AC SAF VALIDATION REPORT

Validated products:

Identifier	Name	Acronym
O3M-112	Reprocessed high-resolution ozone profiles (MetOp-A/B)	MxG-RP1-O3HRPR
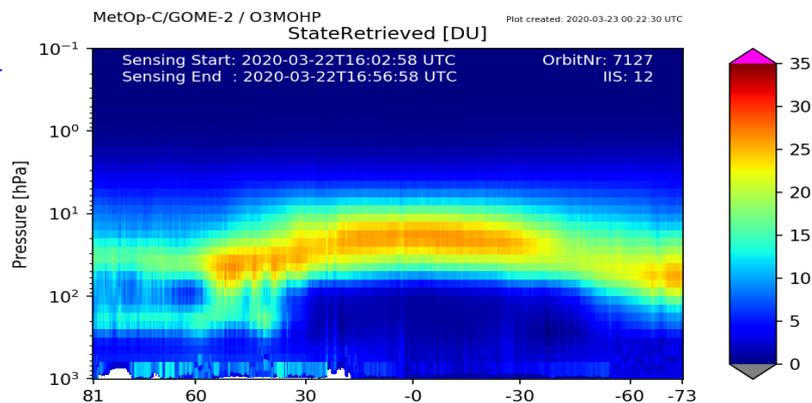
Authors:

Name	Institute
Andy Delcloo	RMI
Katerina Garane	AUTH
Peggy Achtert	DWD

Reporting period: **January 2007 – December 2018**

Validation methods: Lidars and microwave radiometers (altitude range 15 – 60 km)
Balloon soundings (altitude range 0 – 34 km)
Dobson and Brewer observations (total column DU)

Input data versions: Base Algorithm Version:5.3/6.0/6.1/6.2/6.3

Data processor versions: ProductAlgorithmVersion: 2.00/2.00/2.00/2.02/2.02/2.02
Product Software Version: 2.01/2.01/2.01/2.03/2.03/2.03

Table of Contents

Introduction to EUMETSAT Satellite Application Facility on Atmospheric Composition monitoring (AC SAF)

Background

The monitoring of atmospheric chemistry is essential due to several human-caused changes in the atmosphere, like global warming, loss of stratospheric ozone, increasing UV radiation, and pollution. Furthermore, the monitoring is used to react to threats caused by natural hazards as well as to follow up the effects of international protocols.

Therefore, monitoring the chemical composition of the atmosphere and its effect on the Earth's radiative balance is a very important duty for EUMETSAT. The target is to provide information for policy makers, scientists and the general public.

Objectives

The main objectives of the AC SAF is to process, archive, validate and disseminate atmospheric composition products (O_3, NO_2, SO_2, BrO, HCHO, H_2O, OClO, CO, NH3), aerosol products and surface ultraviolet radiation products. The majority of the AC SAF products are based on data from the GOME-2 and IASI instruments onboard *EUMETSAT's* MetOp satellites.

Another important task besides the near real-time (NRT) and offline data dissemination is the provision of long-term, high-quality atmospheric composition products resulting from reprocessing activities.

Product categories, timeliness and dissemination

NRT products are available in less than three hours after measurement. These products are disseminated via EUMETCast, WMO GTS or the internet.

- Near real-time trace gas column (total and tropospheric O_3 and NO_2, total SO_2, total HCHO, CO) and high-resolution ozone profile

- Near real-time absorbing aerosol index (AAI) from main science channels and polarization measurement detectors

- Near real-time UV index, clear-sky and cloud-corrected

Offline products are available within two weeks after measurement and disseminated via dedicated web services at EUMETSAT and AC SAF.

- Offline trace gas column (total and tropospheric O_3 and NO_2, total SO_2, total BrO, total HCHO, total H_2O) and high-resolution ozone profile

- Offline absorbing aerosol index from main science channels and polarization measurement detectors

- Offline surface UV, daily doses and daily maximum values with several weighting functions

Data records are available after reprocessing activities from the EUMETSAT Data Centre and/or the AC SAF archives.

- Data records generated in reprocessing

- Lambertian-equivalent reflectivity

- Total OClO

Users can access the AC SAF offline products and data records free of charge by registering at the AC SAF web site.

More information about the AC SAF project, products and services: https://acsaf.org/

AC SAF Helpdesk: helpdesk@acsaf.org

Twitter: https://twitter.com/Atmospheric_SAF

Applicable AC SAF Documents

[ATBD] Algorithm Theoretical Basis Document for Near Real Time and Offline Ozone profiles, KNMI/GOME/ATBD/01/001, issue 2.0.1, Olaf Tuinder, 20181115.

[PUM] Product User Manual for Near Real Time and Offline Ozone profiles, KNMI/GOME/PUM/001, issue 2.00, Olaf Tuinder, 20181115.

Both documents are available at http://acsaf.fmi.fi in the *Documents* section.

Acronyms and abbreviations

ATBD	Algorithm Theoretical Basis Document
AUTH	Aristotle University of Thessaloniki
B-M	Brewer Mast
DOAS	Differential Optical Absorption Spectroscopy
DWD	Deutscher Wetterdienst
ECC	Electrochemical concentration cell
GAW	Global Atmosphere Watch
GDP	GOME Data Processor
GOME	Global Ozone Monitoring Experiment
LAP/AUTH	Laboratory of Atmospheric Physics/Aristotle University of Thessaloniki
MetOp	Meteorological Operational satellite
MWR	Microwave Radiometers
NDACC	Network for the Detection of Atmospheric Composition Change
NH	Northern Hemisphere
O3-CCI	Ozone – Climate Change Initiative
OMI	Ozone Monitoring Instrument
OPERA	Ozone Profile Retrieval Algorithm
RMI	Royal Meteorological Institute of Belgium
SH	Southern Hemisphere
SZA	Solar Zenith Angle
TOC	Total Ozone Column
TOMS	Total Ozone Mapping Spectrometer
WMO	World Meteorological Organization
WOUDC	World Ozone and UV Data Center

1. General Introduction

This report contains validation results of the reprocessed datasets from GOME-2/MetOp-A and GOME-2/MetOp-B ozone profile products, retrieved by the Ozone Profile Retrieval Algorithm (OPERA) at KNMI. It covers the time period from January 2007 to December 2018. Ozone profiles retrieved from processed level-1b data were retrieved with 80 km x 40 km resolution.

Since this work was carried out in three different institutes, this document is split up into three separate parts. The first part contains the validation of the retrieved GOME-2 ozone profiles using ozonesondes (chapter 2). This part validates the retrieved ozone profiles in the troposphere and the lower stratosphere. The second part (chapter 3) uses measurements with lidars and microwave radiometers to assess the performance of GOME-2 ozone profiles; primarily in the stratosphere from 20 to 60 km altitude (chapter 3). The third part of this report (chapter 4), covers the validation of the integrated ozone profile product through an intercomparison with ground truth data from spectrophotometers (Dobson and Brewer). Additionally, the consistency of the integrated ozone profiles of GOME-2A and GOME-2B is examined by intercomparison to the respective operational products from GOME-2A and GOME-2B (chapter 4). The outcome of the different validation parts is summarized in the summary and conclusions section at the end of this report.

Table 1.1 presents the different accuracies which are taken into account to assess the quality of the product.

Table 1.1: Different intended accuracies for ozone profiles, provided in the Product Requirements Document SAF/AC/FMI/RQ/PRD/001

Accuracy		
Threshold	**Target**	**Optimal**
30 % in stratosphere	15 % in stratosphere	10 % in stratosphere
70 % in troposphere	30 % in troposphere	25 % in troposphere

2. Validation of ozone profiles using ozonesondes

2.1 *Introduction*

This chapter presents validation results for the AC SAF GOME-2 ozone profile product. The validation was carried out using ozone sounding profiles.

Ozonesondes are lightweight balloon-borne instruments which measure ozone concentrations from the surface up to about 30 km with much better vertical resolution than possible from satellite data. In general, measurement precision and accuracy are also better compared to satellite observations, at least in the lower stratosphere and the troposphere. Another advantage is that ozone soundings can be performed at any time and during any meteorological condition.

The precision of ozonesondes varies with altitude and depends on the type of ozonesonde used. Tabel 2.1 shows indicative precision of the Electrochemical Concentration Cell (ECC) and Brewer-Mast (B-M) and the Japanese KC79 ozonesondes (KC79) at different pressure levels of the sounding.

Tabel 2.1: Precision (in percent) of different types of ozonesondes at different pressure levels.

Pressure level (hPa)	ECC	B-M	KC79
10	2	10	4
40	2	4	3
100	4	6	10
400	6	16	6
900	7	14	12

Profiles from ozonesondes are most reliable around the 40 hPa level, which is around the ozone maximum. The error bar of profiles from ozonesondes increases rapidly at levels above the 10 hPa level, which is at around 31 km altitude. For this validation report, only the station of Hohenpeissenberg is using B-M sondes (Brewer and Milford, 1960). The other stations under consideration (*Table A. 3*) use ECC sondes (Komhyr, 1969, 1971). The Japanese stations Tateno-Tsukuba and Sapporo used also Iodine sensors until 2011. KC-79 sondes (Kobayashi and Toyama, 1966) are not launched anymore, they are replaced by ECC sondes.

