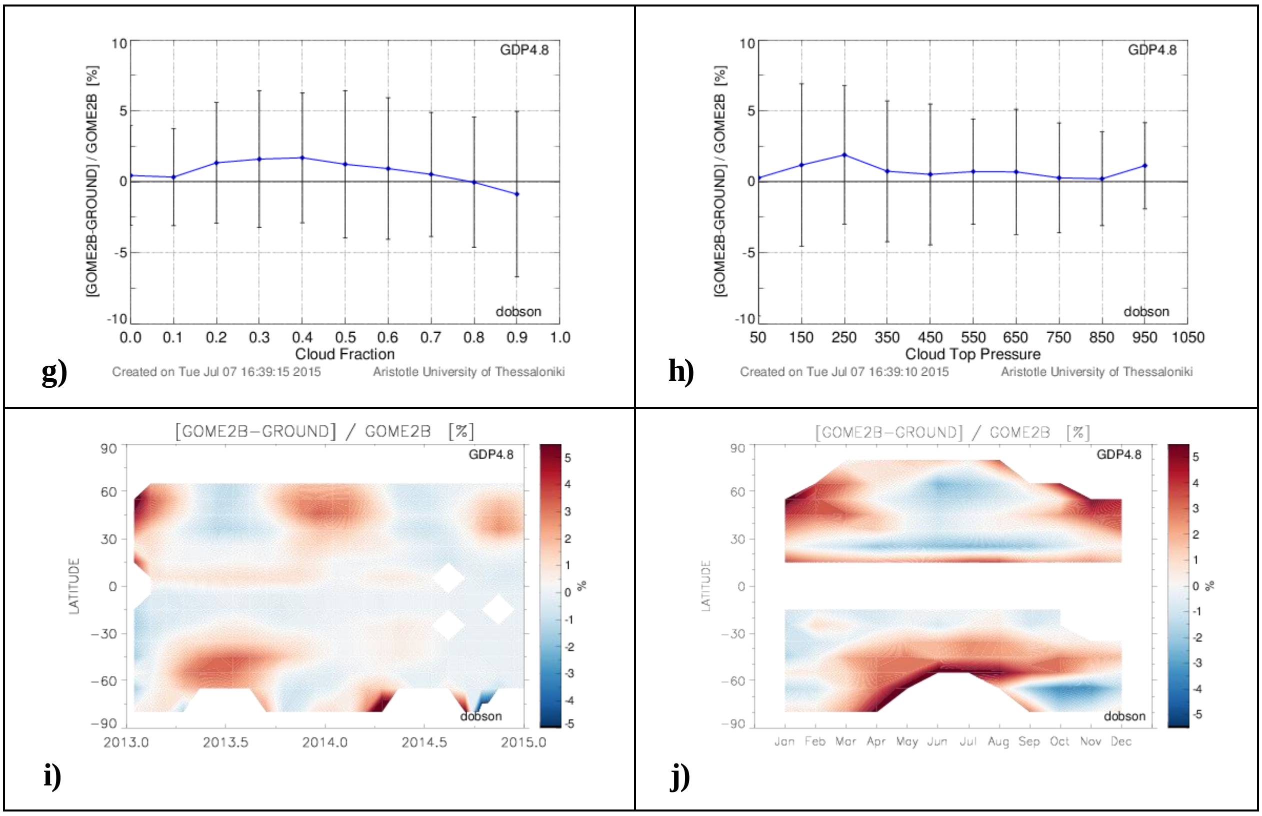


Figure 4.1. The comparison between the GOME-2B and the ground-based Dobson stations is shown in various formats. **a)** Latitudinal average differences between GOME-2B and Dobson instruments based on the period January 2013-December 2014; **b)** Solar zenith angle dependence of the differences; **c)** Time series of the difference between GOME-2B and Dobson instruments for the entire Northern Hemisphere depicted as monthly mean values; **d)** Time series of the difference between GOME-2B and Dobson instruments for the entire Southern Hemisphere depicted as monthly mean values; **e)** Histogram representation of the differences; **f)** Total ozone dependence of the differences; **g)** Cloud fraction dependence of the differences; **h)** Cloud Top Pressure dependence of the differences; **i)** Monthly variations dependence of the differences; **j)** Seasonal variations dependence of the differences.

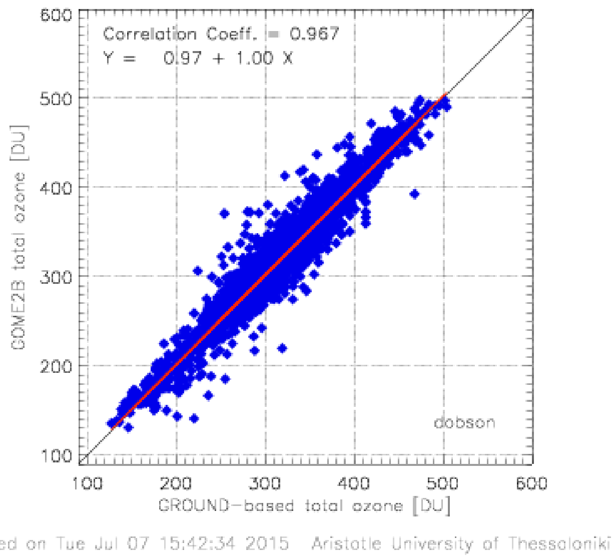


Figure 4.2. Scatter plot of the total ozone values between GOME-2B GDP 4.8 and Dobson instruments.

Discussion on the GOME-2B – Dobson comparisons

a. Based on archived Dobson data for the period January 2013 to December 2014 it is evident from **Figure 4.1**a, in the first row, left, that GOME-2B overestimates ozone by about 1.5% on average over almost all the latitudes of the Northern and Southern Hemisphere. Over the high latitudes of the Northern and Southern Hemisphere GOME-2B overestimates ozone by less than 1.5%. Over the middle latitudes GOME-2B estimates on the average more ozone by 2.0% compared to Dobson measurements. Over tropical latitudes there are no ground-based observations for the two year time period available to draw conclusions from.

b. The time series of the differences over the Northern Hemisphere (**Figure 4.1**c, second row, left) show that the peak to peak variability is almost stable, about ±2%, in Northern Hemisphere.

c. GOME-2B data, when compared to Dobson measurements do not show any dependence on solar zenith angle up to 50°. (**Figure 4.1**b, first row, right). For SZAs greater than 60^{o}, GOME-2B starts overestimating ozone reaching values of 2.0 – 2.5%.

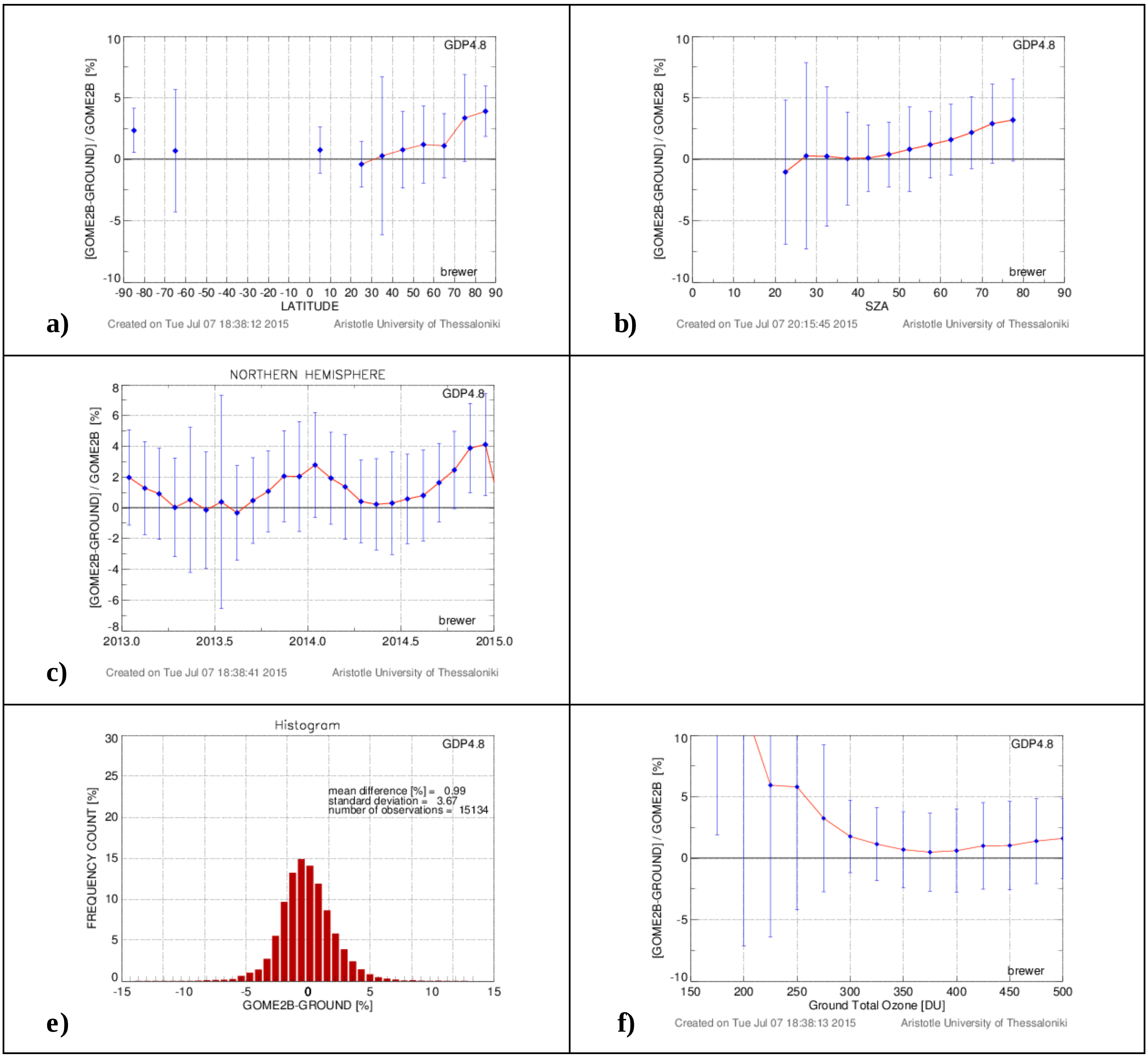
d. The average difference between GOME-2B and Dobson observations is 0.50±4.10% based on 8506 observations, **Figure 4.1**e, third row, left). The two data sets show a remarkably high correlation coefficient of 0.967, as shown in Figure 4.2, with small scatter.

e. As it is evident from **Figure 4.1**f, third row, right, there is no significant dependence of the differences on the total ozone content itself except for the comparisons below 200 DU where there is small amount of co-locations (less that 20).

f. No significant dependencies of the comparisons against cloud cover and cloud top height were found for this period as it evident from, **Figure 4.1**g & h, in the fourth row.

g. There is more seasonal variability of the average difference between GOME-2B and Dobson observations over mid latitudes than over tropics, (**Figure 4.1**i, last row, left) with higher values during the local winter months (**Figure 4.1**j, last row, right) in the NH.