2.2 *Dataset description*

GOME-2 ozone data used in this validation report covers the time period from January 2007 to December 2018. GOME-2 ozone profile data was made available by KNMI at pre-selected site where ozone soundings are performed on a regular basis. Ozonesonde data was made available by the World Ozone and Ultraviolet Data Center (WOUDC). (http://www.woudc.org) and the NILU's Atmospheric Database for Interactive Retrieval (NADIR) at Norsk Institutt for Luftforskning (NILU) (http://www.nilu.no/nadir/). In Table A.3, an overview is shown from the ozonesonde station data used in this report.
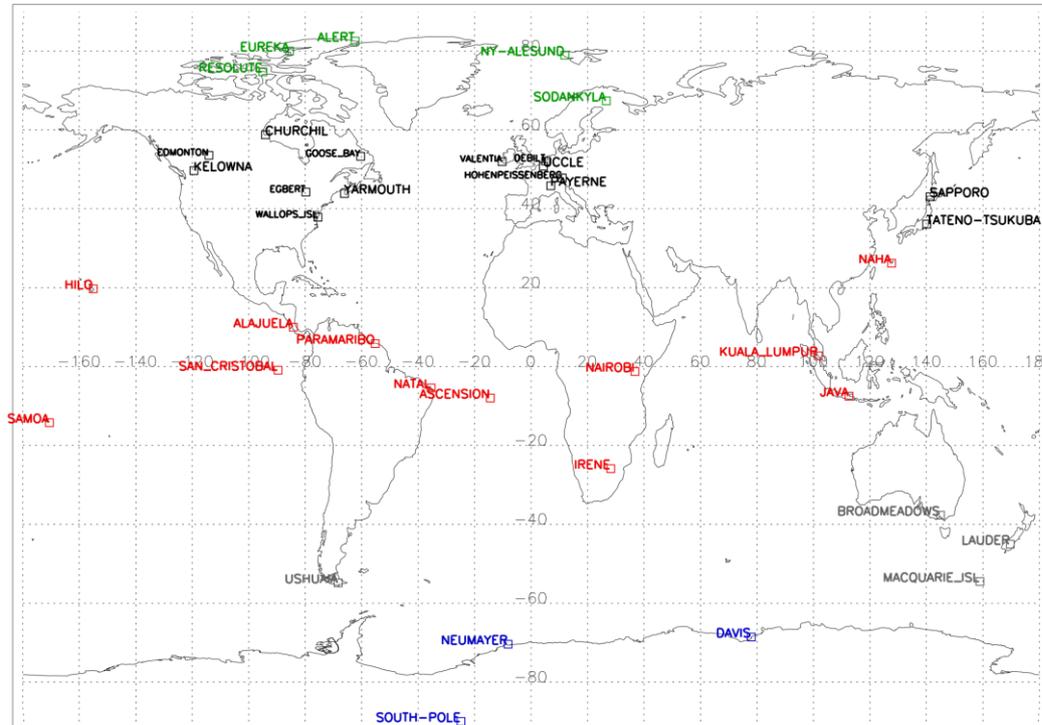
Figure 2.1: Stations consulted for validation. Latitude belts from north to south: polar stations north: green (67N – 90 N), midlatitude stations north: black (30 N – 67 N), Tropical stations: red (30 N – 30 S), midlatitude stations south: grey (30 S – 70 S), polar stations south: blue (70 S – 90 S).

Ozonesonde data are generally made available by the organization carrying out observations after a short delay related to data quality assurance. Nevertheless, some organizations make their ozone profile data readily available for validation purposes. Since the time period under consideration here is between January 2007 and December 2018, more ozonesondes than we usually can consult during an operational review are available.

Table A.3 of the Appendix shows an overview of the station data used in this validation report using ozonesondes and the collocations in space and time are shown in Figure 2.2 for GOME-2A and in Figure 2.3 for GOME-2B.

AC SAF: Validation of reprocessed ozone profiles   9
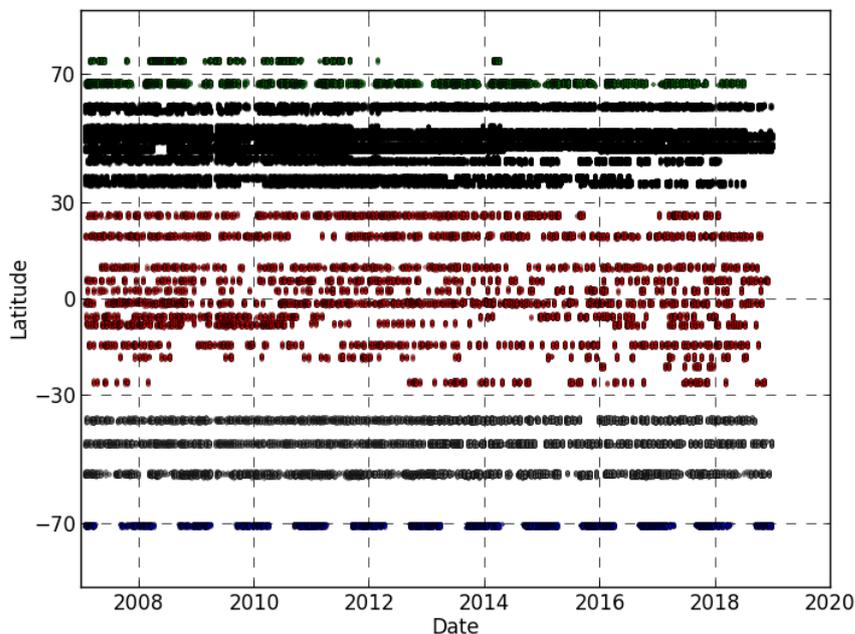
Figure 2.2: Spatial and temporal representation of the collocation data used for the validation with ozonesonde data for the time period January 2007 - December 2018 for GOME-2A.
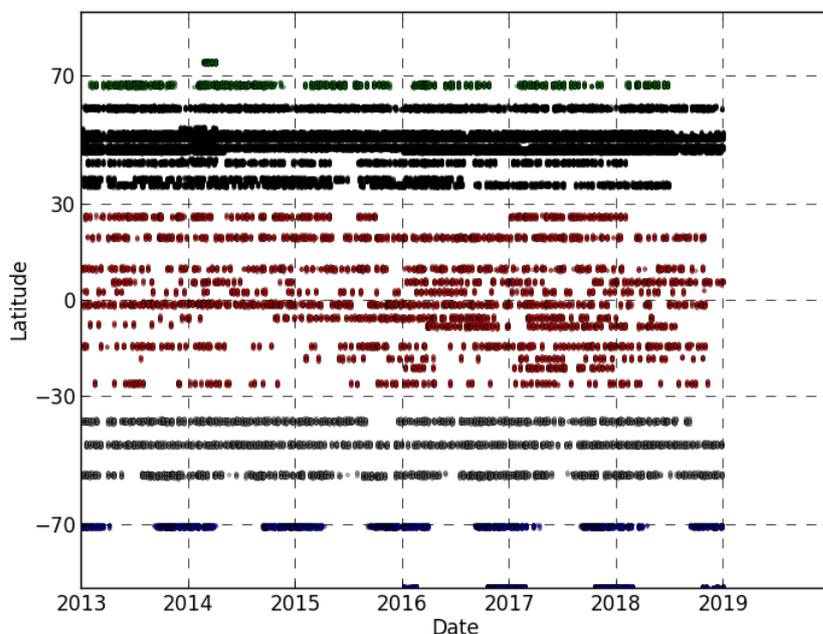
*Figure 2.3: Spatial and temporal representation of the collocation data used for the validation with ozonesonde data for the time period January 2013 - December 2018 for GOME-2B.*

2.3 *Comparison procedure*

2.3.1 **Co-location criteria**

The selection criteria are twofold:

- The geographic distance between the GOME-2 pixel center and the sounding station location is less than 100 km.

- The time difference between the pixel sensing time and the sounding launch time is less than ten hours.

Each sounding that is correlated with a GOME-2 overpass is generally correlated with several GOME-2 pixels if the orbit falls within this 100 km circle around the sounding station. This means that a single ozone profile is compared to more than one GOME-2 measurement.

2.4 *Ozone sounding pre-processing*

GOME-2 ozone profiles are given as partial ozone columns on 40 varying pressure levels calculated by the Ozone Profile Retrieval Algorithm (OPERA) developed by KNMI. Ozone partial columns are expressed in Dobson Units.

Ozonesondes measure ozone concentration along the ascent with a typical vertical resolution of 100 m while GOME-2 profiles consist 40 layers between the ground and 0.001 hPa. Ozonesondes give ozone concentration in partial pressure. The integration requires interpolation, as GOME-2 levels never match exactly ozonesonde layers. This interpolation causes negligible errors given the high vertical resolution of ozonesonde profiles.

For comparison, ozonesonde profiles are integrated between the GOME-2 pressure levels. When a single ozonesonde profile is compared to different GOME-2 profiles, the actual reference ozone values are not the same given that the GOME-2 level boundaries vary from one measurement to another. Integrated ozonesondes data will be referred to in this report as X_{sonde}.

GOME-2 layers are relatively thick and GOME-2 layer boundaries show small variations compared to the layer thickness. Hence, individual layers generally occur around the same altitude. The altitude of those layers can be considered as "fixed" and therefore the center of an "*averaged layer altitude (or pressure)*" is used in plotting the data.

In this report, the validation of the GOME-2 profiles is calculated by using the averaging kernels (AVK) of the GOME-2 profile. The motivation to apply the AVK is to "smooth" the ozone soundings towards the resolution of the satellite:

$$X_{avk\_sonde} = X_{apriori} + A (X_{raw\ sonde} - X_{apriori}) \qquad (1)$$

___

Where A represents the averaging kernel, X_{avk_sonde} is the retrieved ozone sonde profile, X_{sonde} is the ozone sonde profile and $X_{apriori}$ is the a priori profile.

2.5 *Results*

2.5.1 Difference profiles

The relative difference between the ozone profiles from GOME-2 and an ozonesonde is calculated as:

$$(X_{GOME-2} - X_{sonde})/X_{sonde.} \tag{2}$$

For comparing the GOME-2 ozone profile with the smoothed ozonesonde profiles (AVK ozonesondes) the following equation is used:

$$(X_{GOME-2} - X_{AVK-SONDE})/X_{AVK-SONDE} \tag{3}$$

Figure 2.4 shows relative difference profiles between GOME-2 ozone profiles at the one hand and on the other hand ozonesonde-, and AVK ozonesonde profiles for different latitude belts for GOME-2A (time period: 200701 – 201812) and GOME-2B (time period: 201301 – 201812), for the specific time periods under consideration.

In the next sections, we will discuss the seasonal behaviour and other possible influences on the quality of the ozone profile product.

For the polar and midlatitude stations, the difference plots in Figure 2.4 show that GOME-2 ozone profiles are within the optimal error range of 10%, compared to the ozonesonde reference, except for the Upper Troposphere – Low Stratosphere (UTLS) region. For the troposphere, most of the latitude belts show relative differences within 25%. Applying the averaging kernels, improves the comparison significantly. For the tropical stations however, there is a significant overestimation of tropospheric ozone.

Since the tropospheric integrated ozone column (TrOC) is an official operational product, its results are not mentioned in this report and will be part of a different review. There is still some room to improve this product, since the tropospheric part is strongly scan angle dependent (Delcloo, 2020). Here we will focus on the quality of the ozone profiles in the lower and the upper stratosphere as it is reported in the two-yearly operational reports. These documents are available at http://acsaf.fmi.org in the *Documents* section (operational reports).

Table 2.1 provides an overview of the height ranges related to the troposphere, the UTLS-zone and the stratosphere.

Relative difference plot for N_polar

Relative difference plot for N_midlat

Relative difference plot for Tropical

Relative difference plot for S_midlat
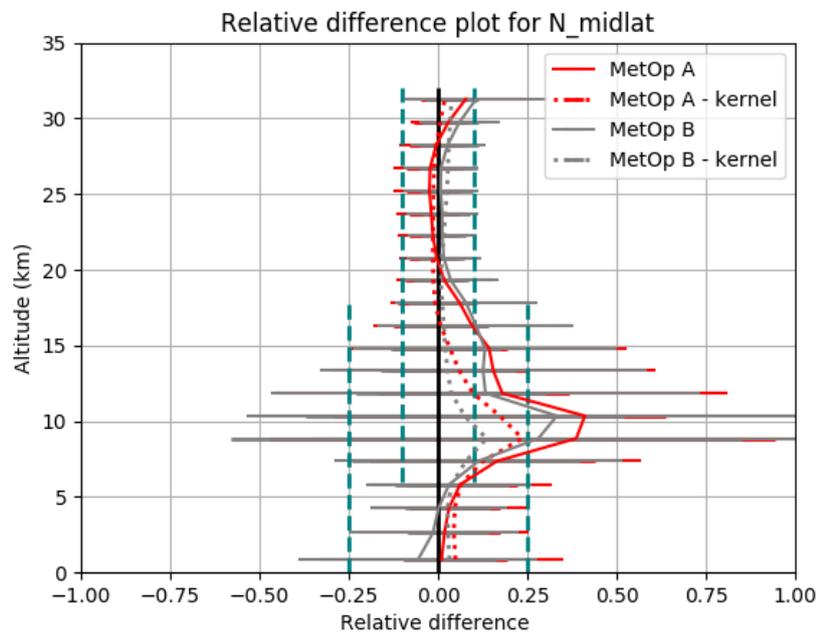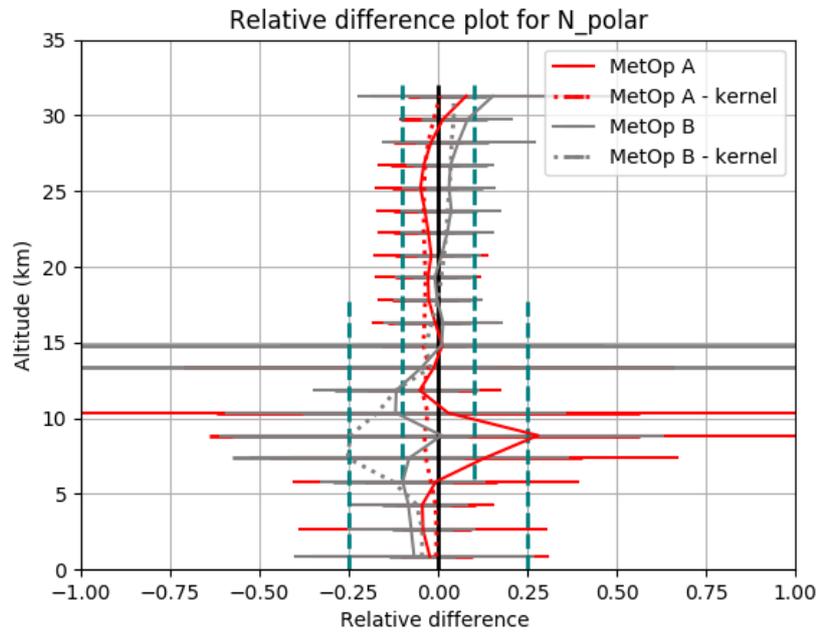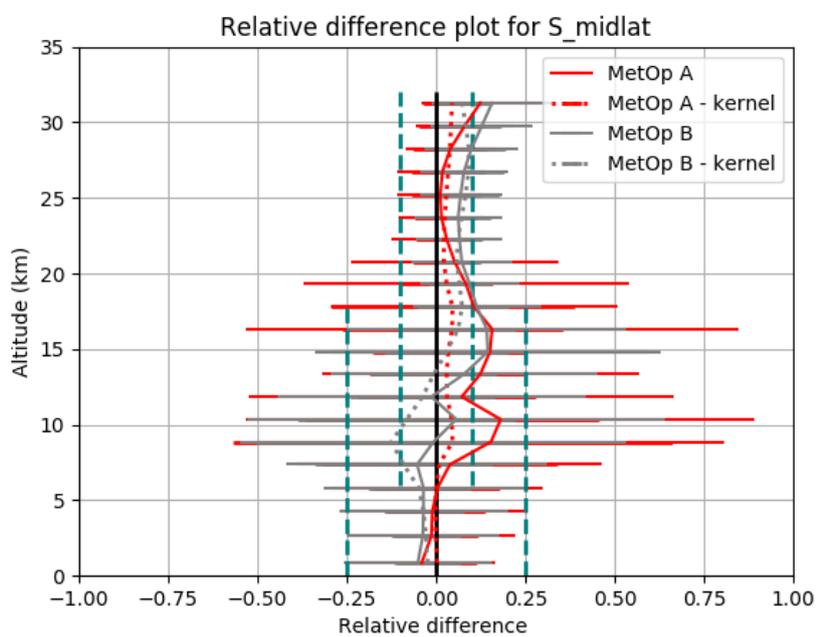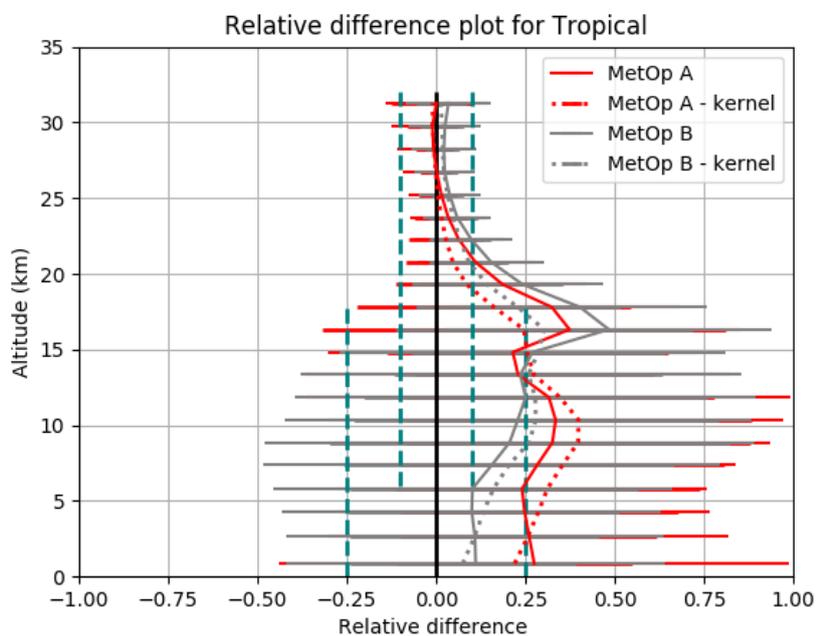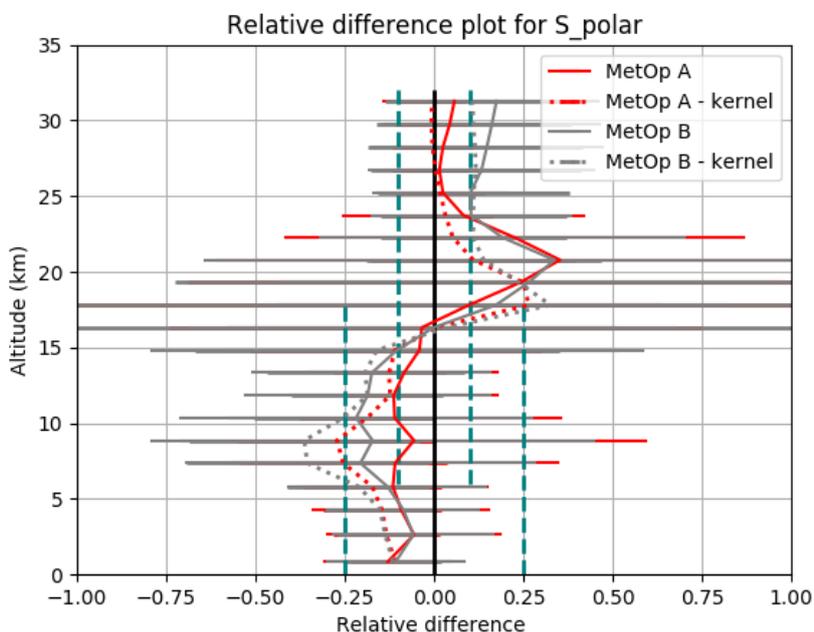
Figure 2.4: Relative difference in ozone profiles from GOME-2, ozonesondes and smoothed ozonesondes according to equations (2) and (3) for different latitude belts and for different sensors (GOME-2A/2B) for the time period January 2007 to December 2018 (GOME-2A) and for the time period January 2013 to December 2018 (GOME-2B). The error bars represent one standard deviation on the mean error. The green dashed lines are the optimal values for stratospheric and tropospheric ozone according to Table 1.1.