4b. GOME-2B comparisons with Brewer instruments



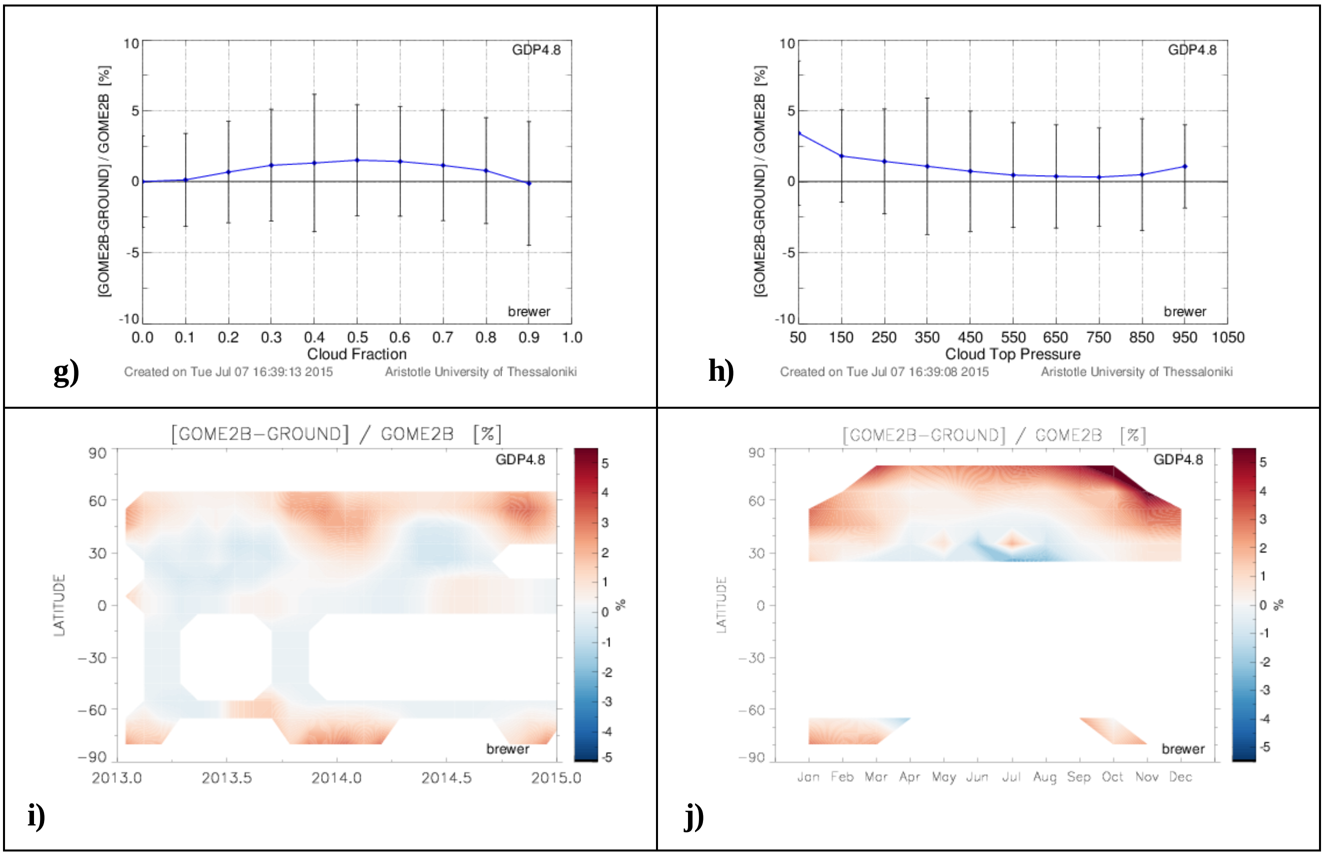


Figure 4.3. The comparison between the GOME-2B and the ground-based Brewer stations is shown in various formats. **a)** Latitudinal average differences between GOME-2B and Brewer instruments based on the period January 2013-December 2014; **b)** Solar zenith angle dependence of the differences; **c)** Time series of the difference between GOME-2B and Brewer instruments for the entire Northern Hemisphere depicted as monthly mean values; **e)** Histogram representation of the differences; **f)** Total ozone dependence of the differences; **g)** Cloud fraction dependence of the differences; **h)** Cloud Top Pressure dependence of the differences; **i)** Monthly variations dependence of the differences; **j)** Seasonal variations dependence of the differences.

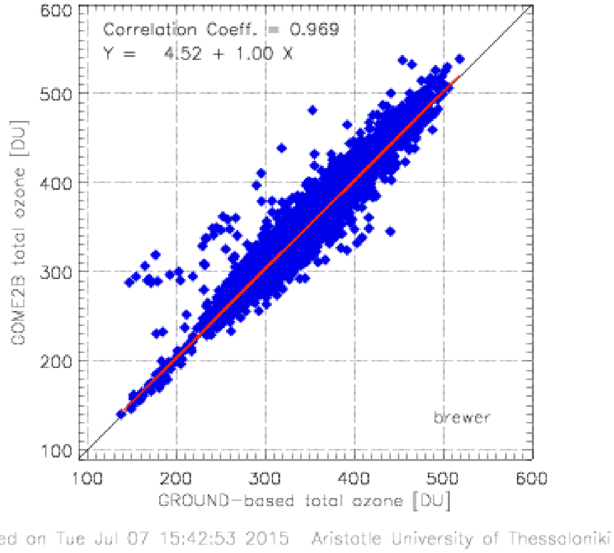


Figure 4.4. The scatter plot of the total ozone values between GOME-2B GDP 4.8 and the Brewer instruments.

Discussion on the GOME-2B – Brewer comparisons

a. Based on archived Brewer data for the period January 2013 to December 2014 it is evident from Figure 4.3**a** (first row, left) that GOME-2B over the Northern Hemisphere show a slight overestimation of 0.5%. There are no enough archived data available for this period for the Southern Hemisphere to draw conclusions.

h. When examining the time series of the differences over the Northern Hemisphere (Figure 4.3c, second row, left) it is shown that the inter-annual variability seasonal dependence is small, with a constant peak-to-peak variability between 0 and 2%.

b. GOME-2B data when compared to Brewer data do not show a significant dependence on solar zenith angle (Figure 4.3b, first row, right) up to 50^{o}. Higher estimates of ozone for SZAs greater than 50^{o} are through to be due to the satellite algorithm.

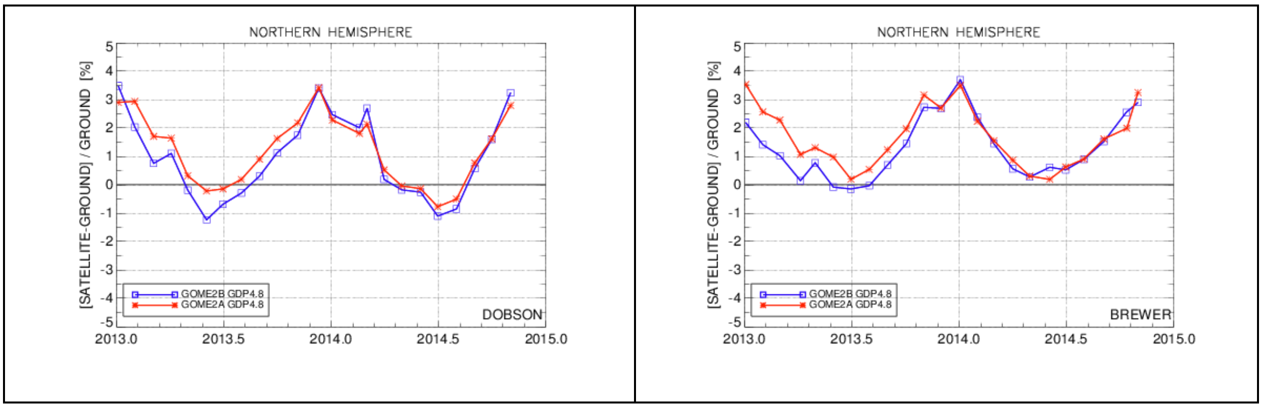
c. The average difference between GOME-2B and Brewer observations is 0.99±3.67% based however only on 15134 observations (Figure 4.3e, third row, left). The two data sets show a remarkably high correlation coefficient of 0.969 as it shown in Figure 4.4 with small scatter.

d. As it is evident from Figure 4.3f, third row, right, there is no significant dependence of the differences on the total ozone content itself except for the comparisons below 275 DU where there is small amount of co-locations (less that 15).

e. No significant dependencies of the comparisons with the Brewer against cloud cover and cloud top height were found for the certain period as it evident from Figure 4.3g & h, fourth row.

5. TIME SERIES OF GOME-2A AND GOME-2B GDP4.8 TOTAL OZONE FOR COMMON DAYS AGAINST SELECTED GROUND-BASED LOCATIONS AND COMPARISON TO THE GDP4.7 ALGORITHM OVER THE NORTHERN HEMISPHERE

The GOME-2A and GOME-2B products for the period examined have been processed both with the GDP4.8 and GDP4.7 versions of the algorithm. To check the consistency of between these two data sets these have been compared with ground-based data for the period 2013-2014. In this section onwards, common co-locations are shown, i.e. days when all three instruments provide a valid TOC estimate. The results for the Northern Hemisphere are shown in Figure 5.1 below with the Dobson comparisons on the left and the Brewer comparisons on the right. In the first row, the GOME-2A [red] and GOME-2B [blue] are both analyzed with the GDP 4.8 algorithm. For the entire year 2013 there is a difference of around 0.5-1% between GOME-2A and GOME-2B GDP4.8 for both Dobson and Brewer instruments. This difference cannot be attributed to the ground-based network and should be investigated further. In the second row, the GOME-2A [red] and GOME-2B [blue] are both analyzed with the GDP 4.7 algorithm. GDP4.7 does not appear to introduce differences between the two sensors apart from some issues around the winter of 2014 of around 0.5% again present in both the Brewer and the Dobson comparisons. In the third row, the GOME-2B GDP4.7 [red] is compared to the GOME-2B GDP4.8 [blue] it is shown that a constant bias is introduced in the new version of the data of between 0.5 and 1% for the Dobson coincidences and between 0 and 0.5% for the Brewer comparisons for the entire two years period. In the fourth row, the GOME-2A GDP4.7 [red] is compared to the GOME-2A GDP4.8 [blue] where, surprisingly, the Brewer comparisons show the higher deviation of between 0.5 and 1% whereas the Dobson comparisons show differences mainly in the winter months of 2010 to 2012 inclusive and a near-perfect agreement for the rest of the time period.



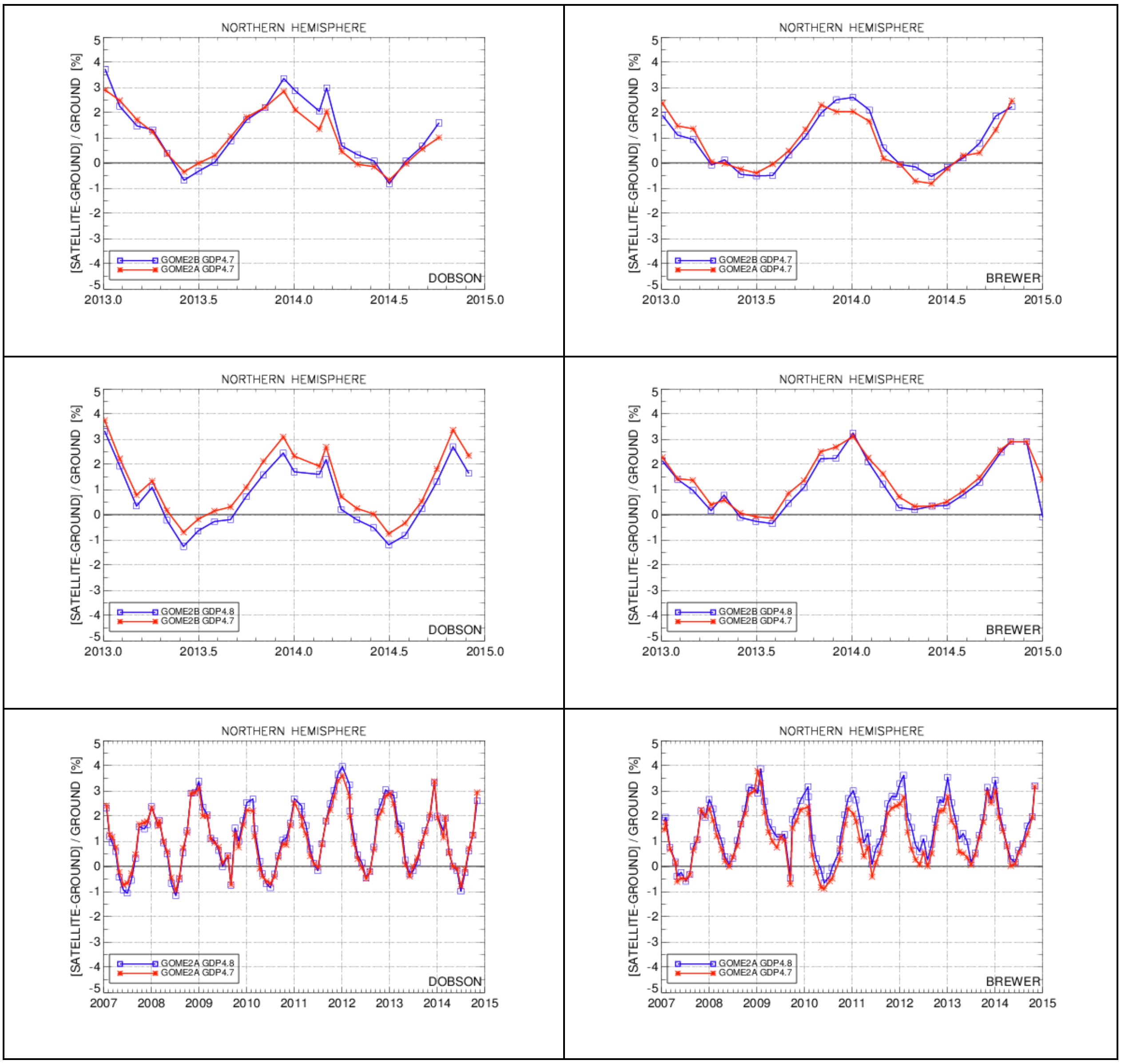


Figure 5.1. Time series of the monthly mean differences between GOME-2A and GOME-2B with GDP 4.7 and GDP 4.8 algorithms (using common co-locations) and ground-based data Dobson (left column) and Brewer (right column) geolocations for the entire Northern Hemisphere.