Table 2.1: Definition of the ranges in km for troposphere, UTLS-zone and stratosphere for the different latitude belts.

	Troposphere	UTLS	Lower Stratosphere
Polar Regions	< 6 km	6 km - 12 km	12 km - 30 km
Mid-Latitudes	< 8 km	8 km - 14 km	14 km - 30 km
Tropical Regions	< 12 km	12 km - 18 km	18 km - 30 km

Table 2.2: Relative Differences (RD) and standard deviation (STDEV) of GOME-2 ozone profiles product with respect to XAVK-sonde for the lower stratosphere, taking into account five latitude belts for the time periods January 2007 - December 2018 (GOME-2A), January 2013 - December 2018 (GOME-2A) and GOME-2B (January 2013 – December 2018).

	Lower stratosphere MetOp-A 2007 - 2018			Lower stratosphere MetOp-A 2013 - 2018			Lower stratosphere MetOp-B 2013 - 2018		
	AD (DU)	RD (%)	STDEV (%)	AD (DU)	RD (%)	STDEV (%)	AD (DU)	RD (%)	STDEV (%)
northern polar region	-9.4	-3.4	14.3	2.56	0.7	14.6	0.1	1.3	27.2
northern midlatitudes	-2.2	-0.4	9.0	1.91	2.2	7.7	3.3	2.0	8.8
tropical regions	1.8	2.1	9.2	0.63	5.1	8.6	7.0	5.5	8.3
southern midlatitudes	4.8	3.3	14.8	1.30	6.2	9.4	13.1	7.0	10.5
southern polar region	-1.5	4.1	38.7	0.92	4.3	33.0	6.8	8.8	59.4

*The relative difference statistics are derived as a weighted average over the lower- and upper stratospheric ozone profile levels. The absolute differences however are integrated over respectively the lower- and upper stratospheric ozone profile levels.

Table 2.2 shows an overview of the obtained results for both sensors for the lower stratosphere. For the ozone profile product, the optimal values are met in the lower stratosphere for both GOME-2A and GOME-2B for the full time series. However, when we take a close look to the full time period of GOME-2A, there is an abrupt change in 2013. Values, obtained after 2013 show slightly higher values in the lower stratosphere, when compared with the full time series of GOME-2A. The main reason for this offset between both sensors can be attributed to a change in the GOME-2A swath width, which changed from 1920 to 960 km on the 15th of July 2013. Figure 4.6 in chapter 4.2.1.1 illustrates this abrupt change very clearly, while it is more difficult to identify in the ozone profile time series in in the lower stratosphere.

Table 2.2 also shows the statistics for GOME-2A for the same time period as GOME-2B (2013 – 2018) to make a consistent intercomparison between both ozone profile products for both sensors. These statistics confirm that both products are showing very similar results.

2.6 *Scatter plots for the retrieved ozone partial columns*

Scatter plots for different altitude levels are plotted in Figure 2.5, showing the retrieved ozone partial columns as a function of the reference partial column measured by ozonesondes. This is shown in Figure 2.5 for the northern midlatitude stations at six different altitude levels. In order to evaluate these ozone profile layers as seen by the satellite, we will smooth the ozone profile layers by applying the averaging kernels. This is shown in Figure 2.6. We observe that the slope values indeed improve significantly (closer to 1) while the intercept values are closer to 0.

The interpretation of "better results" should be taken with care. Applying the kernels using equation 1 is a way to smooth the ozone profile towards a comparable vertical resolution of the retrieved ozone profile. High resolution effects like filaments present for example in secondary ozone maxima are mostly not seen by GOME-2 which results in sometimes large differences between observed and retrieved partial ozone columns. The regression line in the scatter plots show therefore that GOME-2 loses sensitivity in the lower troposphere and around the UTLS-zone (Figure 2.5). We can conclude that upon smoothing matching, the agreement improves.
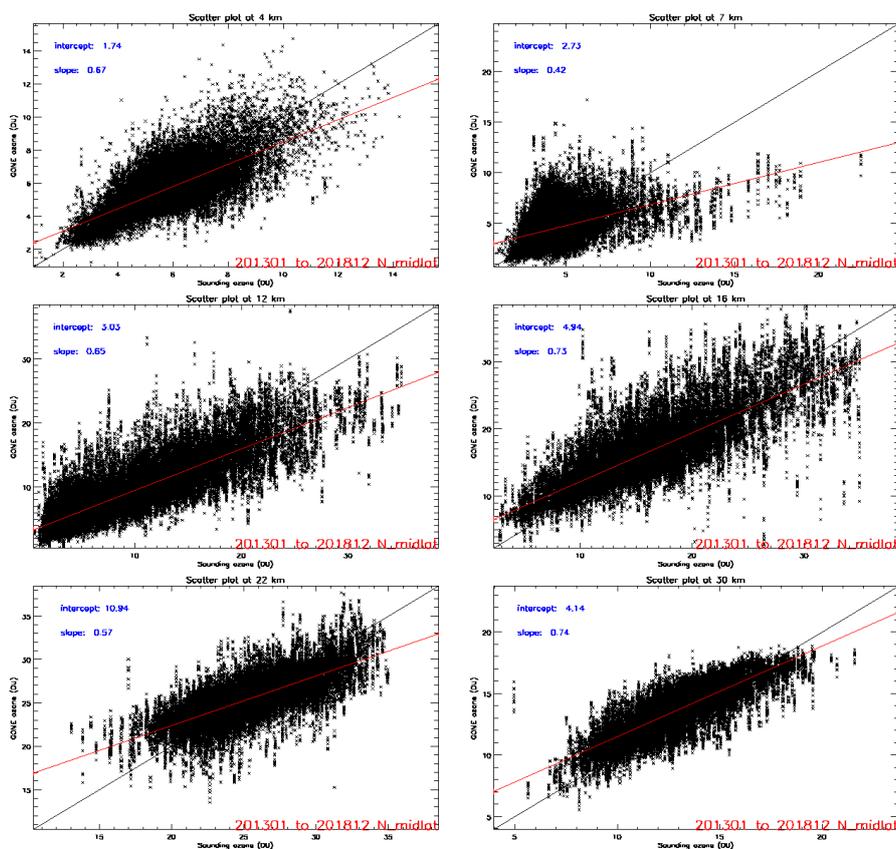

Figure 2.5: Scatter plot at 6 different altitude levels for the stations at northern midlatitudes (January 2013- December 2018, GOME-2B).
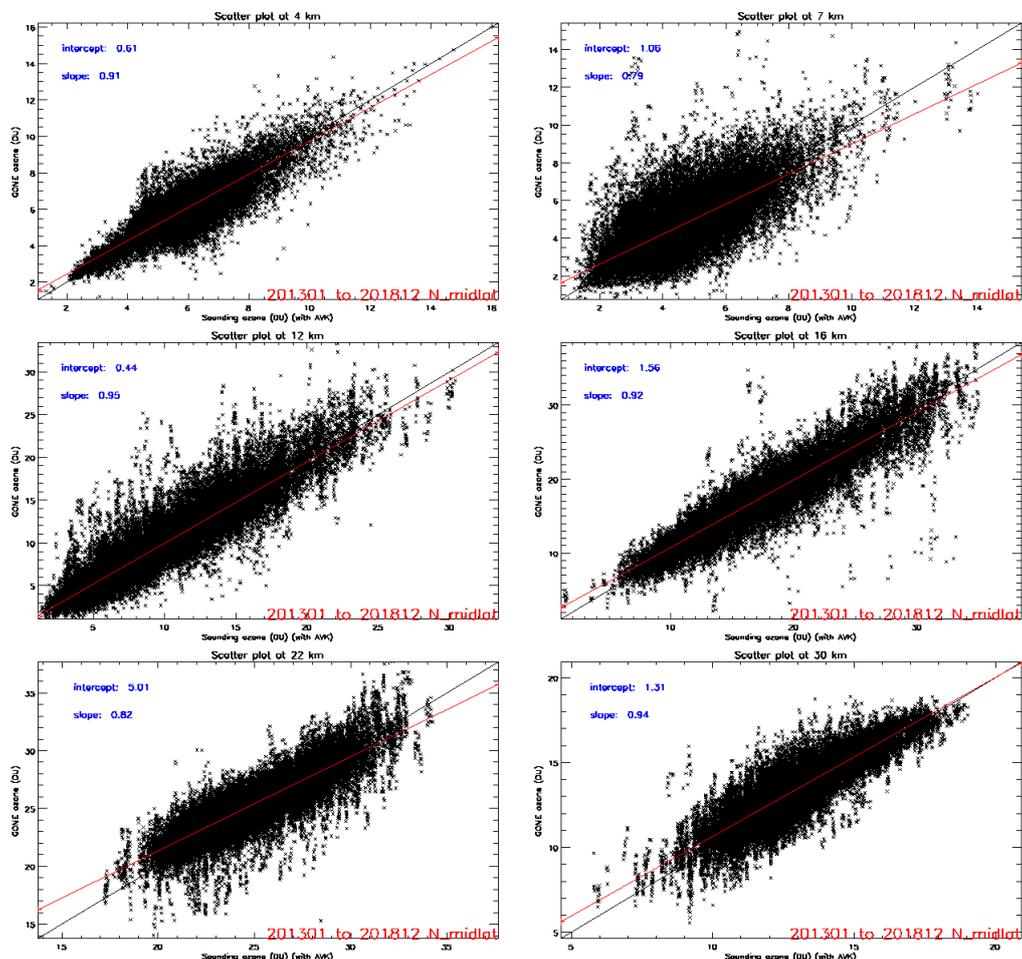
Figure 2.6: Scatter plot at 6 different altitude levels for the stations at northern midlatitudes (January 2013- December 2018, GOME-2B), applying the kernels

Previous studies with GOME-2/MetOp-A data (Delcloo and Kreger, 2013) have shown that the GOME-2 ozone profile retrieval shows a seasonal dependency and is also influenced by the Solar Zenith Angle (SZA), more specifically at higher latitudes (polar stations, not shown here). Besides this influence on SZA, the dependence on cloud cover and seasonal behaviour has been verified. We could not identify any specific dependence on cloud cover. For the seasonal behaviour, it is known from previous reports (Delcloo and Kreher, 2013) that there is some seasonal behaviour present at higher altitudes. This is also true for the lower altitudes as can be seen in Figure 2.7 and Figure 2.8 for the northern midlatitude stations.

When we compare the results from the operational ozone profiles, to be consulted in the operations reports (AC SAF Operations Report, 2020), it is shown that the seasonal behaviour has significantly improved in this reprocessed product, resulting in a lower standard deviation (Table 2.2Table 2.1). More results can be consulted on the official validation website for ozone profiles where the statistics of the operational products are published twice a year.
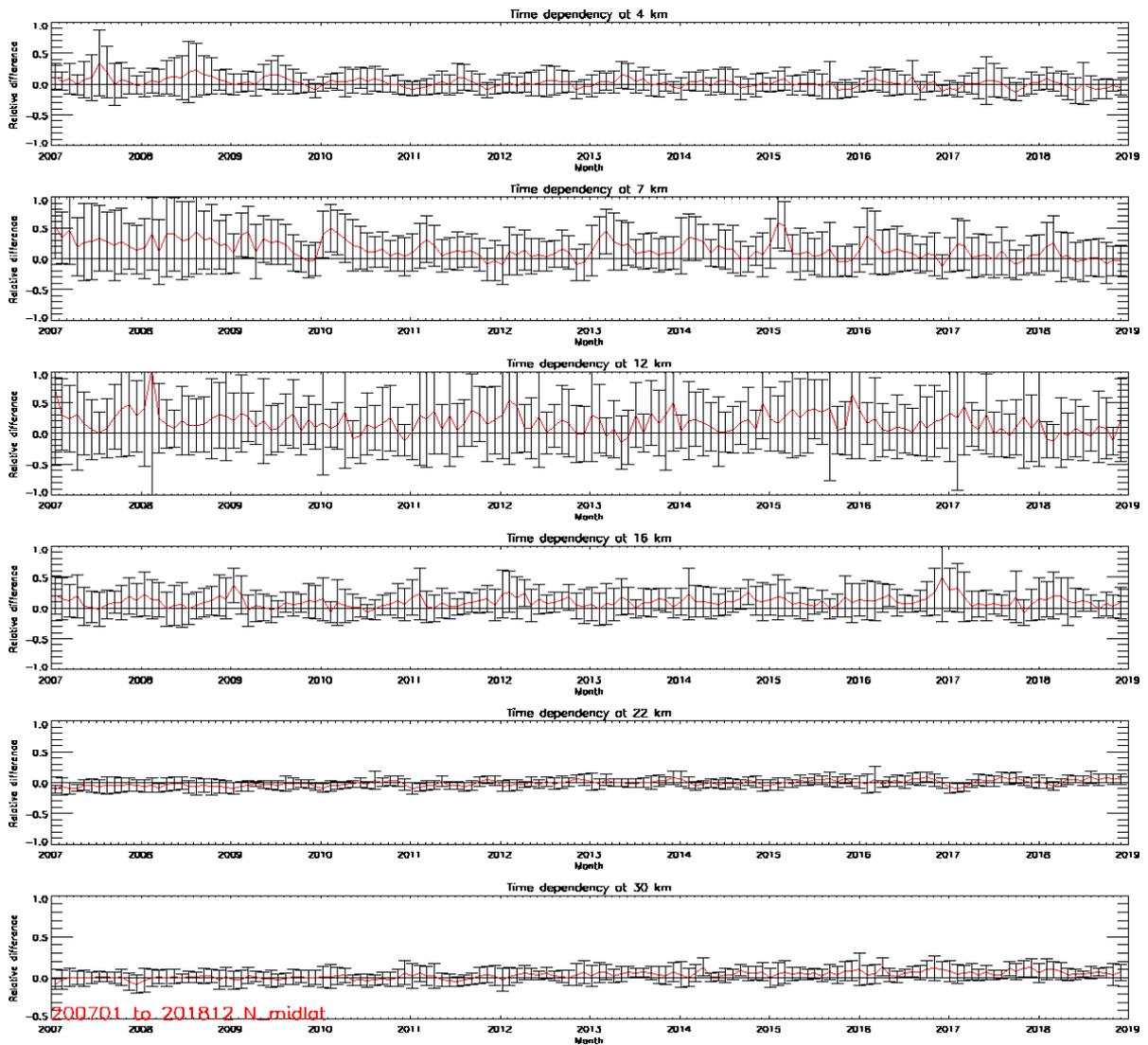
Figure 2.7: Time series at 6 different altitude levels for the stations at northern midlatitudes (January 2007 - December 2018) for the reprocessed GOME-2A time series.
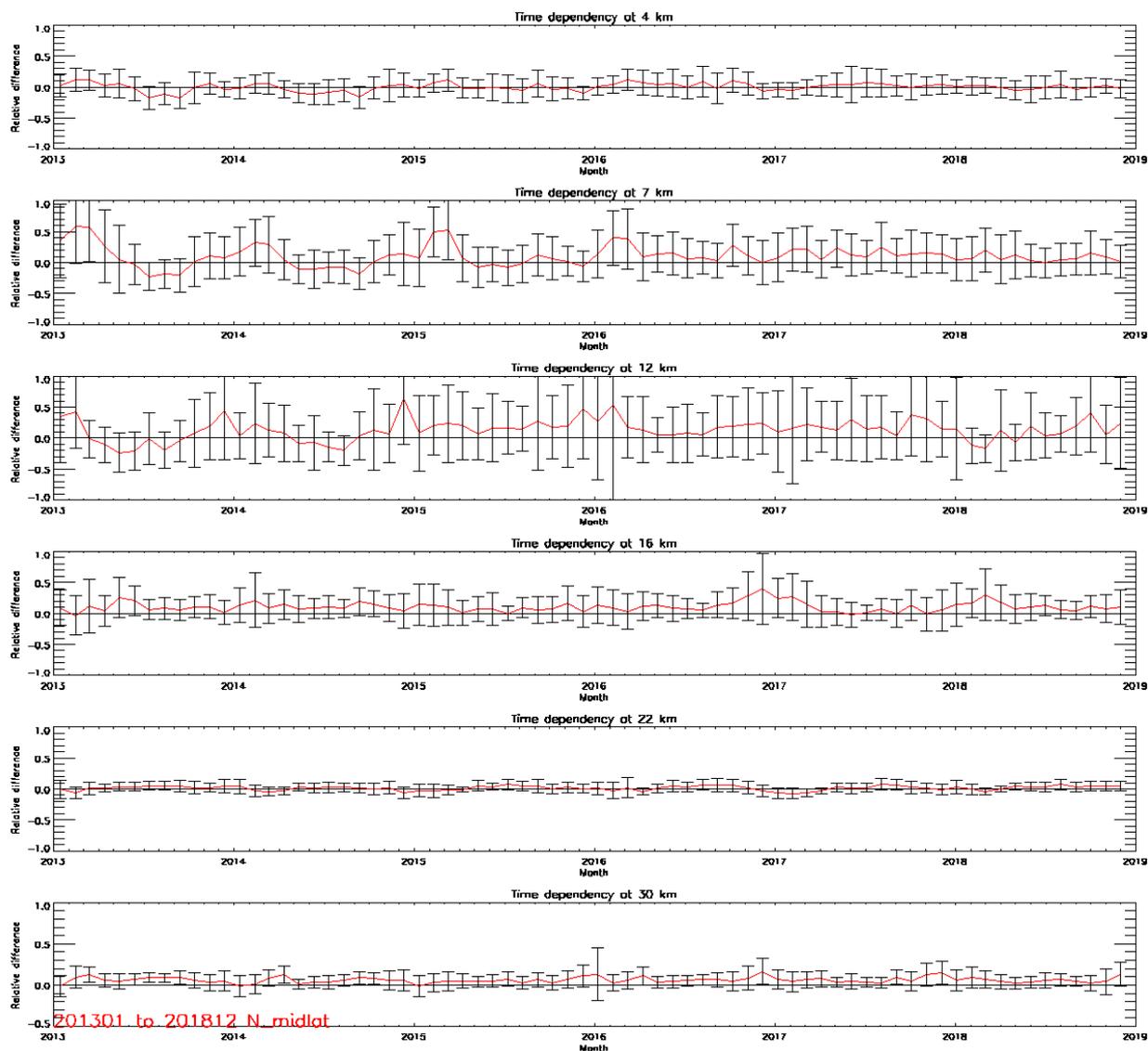
Figure 2.8: Time series at 6 different altitude levels for the stations at northern midlatitudes (January 2013 - December 2018) for the reprocessed GOME-2B time series.