From this Figure, we may conclude that, even though a change is introduced in both GOME-2A and GOME-2B when moving from the current GDP4.7 algorithm to the new GDP4.8 version, these changes are producing a constant bias and no discernible dependencies in time.

In order to explain the discrepancy observed between the Dobson and Brewer comparisons shown in the third and fourth column of Figure 5.1, in both left and right column, whereupon it appears as though GDP4.8 overestimates GDP4.7 for GOME-2A but under-estimates for GOME-2B, we present belt statistics below and hence demonstrate that this issue is possibly both a simple sampling issue as well as an issue of the different implementation of GDP4.8 onto GOME-2A and GOME-2B.

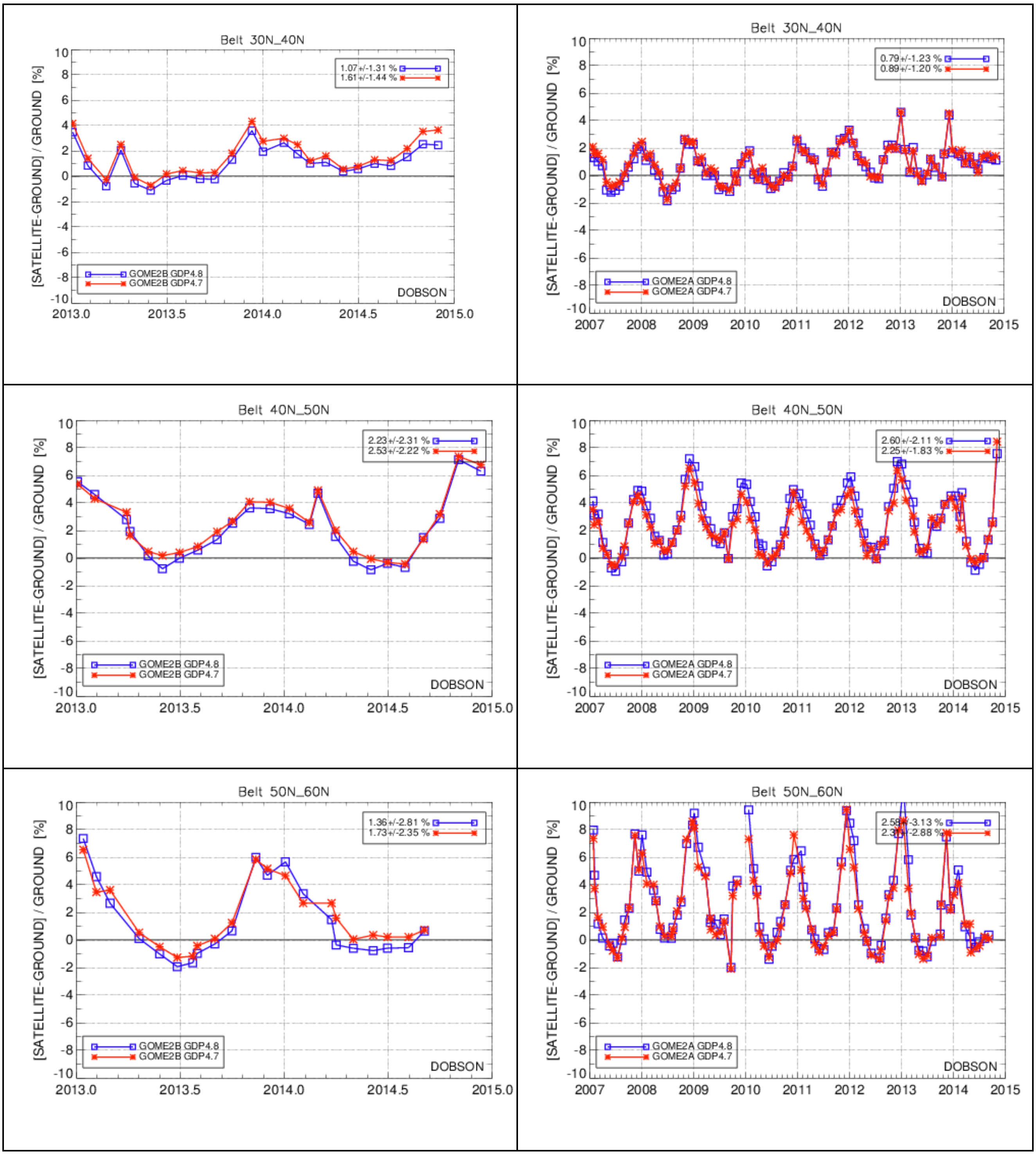


Figure 5.2. Time series of the differences between GOME-2B GDP 4.8 [blue line] and GDP4.7 [red line] for the Dobson [left column] and GOME-2A GDP 4.8 [blue line] and GDP4.7 [red line] for the Dobson [right column.] Top panel : the 30° to 40°N belt, middle panel : the 40° to 50°N belt and bottom panel : the 50° to 60°N belt. The mean value and associated std are also given in the upper right corner.

As noted from Figure 5.2, left column, where the Dobson to GOME-2B comparisons are shown, for all belts the new algorithm GDP4.8 [blue lines] produces smaller differences to the ground-based data than the GDP4.7 algorithm by around 0.3-0.4%, as was also seen in Figure 5.1, left, middle, for the entire hemisphere mean. From Figure 5.2, right column, it is shown that for GOME-2A there are belts where GDP4.7 produces smaller differences to the Dobson spectrophotometers [such as the top row showing the 30° to 40°N belt] whereas there are other belts, which represent a larger number of instruments, such as the 40° to 50°N the 50° to 60°N belts, where the opposite is true. As a result, in the entire hemisphere mean [see Figure 5.1, left, bottom], GOME-2A GDP4.8 appears to produce larger differences than GDP4.7 of around 0.2 to 0.5% depending on the year. These belt statistics are well in-line with what is also shown in Sections 3 and 4 where the global statistics for one day in the year are shown.

The take-away message of this section is that, even though it would appear that there are differences in the validation results between GOME-2A and GOME-2B GDP4.8 these differences are small and should not be the decisive factor for the operational implementation of GDP4.8 for total ozone studies.

6. LATITUDINAL, SZA AND SEASONALCOMPARISONS BETWEEN GOME-2A AND GOME-2B, GDP4.8 AND GDP4.7 ALGORITHMS, FOR COMMON DAYS AGAINST SELECTED GROUND-BASED LOCATIONS

In this section we show comparisons with ground based data (separately for the Dobsons on the left column and the Brewers on the right column) of common days of GOME-2A and GOME-2B, examining possible algorithm effects and biases on the two instruments introduced by both of the two algorithm versions, GDP4.8 and GDP4.7.

In Figure 6.1, the latitudinal dependence of the differences is given. The same findings as discussed for Figure 5.1, first row, may be applied here as well. In general, the bias between GOME-2A and GOME-2B for the SH is larger than the bias for the NH [top row, left] with the differences between the two sensors reach more than 1.5% for certain bins, whereas in the NH Dobson comparisons the differences fall to 0.5%, similar to what is observed for the Brewer comparisons on the top row, right. When moving from GDP4.7 to GDP4.8 [middle row], GOME-2B exhibits the larger differences in the Antarctic latitudes and the Dobson NH tropics and middle latitudes, with however small biases around 0.5%. Similarly for GOME-2A, the move from GDP4.7 to GDP4.8 [bottom row] produces small changes in the Antarctic but also in the Brewer NH middle latitudes of between 0.5 and 1%.

In Figure 6.2, the SZA dependence of the differences is given. Apart from the bias between algorithms and sensors already discussed which may differ from 0 to 1% depending on the sensor and algorithm combination, no other pertinent features can be revealed by this investigation.

In Figure 6.3, the seasonal variability of the differences averaged over the Northern and the Southern Hemisphere is given for the three cases. Again, no obvious features can be deduced from these Figures, apart from the known seasonal variability mainly due to the ground-based instrumentation which appears to be of the same order of magnitude for the two types of instruments.

Latitudinal variability

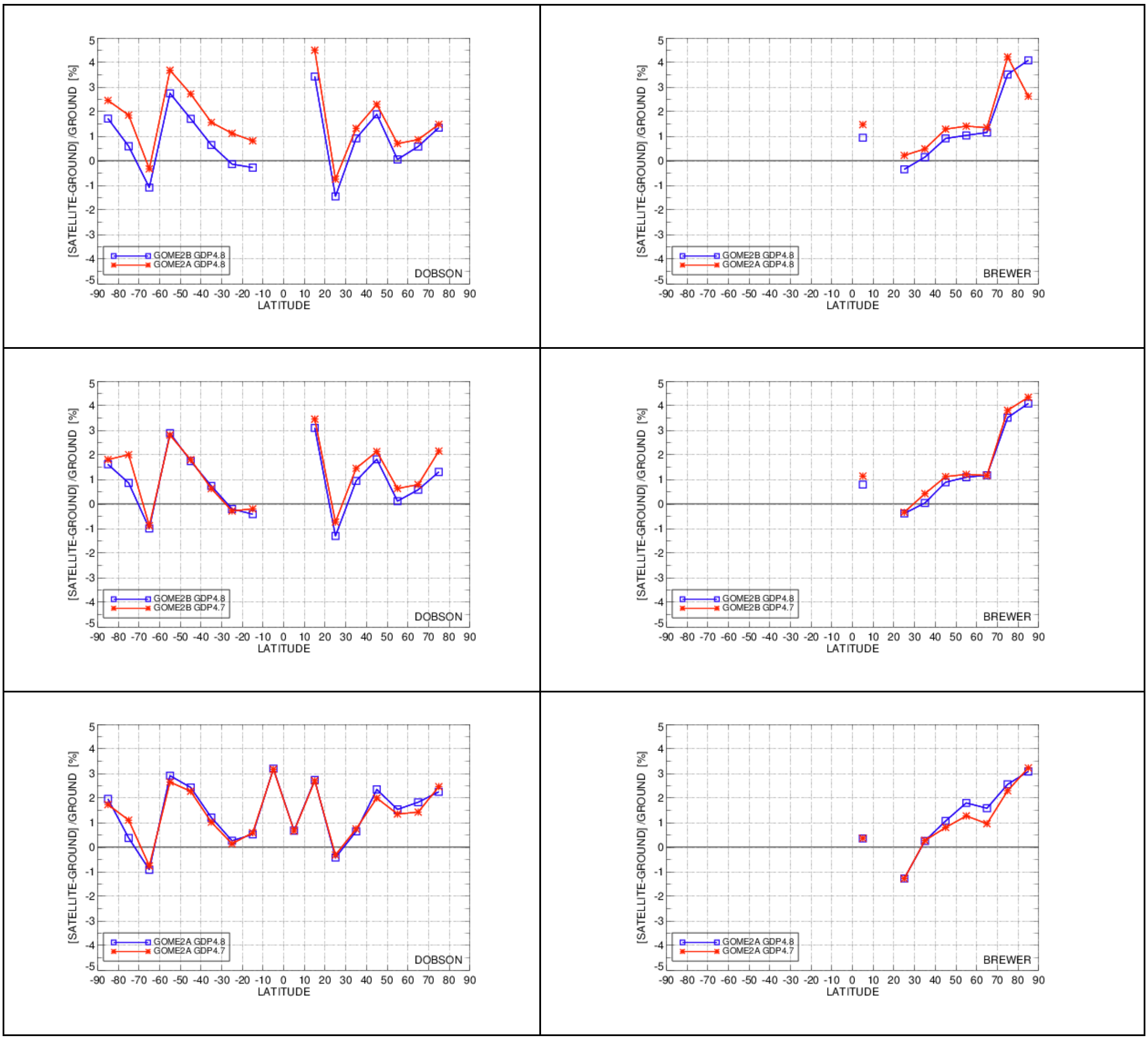


Figure 6.1. Latitudinal dependence of the differences between coincidences between satellite and Dobson [left] and Brewer spectrophotometers [right]. **Top row:** the effect of the GDP4.8 algorithm on GOME-2A [red] and GOME-2B [blue]. **Middle row:** GOME-2B comparisons with GDP4.7 [red] and GDP4.8 [blue]. **Bottom row:** GOME-2A comparisons with GDP4.7 [red] and GDP4.8 [blue].