2.7 *Median sensitivity*

According to Keppens et al., 2015, it is interesting to have a more detailed look to the averaging kernels, which can be seen as the vertical sensitivity of the ozone profile product. The best way to do it in an intuitive way is to observe the evolution of the median sensitivity in function of time and vertical profile. Here we only look at the median sensitivity until an altitude of about 30 km for northern midlatitude station (Figure 2.9). We also observe here that there is a seasonal variation present for the whole profile. When we compare the operational product (upper panels) against the reprocessed values (lower panels) for GOME-2A (left panels) and GOME-2B (right panels), it is shown how the over-sensitivity has been significantly reduced in the reprocessed product. Also the seasonal behaviour, present in the product is visible here. For the lower stratosphere, we observe a reduction in amplitude around the ozone maximum (22 – 23 km) and for the reprocessed version, the over-sensitivity, present in the operational versions is significantly reduced. These figures also show that the retrieval schemes for both sensors are now quiet comparable/homogenized when compared against the median sensitivity plots from the operational versions.
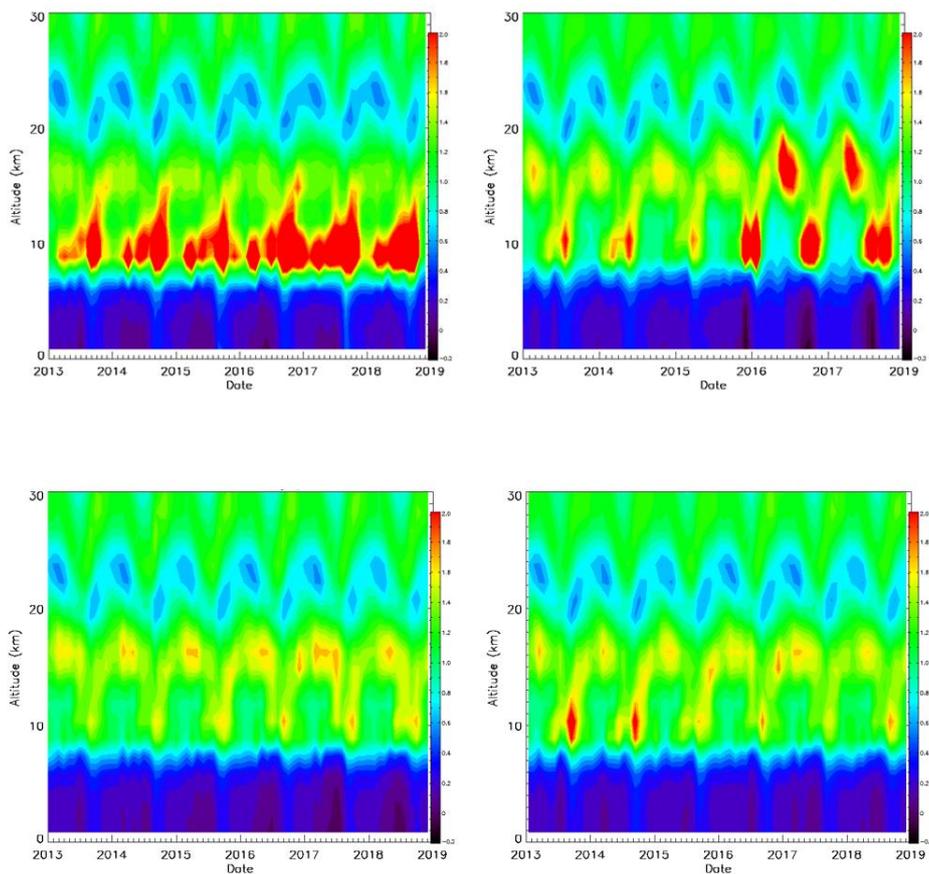
Figure 2.9: Median sensitivity plot (0 -30 km) from collocated data for GOME-2A (left) and GOME-2B (right) for the northern midlatitude stations according to Keppens et al., 2015. Above are the median sensitivity plots, derived from the operational products, below are the median sensitivity plots derived from the reprocessed products.

2.8 *General conclusions for the validation of ozone profiles, using ozonesondes*

The GOME-2A and GOME-2B reprocessed vertical ozone profile products were validated against ozonesonde data. The validation results have revealed the following properties:

- The comparisons of both sensors show comparable results and are all within optimal value for the lower stratosphere.
- GOME-2 ozone profile retrievals show a seasonal dependency, especially for the altitude range 20 – 25 km (region where the ozone maximum is located).
- Besides the influence on SZA, the dependency on cloud cover has been verified. For cloud cover, we could not identify any specific dependency.
- The median sensitivity plots show a significant improvement in the retrieval algorithm, resulting in a stable and comparable 'fingerprint' between both sensors.

It is shown that the optimal value (10% accuracy) is met in the lower stratosphere (Table 2.2) for all belts under consideration.

3. Validation of ozone profiles with lidar and microwave instruments

Lidars and microwave radiometers (MWR) are the main ground-based instruments available for validation purposes in the upper stratosphere. Their altitude range covers typically 15 km to 50 or 60 km (Table 3.1: Typical precision and height resolution of lidars and MWR (Steinbrecht et al., 2006)Table 3.1). This significantly extends the range covered by ozonesondes towards higher altitudes. It also provides a good overlap from 15 to 30 km altitude. Note that there are only about 10 operational lidar and MWR stations on the globe that provide regular data, though not as rapidly and operationally as the ozonesonde stations. Typically, ozone profiles do not become available until several weeks after the measurement.

The Differential Absorption Lidar (DIAL) technique provides accurate vertical profiles of ozone in the altitude range from 15 to 50 km, depending on the individual lidar system (Godin et al., 1989). Clouds and daylight conditions inhibit good measurements (Leblanc and McDermid, 2000; Steinbrecht et al., 2006), so lidar ozone profiles are restricted to cloud free nights. Typically, 5 to 8 lidar measurements per month are taken at a station. Depending on atmospheric conditions and lidar system efficiency, each ozone profile measurement covers several hours. For the lidars, number density versus geometric altitude is the natural coordinate system of the measurement.

MWR measures the thermal radiation of a pressure broadened emission line. Line-shape depends on the pressure/altitude profile of ozone (Lobsiger et al., 1984; Parrish et al., 1988). Measurement of the precise line-shape, thus, allows for retrieving the ozone profile. Similar to many satellite measurements, an optimal estimation retrieval (Rodgers, 1990) provides ozone profiles in various coordinate systems, including number density versus altitude for the NDACC MWR profiles. MWR ozone profiles typically cover 20 to 60 km altitude. In contrast to lidars, MWR has little weather dependence, and measures during daylight as well. On average, MWR profiles are measured on 20 days per month. The integration time of one MWR profile varies from 30 minutes to 5 hours, depending on the individual instrument (Boyd et al., 2007; Hocke et al., 2007).

Table 3.1: Typical precision and height resolution of lidars and MWR (Steinbrecht et al., 2006)

	Lidar		microwave radiometer	
Height [km]	**Precision [%]**	**height resolution [km]**	**precision [%]**	**height resolution [km]**
15	5	1.4		
20	5	1.2	3	10
25	3	1.0	3	10
30	3	1.8	3	10
35	3	4.2	3	14
40	5	7.2	3	14
45	15	8.6	3	20
50	55	8.6	3	20
50-70			3	20

3.1 *Dataset description*

The ground-based validation profiles come from the NDACC (Network for the Detection of Atmospheric Composition Change, http://www.ndsc.ncep.noaa.gov/). NDACC lidar and microwave instruments go through an evaluation process and thorough quality checks (Keckhut et al., 2004). The ozone profiles are not available in near real time. A minimum of one month is necessary before profiles become available but most stations need three or more months. NDACC demands that ozone profiles are submitted at least once per year to their database.
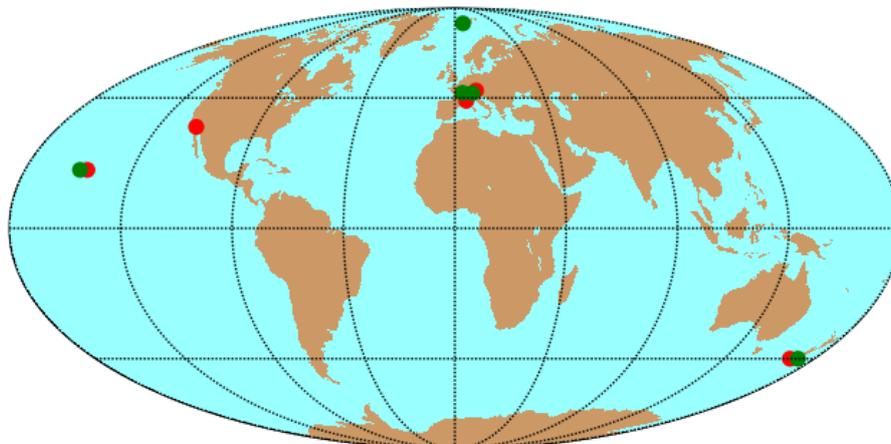
Figure 3.1: Stations consulted for validation. Lidar station in red and mircrowave radiometer station in green.

The stations (Fig. 3.1) used in this validation for the lidar/microwave data are from the Network for the Detection of Atmospheric Composition Change (NDACC): Ny-Ålesund (microwave, 78.92° N, 11.93° E), Payerne (microwave, 46.82° N, 6.95° E), Hohenpeissenberg (lidar, 47.8° N, 11.0° E), Bern (microwave, 46.95° N, 7.45° E), Haute-Provence (lidar, 43.94° N, 5.71° E), Table Mountain (lidar, 34.4° N, 117.7° W), Mauna Loa (lidar and microwave, 19.54° N, 155.58° W), and Lauder (lidar and microwave, 45.04° S, 169.68° E). Polar stations north are located between 65N and 90 N, the mid-latitude stations north are between 25° N and 65° N, and the tropical stations are located between 25° N and 25° S.

3.2 *Comparison procedure*

Generally, the comparison procedure is the same as for the ozonesondes, outlined in Section 2 (see also Delcloo and Kins, 2009; 2012). Different temporal resolution and measurement frequency of the ground-based instruments, however, require some minor changes.

3.3 *Co-location criteria in time and space*

Only ground-based and satellite profiles that are close in space and in time to a GOME-2 profile are compared. Nightly mean lidar measurements are compared to GOME-2 profiles measured either the morning after or the morning before the lidar profile. This means that a maximum time difference of 20 hours is allowed.

MWR measure around the clock, typically one profile every hour. So usually, MWR profiles can be compared with GOME-2 ozone profiles measured within less than 2 hours. Usually all GOME-2 measurements are made in the local morning.

Only GOME-2 profiles with ground pixels centers closer than 200 km to the validation stations are considered. A 200 km radius typically gives about 50 co-located GOME-2 high-

resolution profiles per station and per day. Larger co-location radii result in larger geophysical differences, smaller radii result in too few comparisons cases.

3.4 *Pre-processing of the ground-based ozone profiles.*

Like the ozonesonde data, lidar and MWR ozone number density profiles are first averaged over the GOME-2 retrieval layers, usually 40 layers, about 2 km wide. The resulting slightly smoothed profiles are called X_{ref}.

Since the GOME-2 measurement alone does not fully constrain the retrieved ozone profile, GOME-2 profiles are a mix of measured information and a-priori "climatological" ozone profiles. At altitudes where the measurement provides tight constraints, the retrieved ozone comes to 80% or 90% of the measurement. At other altitudes (usually the troposphere and mesosphere), the GOME-2 profile comes to 80% or 90% from an a-priori profile. For the validation of the retrieval process, it makes sense to also consider reference profiles X_{ref} that have been smoothed by the averaging kernels, and have the same mix of measured and a-priori profile as the GOME-2 profiles. Eq. 1 (see Section 2.3) describes the underlying mathematics. These resulting profiles are called $X_{AVK\ apriori}$ in the following.

An intermediate smoothed profile X_{AVK} is obtained by applying the GOME-2 averaging kernels, without adding the a-priori profile information. This is achieved by scaling the averaging kernels to unit vertical sensitivity, before applying them to X_{ref}. The resulting intermediate profiles X_{AVK} have altitude resolution comparable to the GOME-2 profiles (or coarser), but do not use the a-priori profile.

In nearly all cases, the validation of GOME-2 profiles (*Figure 3.2*) gives almost the same results for the three versions of smoothed reference profiles X_{ref}, X_{AVK}, and $X_{AVK\ apriori}$.
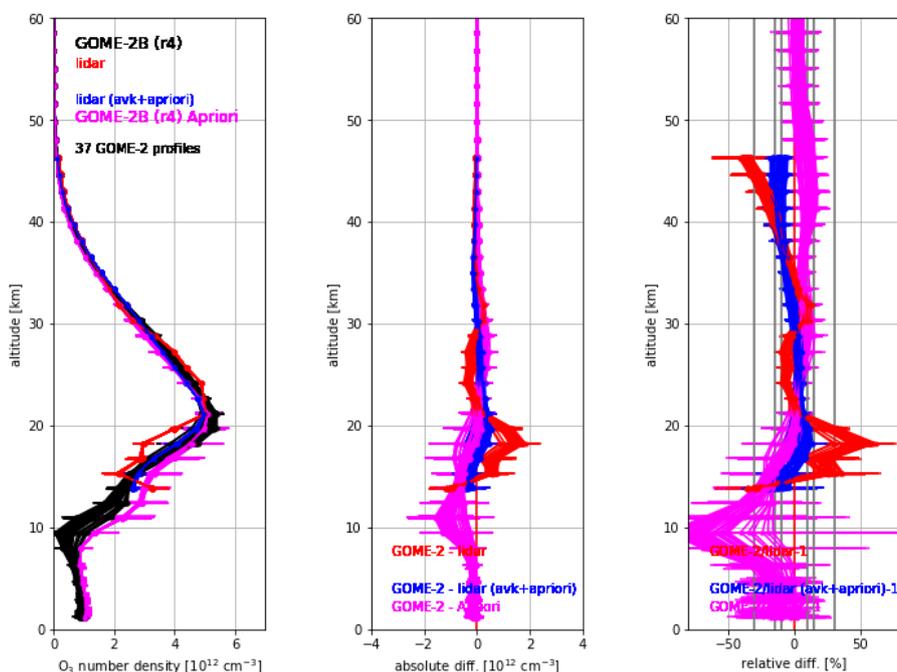
Figure 3.2: Example for the comparison of a lidar profile at Hohenpeissenberg (15.4.2013), Germany, (red Xref, purple Xavk, blue Xavk, apriori) with the matching GOME-2 MetOp-B high-resolution profiles (black). Left panel: Profiles. Middle panel: Absolute differen differences. Right panel: Relative differences. Note that the GOME-2 layer altitudes and averaging kernels vary slightly from profile to profile. This results in small differences in the smoothed lidar profiles. Error bars (1σ) are from the reported measurement uncertainties for GOME-2 and lidar. The vertical lines at ±30%, ±15%, and ±10% in the right panel are the threshold, target, and optimum accuracies specified for the GOME-2 product.

3.5 *Results*

3.5.1 Validation of GOME-2 MetOp-A and B reprocessed ozone profile products:

This summary contains validation results for the reprocessed GOME-2A/B high resolution (HR) ozone profile products, retrieved by the Ozone Profile Retrieval Algorithm (OPERA) at KNMI. The validation period covers from January 2007 to December 2018 for MetOp A and from January 2013 to December 2018 for MetOp-B.

To report the quality of GOME-2 ozone profile products in a very condensed way, the statistics for the different output levels of GOME-2 can be reduced to two layers: Lower Stratosphere (up to an altitude of 30 km) and Upper Stratosphere (above 30 km, up to 50 or 60 km). Table 3.2 shows the definition of the height ranges for lower and upper stratosphere for different latitude belts used in this report.

Table 3.2: Definition of the ranges in km for lower and higher stratosphere for the different latitude belts.

	Lower Stratosphere	Upper Stratosphere
Polar Region	12 km – 30 km	30 km – 50 km
Mid-Latitudes	14 km – 30 km	30 km – 50 km
Tropical Region	18 km – 30 km	30 km – 50 km

The validation for the lower stratosphere is made using ground-based ozonesonde data as a reference. For the upper stratosphere, ground-based lidar and microwave radiometer data are used as reference.

Relative differences (Eq. 1) are calculated against the ground-based reference data. Usually these are also convolved with the averaging kernels, including the a-priori contribution (Smoothed ground-based):

$$(X_{GOME-2} - X_{gb\_smoothed})/X_{gb\_smoothed} \qquad (3)$$

Table 3.3 and Table *3.4* summarize the overall difference between GOME-2A and B ozone profiles and ground-based reference profiles for the time periods from January 2007 to December 2018 (Table 3.3, MetOp-A), and from January 2013 to December 2018 (*Table 3.4*, MetOp-B). The statistics are shown for the lower and upper stratosphere, and for operational near-real-time (NRT) data, and for the reprocessed data record validated here. Tropospheric ozone is discussed earlier in this report. The statistics for the lower stratosphere are obtained by KMI, the statistics for the upper stratosphere by DWD.

Table 3.3: Absolute Differences (AD), Relative Differences (RD) and standard deviation (STDEV) of GOME-2A HR ozone profile products versus ground-based reference profiles for lower and upper stratosphere and different latitude belts. Results are for the time period January 2007 to December 2018. The upper table is for the operational NRT data and the lower table for the reprocessed data record.

	GOME-2A HR operational NRT					
	Lower Stratosphere			Upper Stratosphere		
	AD (DU)	RD (%)	STDEV (%)	AD (DU)	RD (%)	STDEV (%)
Northern Polar Region				-19.9	-43.7	21.4
Northern Mid-Latitudes				-17.9	-35.2	11.9
Tropical Region				-19.4	-3.6	8.9
Southern Mid-Latitudes				-21.2	-34.6	12.6
Southern Polar Region				-	-	-

	GOME-2A HR reprocessed					
	Lower Stratosphere			Upper Stratosphere		
	AD (DU)	RD (%)	STDEV (%)	AD (DU)	RD (%)	STDEV (%)
Northern Polar Region	-9.4	-3.4	14.3	0.5	0.4	19.9
Northern Mid-Latitudes	-2.2	-0.4	9.0	-1.8	-2.5	6.3
Tropical Region	1.8	2.1	9.2	-2.8	-4.7	4.9
Southern Mid-Latitudes	4.8	3.3	14.8	-1.5	4.2	6.1
Southern Polar Region	-1.5	4.1	38.7	**-**	**-**	**-**

[*]The relative difference statistics are derived as a weighted average over the lower- and upper stratospheric ozone profile levels. The absolute differences however are integrated over respectively the lower- and upper stratospheric ozone profile levels.