Solar Zenith Angle variability

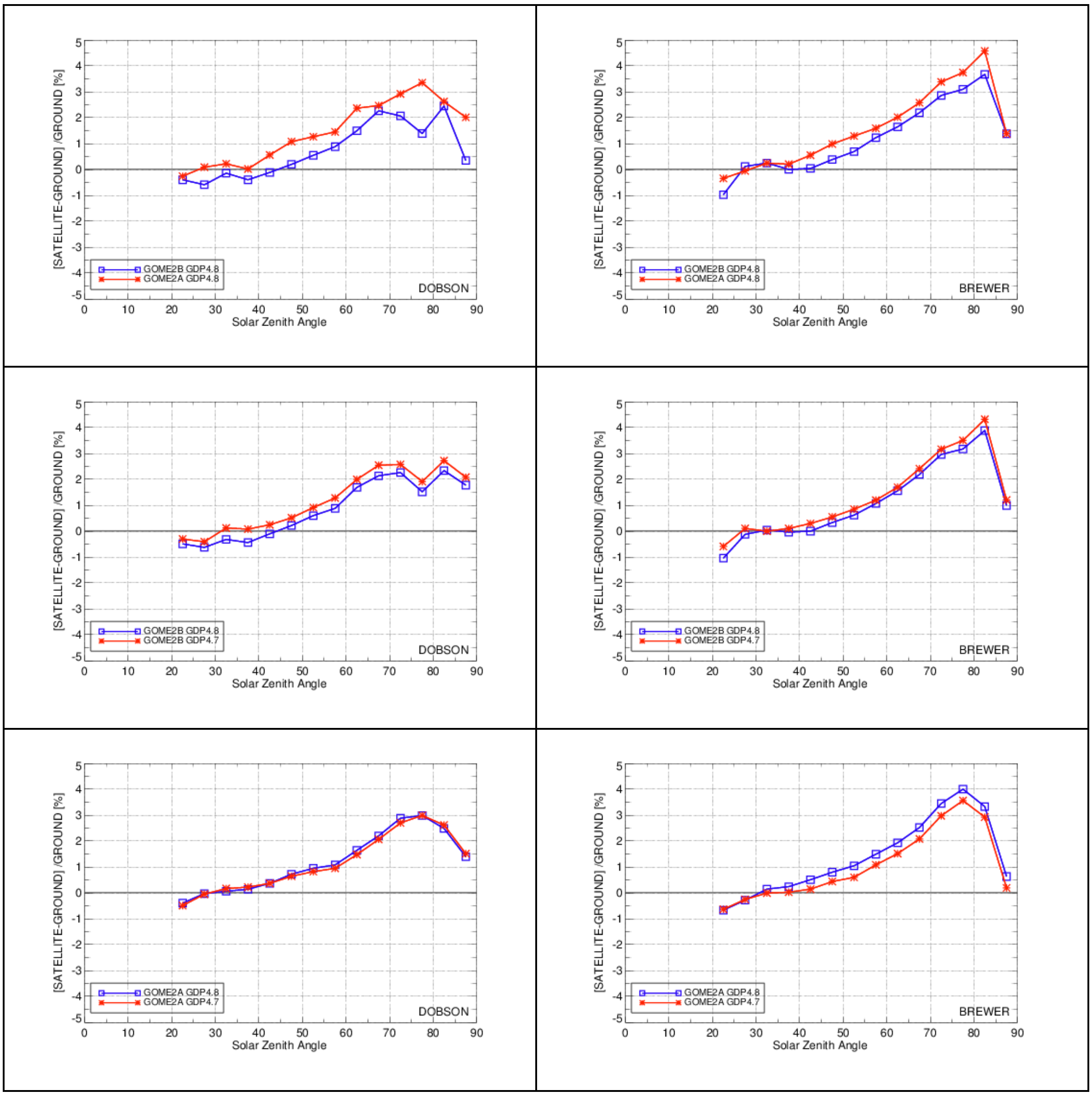
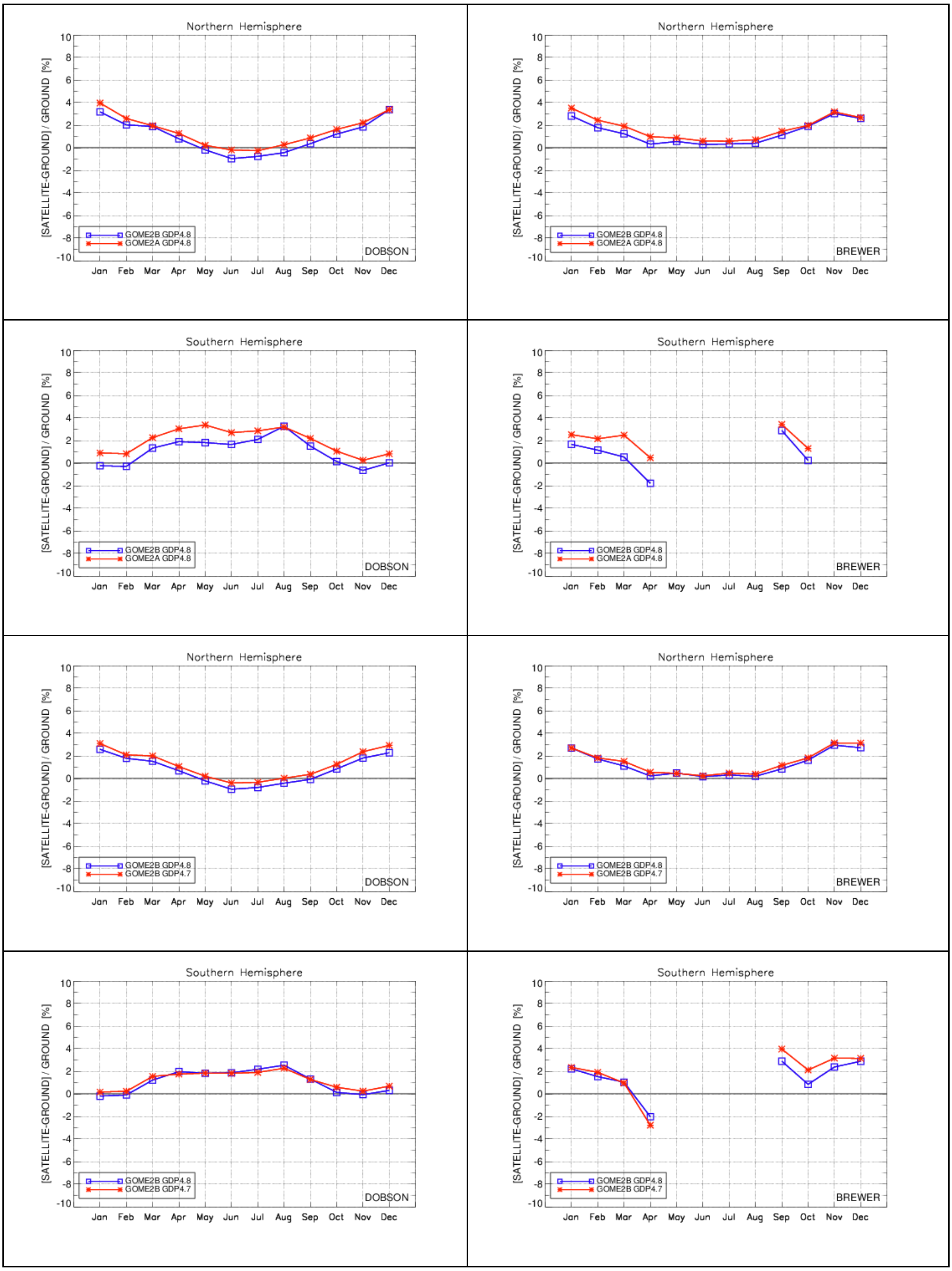


Figure 6.2. Solar Zenith Angle dependency of the differences between coincidences between satellite and Dobson [left] and Brewer spectrophotometers [right]. **Top row:** the effect of the GDP4.8 algorithm on GOME-2A [red] and GOME-2B [blue]. **Middle row:** GOME-2B comparisons with GDP4.7 [red] and GDP4.8 [blue]. **Bottom row:** GOME-2A comparisons with GDP4.7 [red] and GDP4.8 [blue].

Seasonal variability



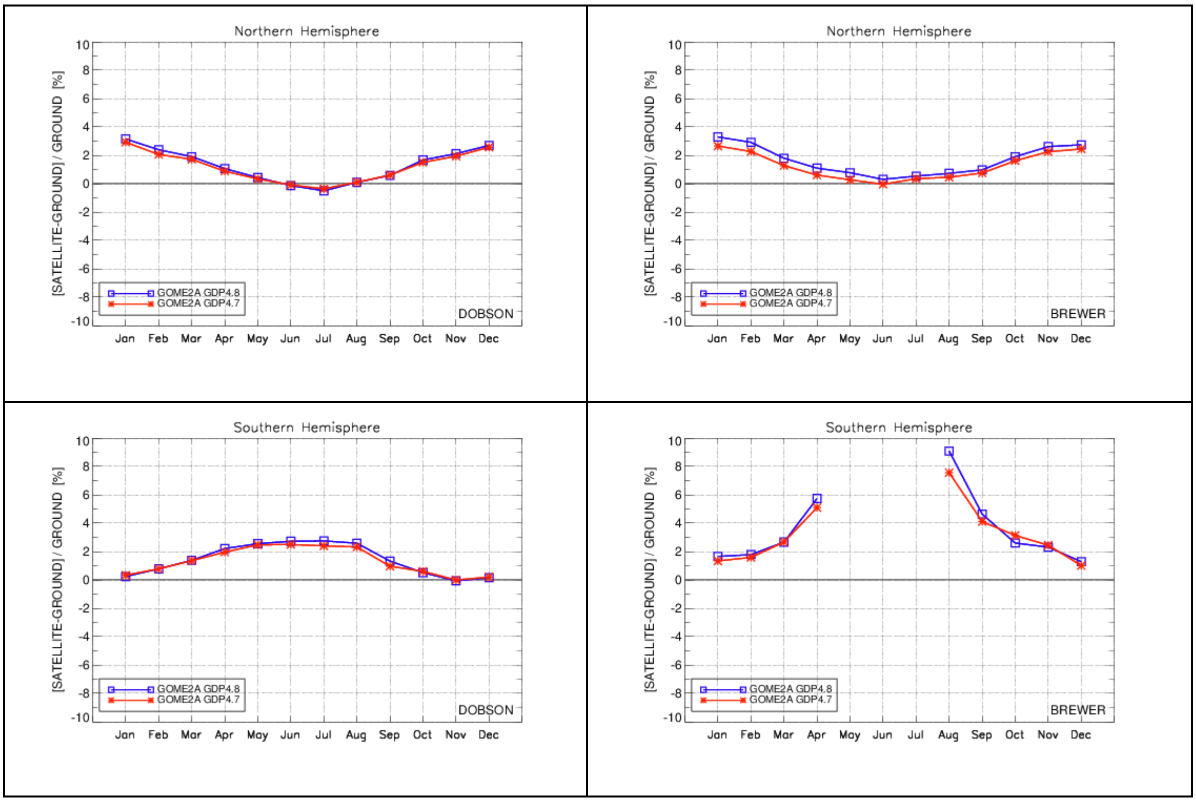


Figure 6.3. Seasonal dependence of the differences between coincidences between satellite and Dobson [left] and Brewer spectrophotometers [right]. **First row:** the effect of the GDP4.8 algorithm on GOME-2A [red] and GOME-2B [blue] for the NH. **Second row:** Same as the first row for the SH. **Third row:** GOME-2B comparisons with GDP4.7 [red] and GDP4.8 [blue] for the NH. **Fourth row:** Same as the third row for the SH. **Fifth row:** GOME-2A comparisons with GDP4.7 [red] and GDP4.8 [blue] for the NH. **Sixth row:** Same as the fifth row for the SH.

The seasonal variability for the NH (Figure 6.3, first row) and the SH (Figure 6.3, second row) is also shown where, for both sensors, the agreement to the ground is best and around 0% during the local summer months of May, June, July and August for the N.H. and November, December and January for the S.H. This difference may reach the 5% difference level for the local winter months of the SH for GOME-2A [bottom right, Brewer comparisons] however note that there are only three Brewer stations considered for the SH. The main findings remain the same for both Dobson and Brewer comparisons.

Scan angle dependency

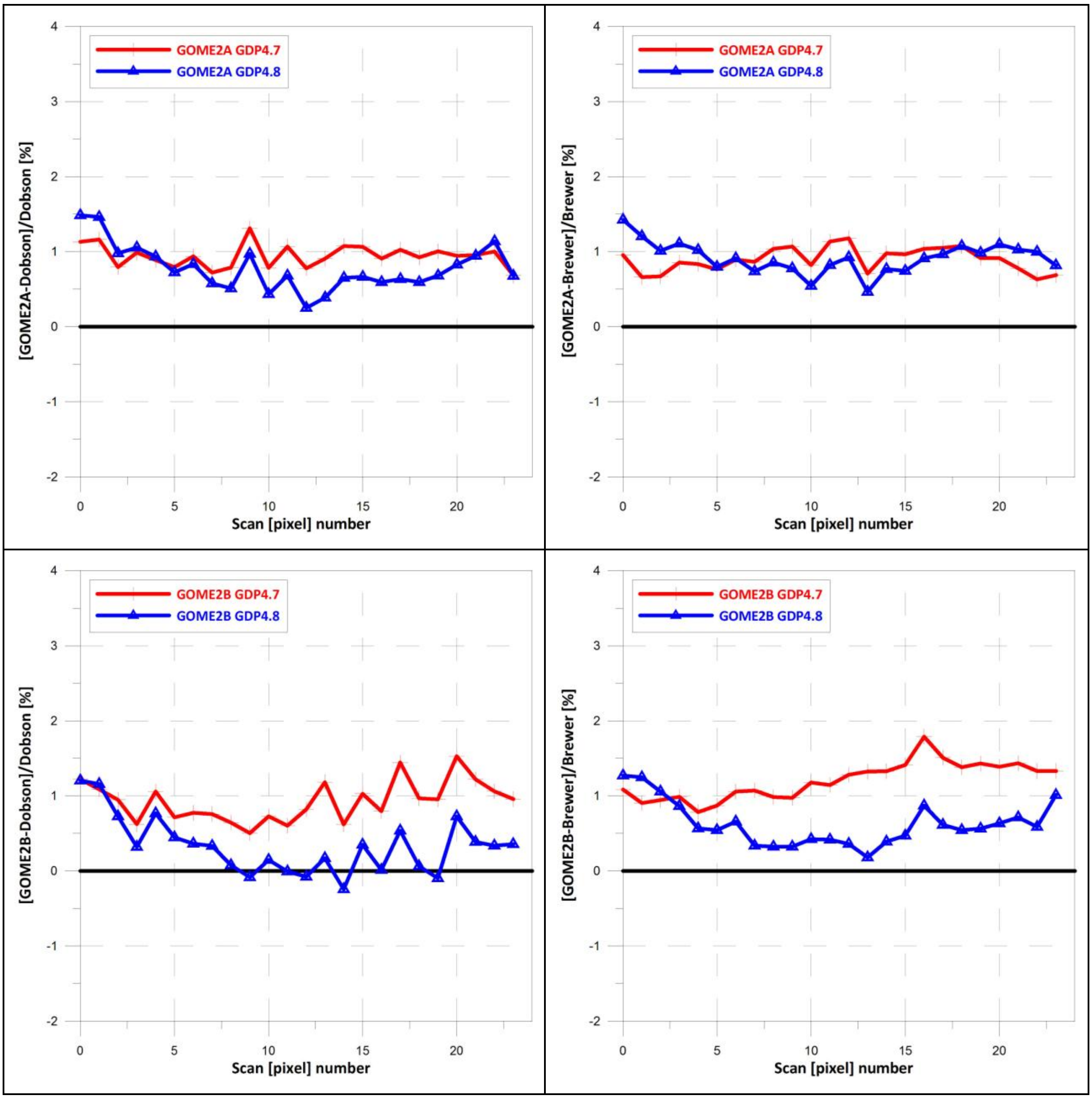
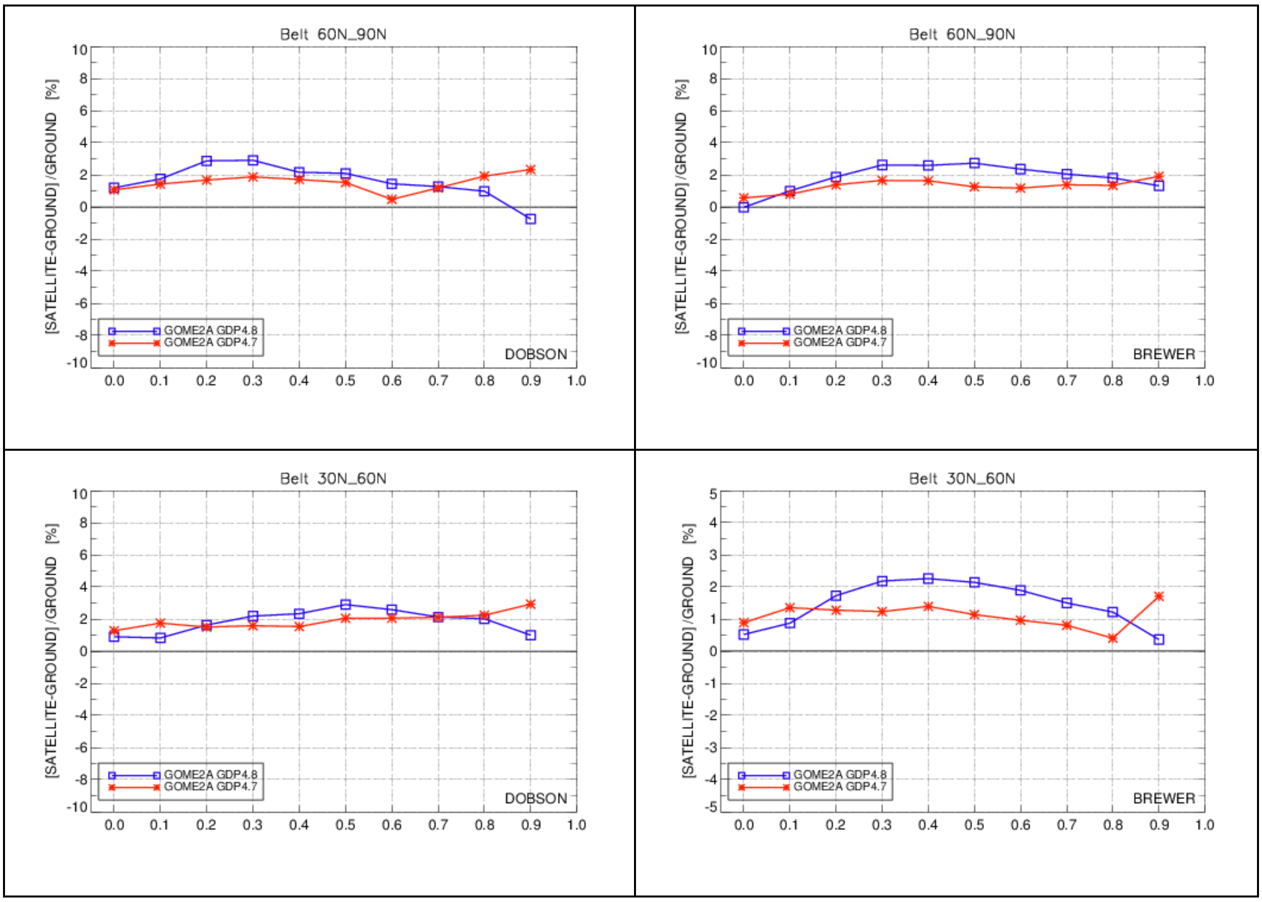


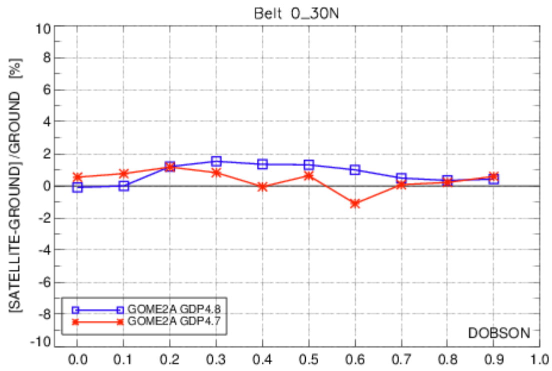
Figure 6.4. The scan angle dependency for the GOME2A [top row] and the GOME2B [bottom row] comparisons to the Dobson network [left column] and the Brewer network [right column.] The red lines depict the GDP4.7 TOCs and the blue lines the GDP4.8 TOCs. Only forward scans are shown.

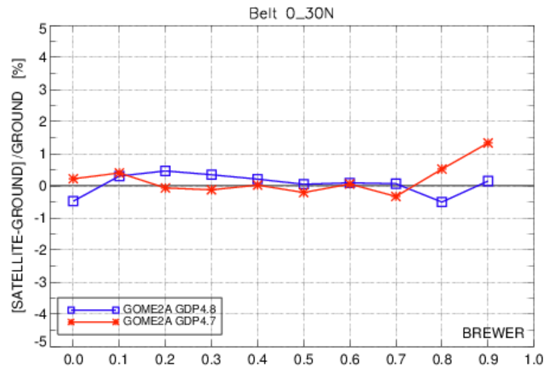
In Figure 6.4, the scan angle dependency for GDP4.7 [in red] and GDP 4.8 [in blue] is shown for the GOME2A [top row] and the GOME2B [bottom row] comparisons to the Dobson network [left column] and the Brewer network [right column.] Note that the y-axis span from -2 to 4% only. Applying the GDP4.8 algorithm on the GOME2B observations seems to improve the scan angle dependency to a larger extend than for the GOME2A, for both Dobson and Brewer collocations.

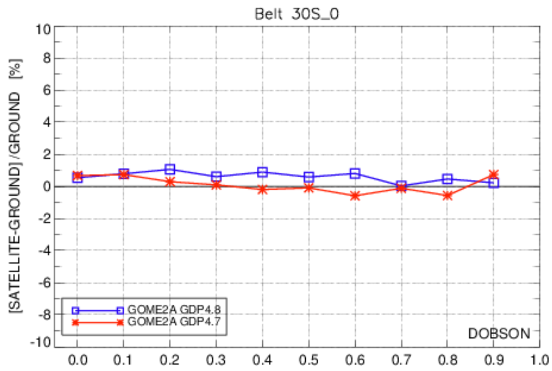
7. CLOUD PARAMETER DEPENDENCY

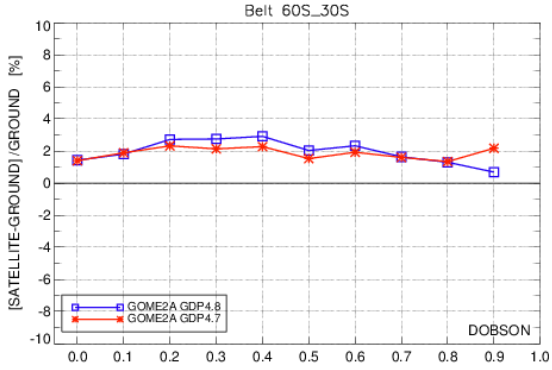
In this section we show the cloud fraction and cloud top pressure dependence of the GOME-2 TOC difference between the two versions of the algorithm for each sensor compared to Dobson and Brewer measurements, in order to study the effect of the new cloud (version 3.0) algorithms applied to GDP4.8 data sets [Lutz et al., 2015]. Figure 7.1 shows the cloud fraction dependence of the TOC difference for GOME-2A for the GDP 4.8 and 4.7, and Figure 7.2 shows the cloud fraction dependence for GOME-2B. Note that, as for the Figures above where both sensors are shown on the same plot, if not stated otherwise, only common coincidences between all three GOME-2A, GOME-2B and ground-based spectrophotometer are shown. Comparing Figure 7.1, the comparisons for GOME-2A GDP4.7 and GDP4.8,to Figure 7.2, the comparisons for GOME-2B GDP4.7 and GDP4.8, it is obvious that the scatter is far larger for GOME-2A for most belts than for GOME-2B. This is possibly due to the larger degradation effects of GOME2A demonstrated in Figures shown above as well. In general, we note that for most belts, the two algorithms shown approximately the same behaviour with apparently random differences that cannot point in one direction or the other.

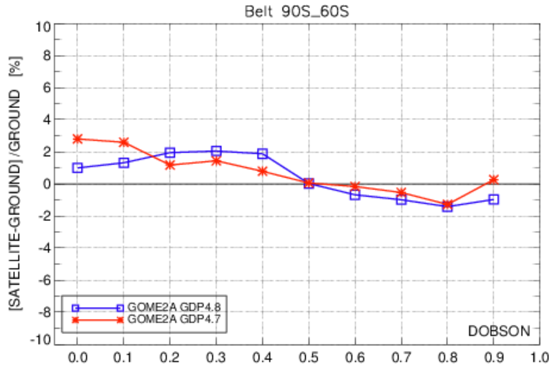



Belt 0_30N
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 0_30N
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
BREWER


Belt 30S_0
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 60S_30S
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 90S_60S
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON

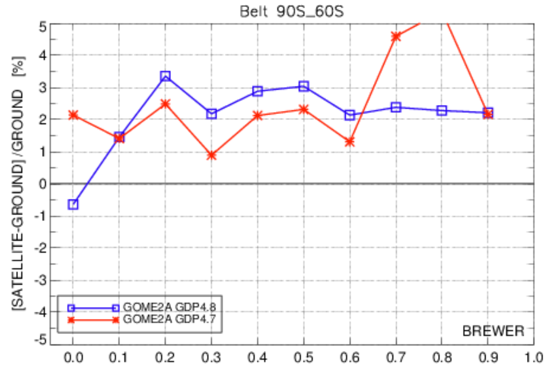
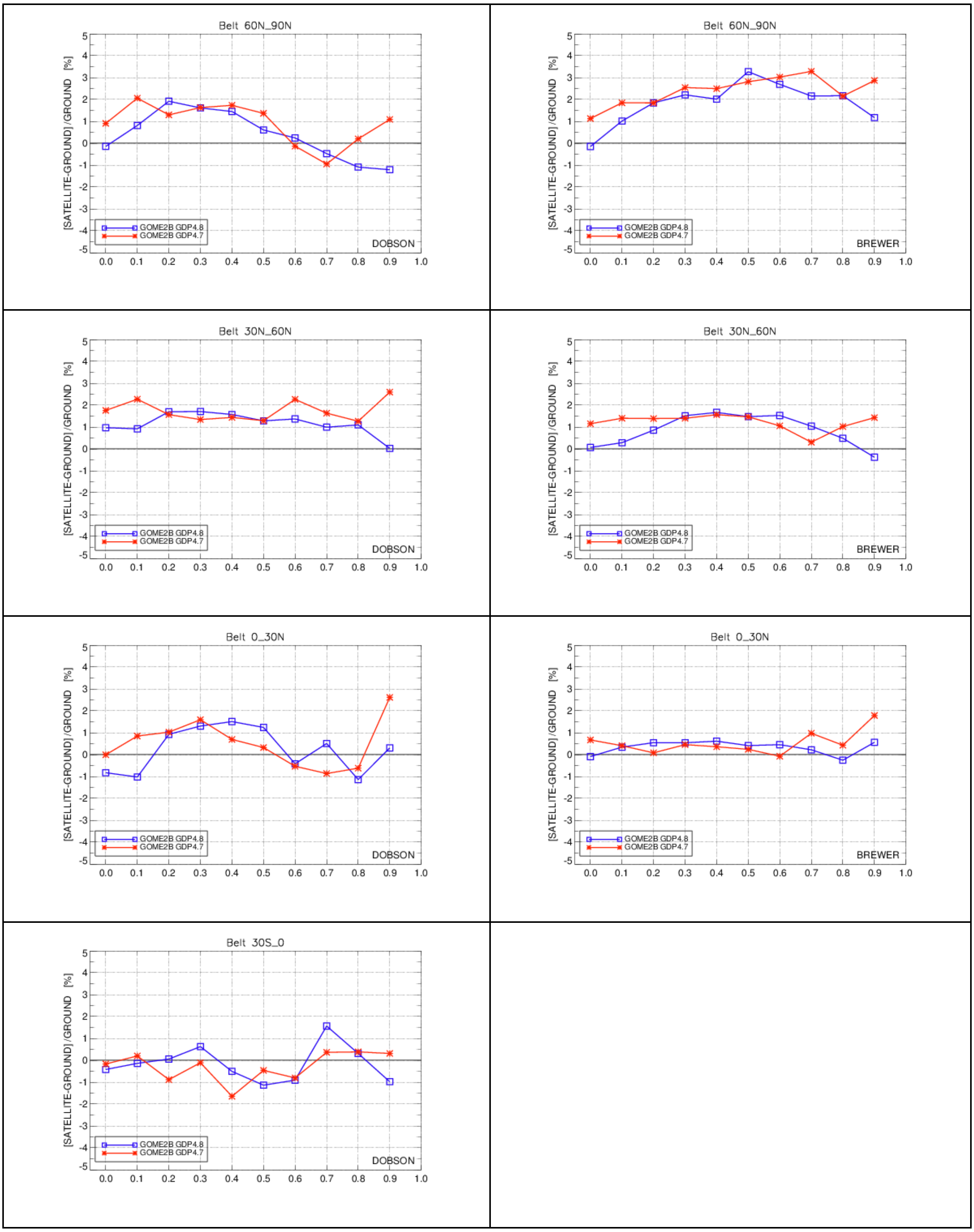

Belt 90S_60S
[SATELLITE-GROUND]/GROUND [%]
GOME2A GDP4.8
GOME2A GDP4.7
BREWER

Figure 7.1. Cloud fraction dependence of the differences per latitudinal belt between GOME-2A GDP4.8 [blue] and GOME-2A GDP4.7 [red] for Dobson (left) and Brewer (right) instruments. From top to bottom: 90° to 60°N, 60° to 30°N, 30° to 0°, 0° to -30°S, -30° to -60°S and -60° to -90°S.



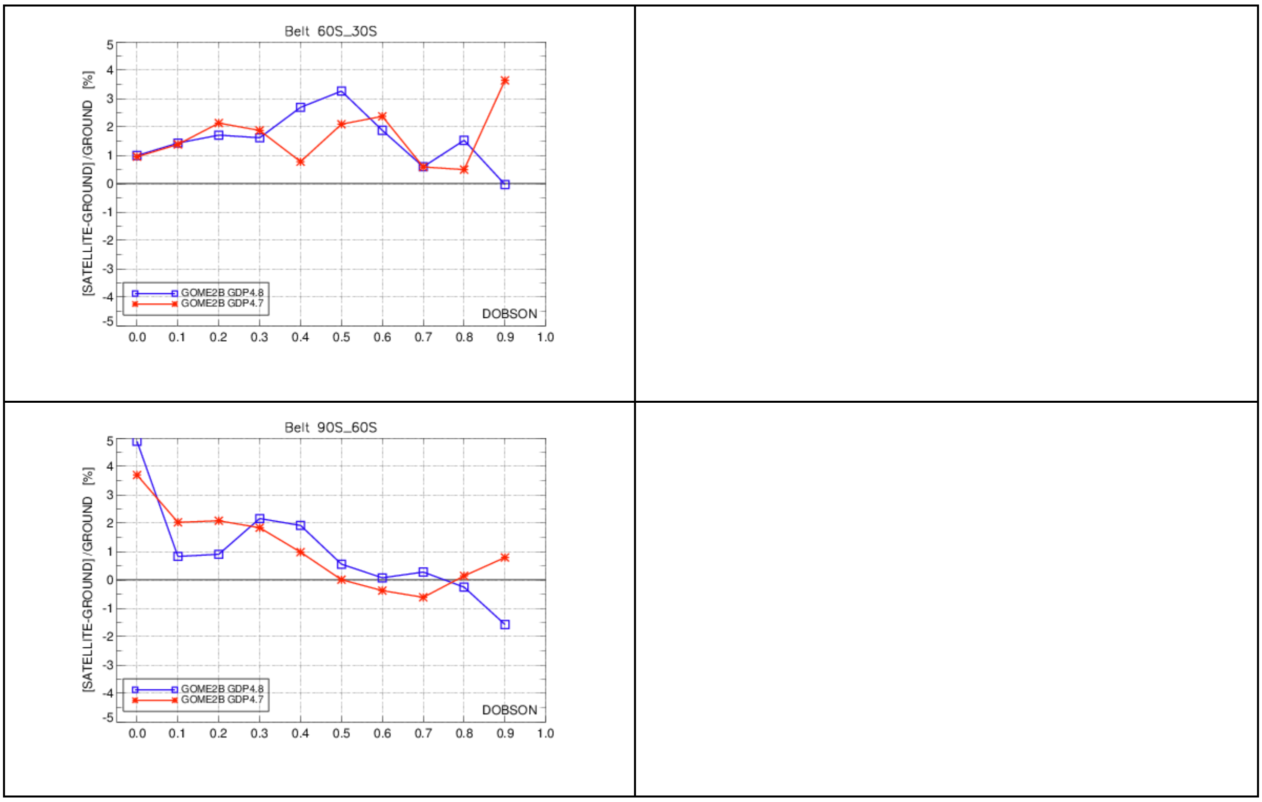
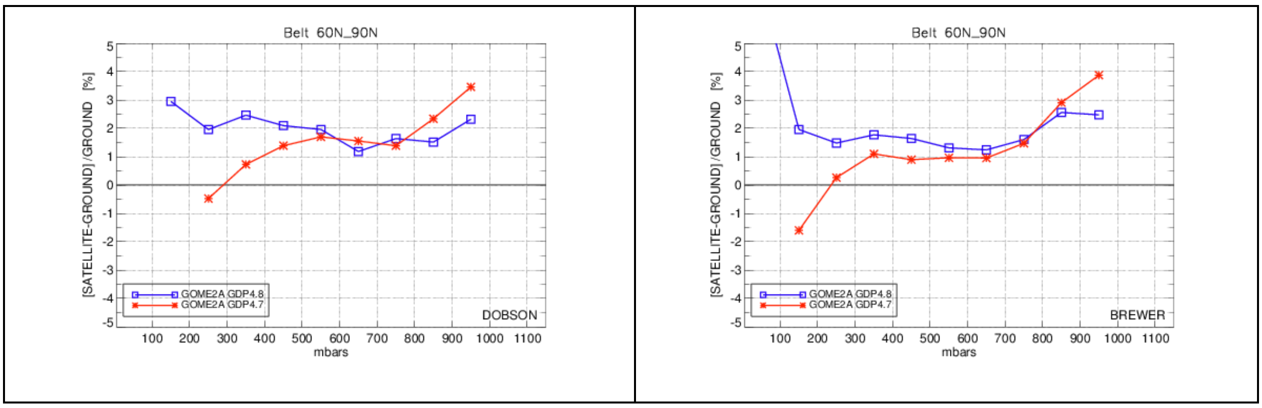
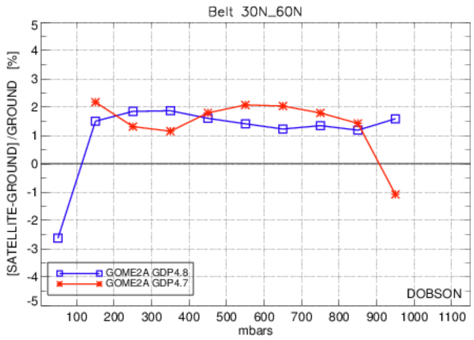
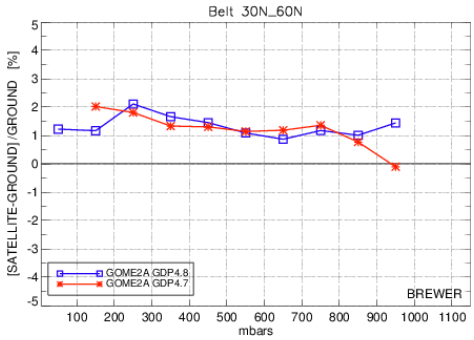


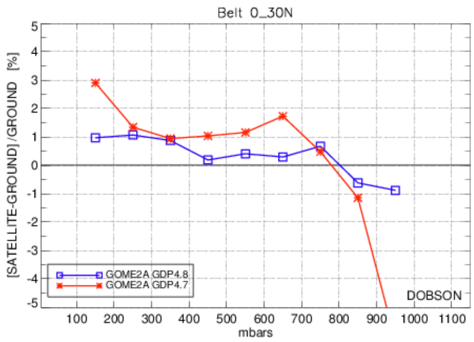
Figure 7.2. Cloud fraction dependence of the differences per latitudinal belt between GOME-2B GDP4.8 [blue] and GOME-2B GDP4.7 [red] for Dobson (left) and Brewer (right) instruments. From top to bottom: 90° to 60°N, 60° to 30°N, 30° to 0°, 0° to -30°S, -30° to -60°S and -60° to -90°S.

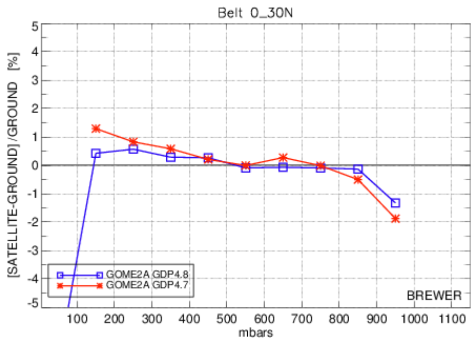
In Figure 7.1, the comparisons for GOME-2A GDP4.7 and GDP4.8 and in Figure 7.2, the comparisons for GOME-2B GDP4.7 and GDP4.8, are given. For both GOME-2A and GOME-2B, the new cloud algorithm applied to GDP4.8 data seems to reduce the differences over most of the belts for both Dobson and Brewer instruments. The cloud fraction variability is more similar between GDP4.8 and GDP4.7 for GOME-2B than for GOME-2A. This result is highly consistent with what has been observed in the analyses shown in Section 6.

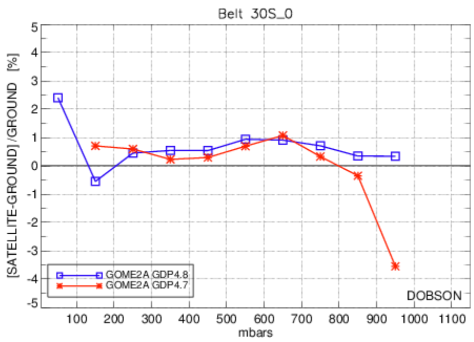


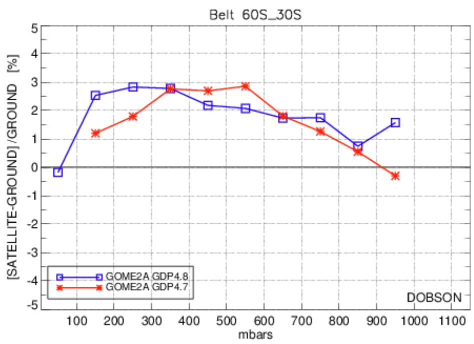

Belt 30N_60N
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 30N_60N
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
BREWER


Belt 0_30N
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 0_30N
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
BREWER


Belt 30S_0
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON


Belt 60S_30S
[SATELLITE-GROUND]/GROUND [%]
mbars
GOME2A GDP4.8
GOME2A GDP4.7
DOBSON

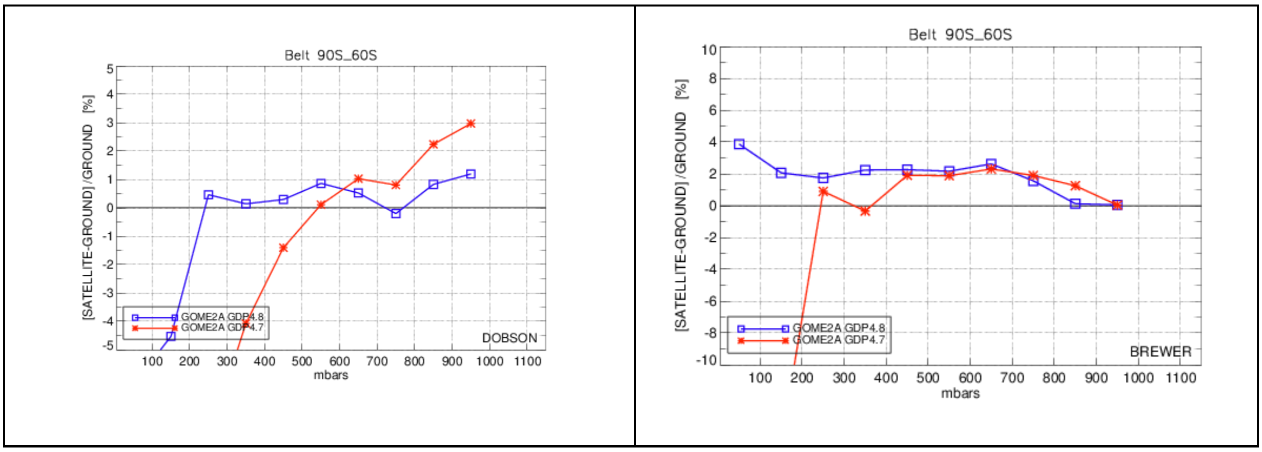
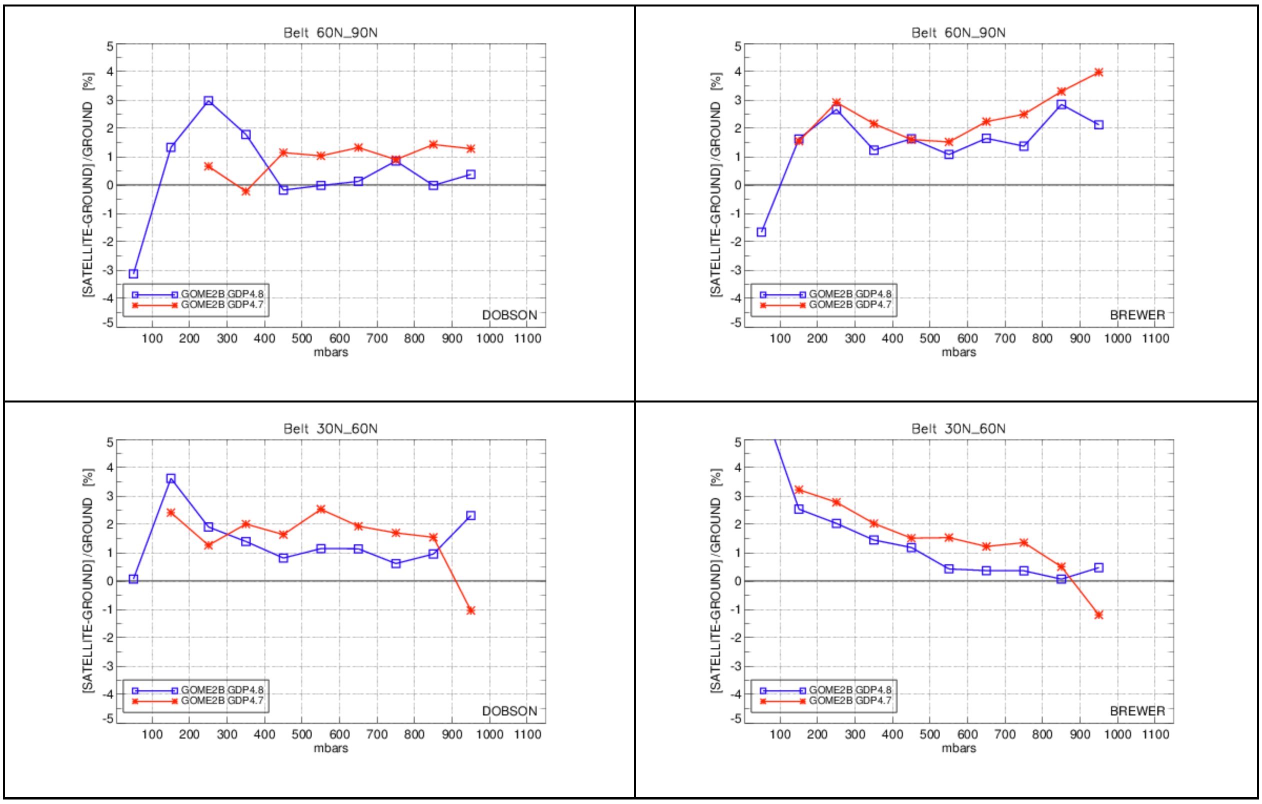


Figure 7.3. Cloud top pressure dependence of the differences per latitudinal belt between GOME-2A GDP4.8 [blue] and GOME-2A GDP4.7 [red] for Dobson (left) and Brewer (right) instruments. From top to bottom: 90° to 60°N, 60° to 30°N, 30° to 0°, 0° to -30°S, -30° to -60°S and -60° to -90°S.



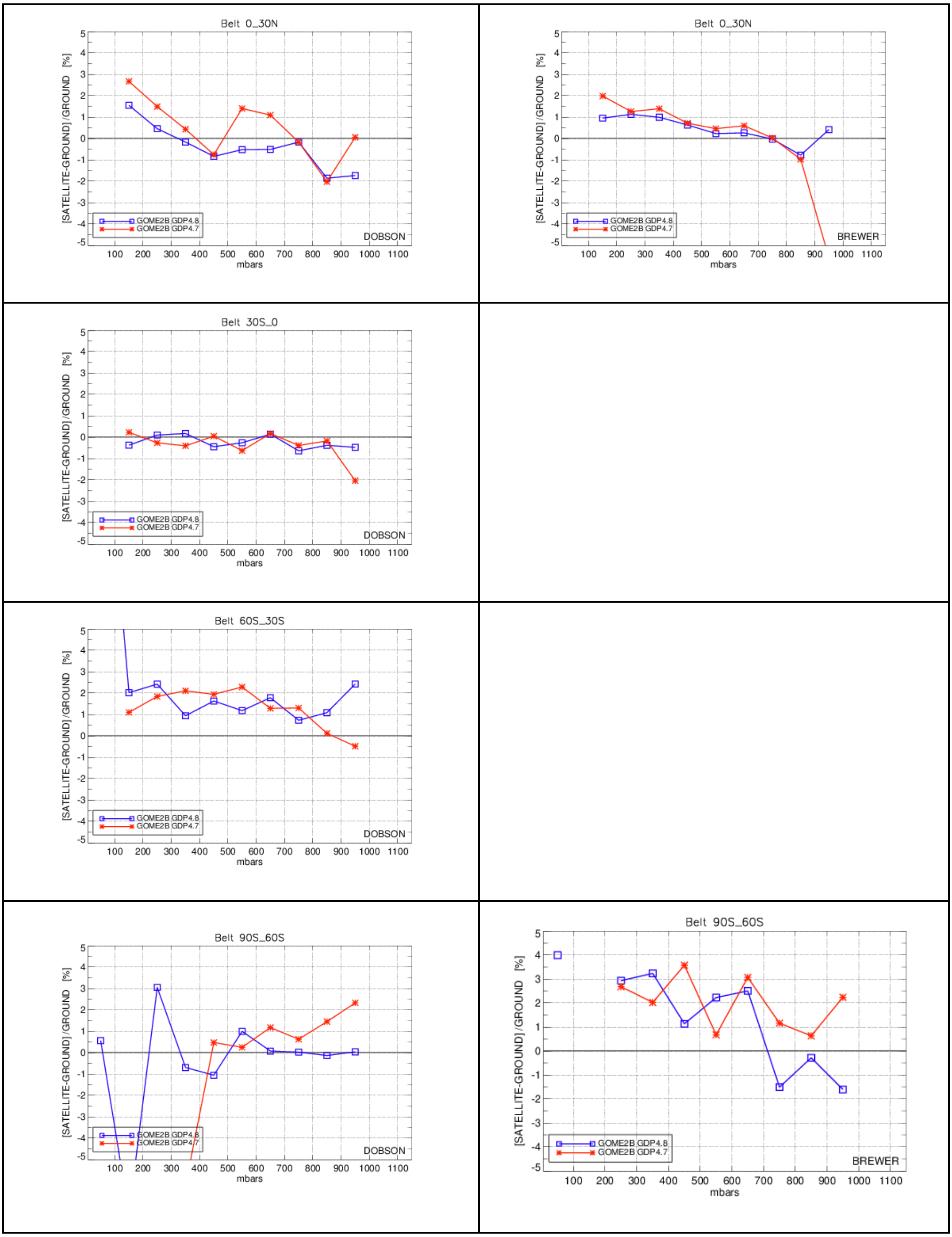


Figure 7.4. Cloud top pressure dependence of the differences per latitudinal belt between GOME-2B GDP4.8 [blue] and GOME-2B GDP4.7 [red] for Dobson (left) and Brewer (right) instruments. From top to bottom: 90° to 60°N, 60° to 30°N, 30° to 0°, 0° to -30°S, -30° to -60°S and -60° to -90°S.

A very similar picture as for the cloud fraction dependency plots shown in Figure 7.1 and Figure 7.2 is given when examining the cloud-top pressure dependency shown in Figure 7.3 and Figure 7.4. For GOME-2A [Figure 7.3] the two algorithms provide indistinguishable results for almost all belts apart from Antarctica which merits a closer examination and is to be performed in the future. For GOME-2B though [Figure 7.4] most belts appear to show a dependency with comparisons moving towards the negative values with increasing cloud top pressure, more pronounced for the Brewer coincidences.

8. SUMMARY STATISTICS

The comparison results of the GOME2/Metop-A and GOME2/Metop-B GDP4.8 TOC product are summarized in the following tables.

The mean differences presented in Tables 8.1 and 8.2 are based on period January 2007 to December 2014 for GOME-2A and January 2013 to December 2014 for GOME-2B. The main purpose of this analysis is to examine the consistency between GOME-2A and GOME-2B and the systematic differences with the previous version of the algorithm. The direct comparisons between GOME-2A and GOME-2B are consistent with the comparisons of each version of algorithm versus the ground-based data. It is indicated that GOME-2B GDP4.8 data are about 0.5% smaller than GOME-2A GDP4.8. Considering that GDP4.8 algorithm introduces a 0.3% positive offset (compared to GDP4.7) to GOME-2A for Brewers and a 0.3% negative offset (compared to GDP4.7) to GOME-2B for Dobsons. In Table 8.2, GOME-2A is used as the reference when the percentage difference between GOME-2A and GOME-2B GDP4.8 is computed. In case of comparisons between the 2 version of the algorithm the GDP4.7 is used as the reference measurements.

Table 8.1 Mean global differences between the various instruments examined for coincident *monthly* measurements only. The standard deviation represents only the latitudinal variability of the differences.

	mean diff [%]	***std [%]***	**correlation**	**Nobs**
100* [GOME-2A GDP4.8 – Dobson] / GOME-2A GDP4.8	0.89	4.27	0.964	53566
100* [GOME-2A GDP4.8 – Brewer] / GOME-2A GDP4.8	1.10	3.41	0.975	80273
100* [GOME-2B GDP4.8 – Dobson] / GOME-2B GDP4.8	0.50	4.10	0.967	8506
100* [GOME-2B GDP4.8 – Brewer] / GOME-2B GDP4.8	0.99	3.67	0.969	15134

Table 8.2 Mean global differences between the various instruments examined for coincident *daily* measurements only. The standard deviation represents only the latitudinal variability of the differences.

	mean diff [%]	***std [%]***	**correlation**	**Nobs**
For co-locations with Dobsons: 100* [GOME-2A GDP4.8 – GOME-2B GDP4.8] / GOME-2A GDP4.8	0.71	2.66	0.986	10618
For co-locations with Brewers: 100* [GOME-2A GDP4.8 – GOME-2B GDP4.8] / GOME-2A GDP4.8	0.56	2.62	0.985	19826
For co-locations with Dobsons: 100* [GOME-2A GDP4.8 – GOME-2A GDP4.7] / GOME-2A GDP4.8	0.126	1.35	0.999	86109
For co-locations with Brewers: 100* [GOME-2A GDP4.8 – GOME-2A GDP4.7] / GOME-2A GDP4.8	0.347	1.73	0.999	133804
For co-locations with Dobsons: 100* [GOME-2B GDP4.8 – GOME-2B GDP4.7] / GOME-2B GDP4.8	-0.312	1.38	0.996	15720
For co-locations with Brewers: 100* [GOME-2B GDP4.8 – GOME-2B GDP4.7] / GOME-2B GDP4.8	-0.22	1.78	0.993	27400

9. SUMMARY AND CONCLUSIONS

The main aim of the O3MSaf Total Ozone Column validation report is to assess the new operational algorithm GDP4.8 and its impact on the reported total ozone columns from GOME2/MetopA and GOME2/MetopB. In Section 2, global, inter-satellite comparisons for both sensors are presented as means to demonstrate the inter-sensor behavior of GDP4.8. Then, in Sections 3, 4 and 5, three type of comparisons were performed; the GOME-2A and GOME-2B GDP4.8 TOC data were firstly compared with archived ground-based total ozone measurements, then compared together using only concurrent measurements and thirdly they were compared correspondingly with the previous version (GDP4.7) of the algorithm, over the same ground-based observational network. Different comparative avenues were explored, as is common when validating TOC from satellite using ground-based measurements: time series, scatter plots as well as latitudinal, solar zenith angle and cloud top pressure dependencies where investigated. Of particular interest are the scan angle dependency comparisons where the GOME2B GDP4.8 TOCs show a marked improvement of the absolute differences against both Dobson and Brewer TOCs contrasted to the GD4.7 TOCs.

The algorithm differences between GDP 4.8 and GDP 4.7 are due to changes of the DOAS algorithm (improved Kurucz solar reference spectrum), the cloud treatment (for OCRA see Lutz et al., 2015 and for ROCINN see Gimeno Garcia et al., 2015) as well as the scan angle corrections. Since these changes were applied simultaneously, a posteriori it is not as straightforward to separate their relative effects on the final product. However, the effect of the change of the DOAS algorithm and the cloud treatment on the differences between GDP4.7 and 4.8 may be assessed by calculating an O_3 VCD using a simple geometric AMF. These comparisons [not shown in the report] have shown that in high latitudes, the cloud treatment plays a major role for the larger GDP 4.7 vs 4.8 differences. Furthermore, the relative smaller DOAS fitting residuals also point to the improvement of DOAS fit in GDP 4.8.

The overall effect of the new cloud algorithm in GDP4.8 may be further studied in Section 7, and in particular Figure 7.3 and Figure 7.4 where the cloud-top pressure dependency of the satellite-to-ground TOC differences are shown per belts for the whole datasets. There it can be observed that the GDP4.8 TOCs are more stable in all belts and some dependencies observed for GDP4.7 are extinct. The effect of the new algorithm is stronger for GOME2A than for GOME2B and stronger for the high altitude belts, compared to the tropics and the middle latitudes.

The main message resulting from the validation results presented above is that the new algorithm, GDP4.8, appears to produce total ozone column of equal quality as the current algorithm version GDP4.7. No strange artifacts or bizarre dependencies have been introduced by the transition to the new algorithm settings compared to the operational algorithm. As a whole, the new algorithm GDP4.8 does not seem to improve issues already identified in the previous version of the algorithm, such as the solar zenith angle dependency and the GOME-2A to GOME-2B systematic differences. A small bias ranging between 0 and 1%, constant for most latitudes, time periods and algorithm parameters has been introduced which, however, is well within the ground-based network's capability for validation. The new algorithm appears to affect GOME-2A to a larger extend than GOME-2B, possibly due to the larger instrument degradation suffered by the older sensor. As a whole, the new GDP4.8 algorithm appears to slightly improve the GOME-2B observations but slightly deteriorate the GOME-2A TOCs. The scan angle dependency issues are largely removed for the GOME-2B GDP4.8 TOCs and somewhat improved for the GOME-2A GDP4.8 TOCs compared to GDP4.7. The new cloud algorithm does introduce changes compared to the old settings but further investigation, on a global scale, using independent of the Brewer and Dobson co-locations is required in order to ascertain which algorithm produces the more stable TOCs.

Concluding, from the validation using Brewer and Dobson ground-based spectrophotometer TOC measurements, the transition for Total Ozone Columns from GOME-2/MetopA and GOME-2/MetopB to the new GDP4.8 algorithm is recommended, since according to the algorithm team, the new settings approach the "true" physics of the atmosphere in a better justifiable way than the GDP4.7 settings. It is also true that both algorithms compared to the ground-based data provide similar accuracy comparisons, however the algorithm that better approximates the real state of the atmosphere is obviously favored. It does appear, irrespective of the validation results shown in this report, complicated to provide a concrete scientific reason to explain why GDP 4.8 does not provide better TOC results compared to GDP 4.7 apart from a possible 'cancellation of errors' effect in GDP 4.7. This effect is not unknown in long-term algorithm development and its validation.

REFERENCES

I. O3MSAF TECHNICAL REPORTS

- [**ATBD**] Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, NO_2, BrO, SO_2, H_2O, HCHO and Cloud Properties (GDP 4.8 for O3M-SAF OTO and NTO), DLR/GOME-2/ATBD/01, Rev. 3/A, Valks, P., et al., March 2015.
- [**PUM**] Product User Manual for GOME Total Column Products of Ozone, NO_2, BrO, SO_2, H_2O, HCHO and Cloud Properties, DLR/GOME/PUM/01, Rev. 3/A, Valks, P., et. al., 2015.
- [**PRD**] Product Requirements Document, SAF/O3M/FMI/RQ/PRD/001/Rev. 1.7, D. Hovila, S. Hassinen, D. Loyola, P. Valks, J., S. Kiemle, O. Tuinder, H. Joench-Soerensen, F. Karcher, 2015.

II. SCIENTIFIC REPORTS

- Staehelin J., J. Kerr, R. Evans and K. Vanicek, Comparison of total ozone measurements of Dobson and Brewer spectrophotometers and recommended transfer functions, WMO TD N. 1147, No 149, 2003.
- World Meteorological Organization, Scientific assessment of ozone depletion: 1998, Global Ozone Res. Monit. Proj., Rep. 44, Geneva, Switzerland, 1998.
- World Meteorological Organization, Scientific assessment of ozone depletion: 2002, Global Ozone Res. Monit. Proj., Rep. 47, Geneva, Switzerland, 2002.
- World Meteorological Organization, Scientific assessment of ozone depletion: 2006, Global Ozone Res. Monit. Proj., Rep. 50, Geneva, Switzerland, 2006.

III. PEER-REVIEWED ARTICLES

- **Antón**, M., D. Loyola, M. López, J. M. Vilaplana, M. Bañón, W. Zimmer, and A. Serrano, Comparison of GOME-2/MetOp total ozone data with Brewer spectroradiometer data over the Iberian Peninsula, Ann. Geophys., 27, 1377–1386, doi:10.5194/angeo-27-1377-2009, 2009.
- **Balis** D, Koukouli ME, Brinksma EJ, Kroon M, Veefkind JP, Labow G, McPeters RD, Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, Journal of Geophysical Research, 112 (D24): Art. No. D24S46, 2007a.
- **Balis** D., J-C. Lambert , Van Roozendael M., D. Loyola, R. Spurr, Y. Livschitz, P. Valks, T. Ruppert, P. Gerard, J Granville and V. Amiridis, Reprocessing the 10-year GOME/ERS-2 total ozone record for trend analysis: the new GOME Data Processor Version 4.0, Validation, J. Geophys. Res., vol. 112, D07307, doi:10.1029/2005JD006376, 2007b.

- **Bernhard**, G., R. D. Evans, G. J. Labow, and S. J. Oltmans, Bias in Dobson total ozone measurements at high latitudes due to approximations in calculations of ozone absorption coefficients and air mass, J. Geophys. Res., 110, D10305, doi:10.1029/2004JD005559, 2005.
- **Fioletov**, V., J. Kerr, E. Hare, G. Labow, and R. McPeters, An assessment of the world ground-based total ozone network performance from the comparison with satellite data, J. Geophys. Res., 104, 1737–1747, 1999.
- **Hao**, N., Koukouli, M. E., Inness, A., Valks, P., Loyola, D. G., Zimmer, W., Balis, D. S., Zyrichidou, I., Van Roozendael, M., Lerot, C., and Spurr, R. J. D.: GOME-2 total ozone columns from MetOp-A/MetOp-B and assimilation in the MACC system, Atmos. Meas. Tech., 7, 2937-2951, doi:10.5194/amt-7-2937-2014, 2014.
- **Hassinen**, S., Balis, D., Bauer, H., et al., Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability, Atmos. Meas. Tech. Discuss., 8, 6993-7056, doi:10.5194/amtd-8-6993-2015, 2015.
- **Kerr**, J. B., I. A. Asbridge, and W. F. J. Evans, Intercomparison of total ozone measured by the Brewer and Dobson spectrophotometers at Toronto, J. Geophys. Res., 93(D9), 11,129–11,140, 1988.
- **Koukouli**, M., D. Balis, D. Loyola, P. Valks, W. Zimmer, N. Hao, J.-C. Lambert, M. Van Roozendael, C. Lerot, and R. Spurr, Geophysical validation and long-term consistency between GOME-2/ MetOp-A total ozone column and measurements from the sensors GOME/ERS-2, SCIAMACHY/ENVISAT and OMI/Aura, Atmos. Meas. Tech., 5, 2169–2181, 2012.
- **Labow**, G. J., R. D. McPeters, P. K. Bhartia, and N. Kramarova, A comparison of 40 years of SBUV measurements of column ozone with data from the Dobson/Brewer network, J. Geophys. Res. Atmos., 118, 7370–7378, doi:10.1002/jgrd.50503, 2013.
- **Lambert**, J.-C., M. Van Roozendael, M. De Mazière, P. C. Simon, J.-P. Pommereau, F. Goutail, A. Sarkissian, and J. F. Gleason, Investigation of pole-to-pole performances of spaceborne atmospheric chemistry sensors with the NDSC, J. Atmos. Sci., 56, 176–193, 1999.
- **Loyola**, D., et al., The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation, J. Geophys. Res., 116, D07302, doi:10.1029/2010JD014675, 2011.
- **Lutz**, R., D. Loyola, S. Gimeno Garcia and F. Romahn, OCRA radiometric cloud fractions for GOME-2A/B, in preparation for Atmos. Meas. Tech., 2015.
- **Van Roozendael**, M., et al., Validation of ground-based UV-visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, J. Atmos. Chem., 29, 55–83, doi:10.1023/ A:1005815902581, 1998.
- **Weber**, M., L. N. Lamsal, M. Coldewey-Egbers, K. Bramstedt, and J. P. Burrows, Pole-to-pole validation of GOME WFDOAS total ozone with ground-based data, Atmos. Chem. Phys., 5, 1341– 1355, 2005.