Table 3.4: Absolute Differences (AD), Relative Differences (RD) and standard deviation (STDEV) of GOME-2B HR ozone profile products versus ground-based reference profiles for lower and upper stratosphere and different latitude belts. Results are for the time period January 2013 to December 2018. The upper table is for operational NRT data and the lower table for the reprocessed data record.

GOME-2B HR operational NRT

	Lower Stratosphere			Upper Stratosphere		
	AD	RD	STDEV	AD	RD	STDEV
	(DU)	(%)	(%)	(DU)	(%)	(%)
Northern Polar Region				-10.9	-23.3	29.6
Northern Mid-Latitudes				-2.6	-7.7	15.8
Tropical Region				-6.3	13.3	7.4
Southern Mid-Latitudes				-7.7	-11.3	16.4
Southern Polar Region				-	-	-

GOME-2B HR reprocessed

	Lower Stratosphere			Upper Stratosphere		
	AD	RD	STDEV	AD	RD	STDEV
	(DU)	(%)	(%)	(DU)	(%)	(%)
Northern Polar Region	0.1	1.3	27.2	-1.2	-1.6	10.9
Northern Mid-Latitudes	3.3	2.0	8.8	-1.9	-2.0	7.6
Tropical Region	7.0	5.5	8.3	-4.8	-5.8	6.3
Southern Mid-Latitudes	13.1	7.0	10.5	-1.1	-4.5	8.7
Southern Polar Region	6.8	8.8	59.4	-	-	-

*The relative difference statistics are derived as a weighted average over the lower- and upper stratospheric ozone profile levels. The absolute differences however are integrated over respectively the lower- and upper stratospheric ozone profile levels.

The optimal goal (10% accuracy), stated in the GOME-2 ozone profile ATBD is met by the reprocessed data in both lower and upper stratosphere for nearly all belts under consideration (but was not always achieved by the operational NRT data).

The time series in Figure 3.3 and Figure 3.4 demonstrate that the reprocessed ozone data from GOME-2A improved the monthly difference between GOME-2A and the reference ground-based MWR and lidars, especially in the upper stratosphere above 40 km. There are two major improvements: The degradation correction applied in the reprocessed data (red lines) removed the more than -60% per decade drift of the operational NRT data (blue lines) in the upper stratosphere. This large drift was caused by aging and degradation of the satellite instrument, which were not accounted for well enough in the operational NRT processing. In the winter months at the Ny Alesund, higher deviations between GOME-2A and the ground-based measurements remain for the reprocessed data (red lines). Fig. 3.5 shows an example, where the GOME-2A retrieval clearly does not reproduce the ground-based profiles very well, typically in late winter / early spring.
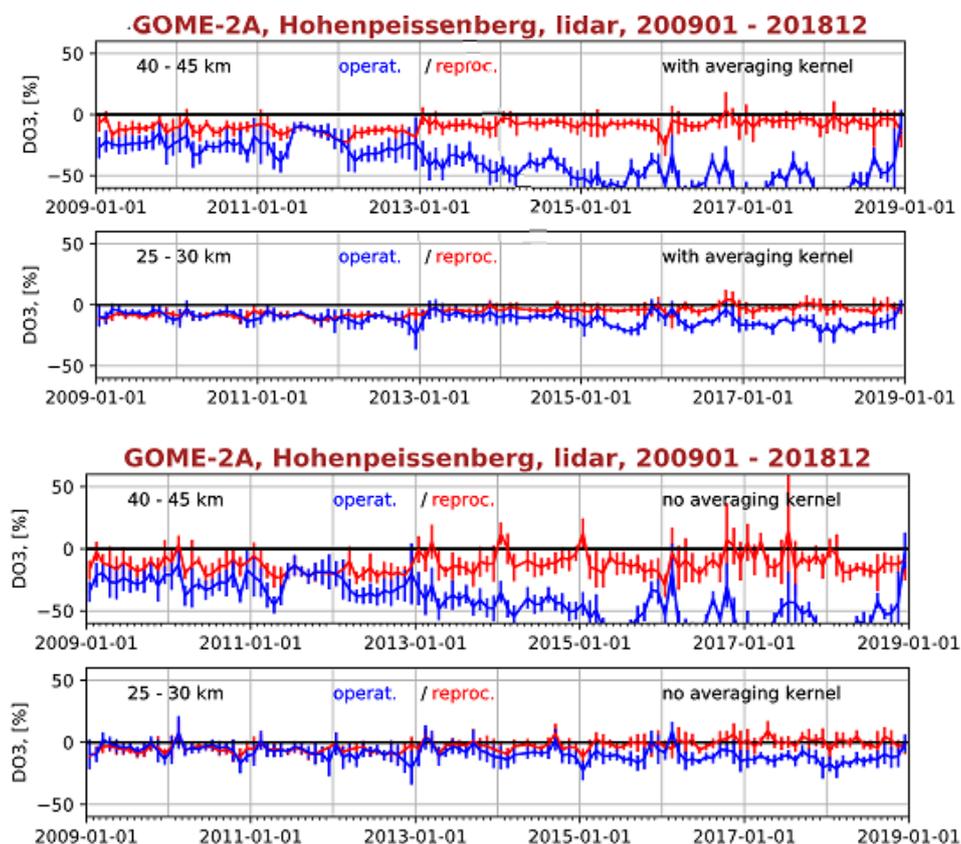
***Figure 3.3:** Time series of monthly mean difference with averaging (upper panel) and without averaging kernel (lower panel) between GOME-2 MetOp-A operational (blue) as well as reprocessed data record (red) and NDACC lidar ground-based ozone measurements at Hohenpeissenberg.*
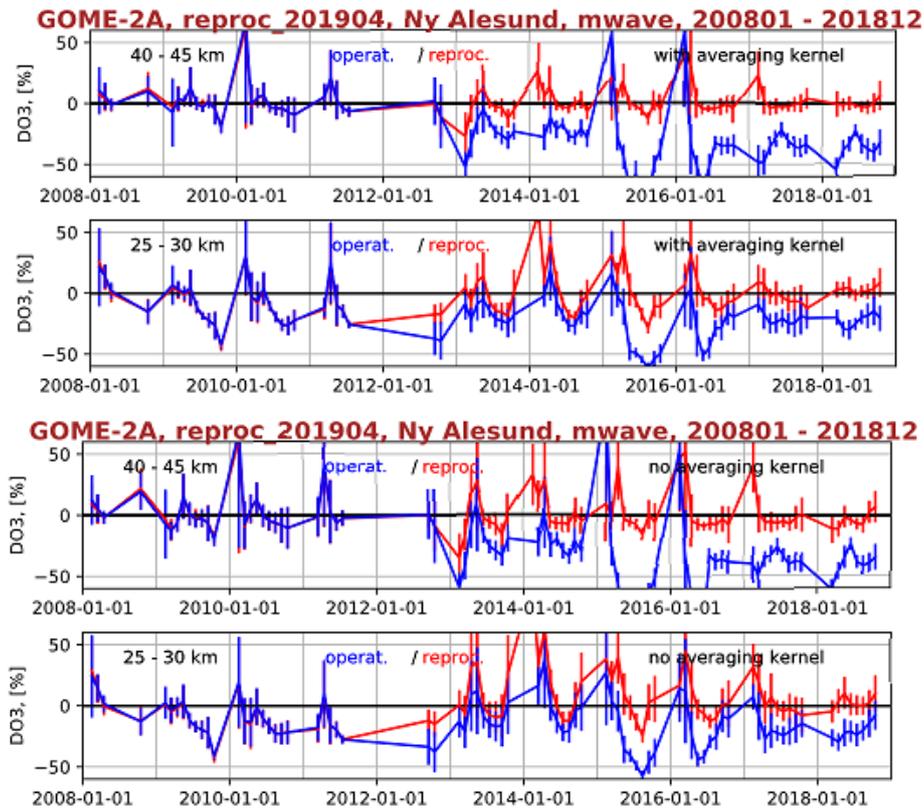
Figure 3.4: Time series of monthly mean difference with averaging (upper panel) and without averaging kernel (lower panel) between GOME-2 MetOp-A operational (blue) as well as reprocessed data record (red) and NDACC ground-based MWR ozone measurements at Ny Alesund.

The second major improvement is most visible at higher latitudes from 01/2015 to 03/2016, e.g. in the upper panels of Figure 3.4. There the operational NRT data (blue lines) show a much larger seasonally varying bias, which is also different from the seasonal bias in other years. This does not occur in the reprocessed data (red lines), which show similar seasonal bias in all years. The 2015/2016 change in the operational data is due to a change in the meteorological data files provided by ECMWF. It resulted in using data from the wrong hemisphere, from 01/2015 (when the ECMWF data format changed), until 03/2016 (when the error was noticed and corrected in the operational processing). In the reprocessed data, this error does not occur, and the reprocessed 2015/2016 results are consistent with other years.

Figure 3.5 shows an example of an individual comparison between a microwave profile in early spring measured at Ny Alesund with the matching GOME-2 MetOp-A high resolution profiles. In nearly all cases, the validation of GOME-2 profiles (*Figure 3.2*) does not agree with the three versions of smoothed reference profiles X_{ref}, X_{AVK}, and $X_{AVK\ apriori}$.
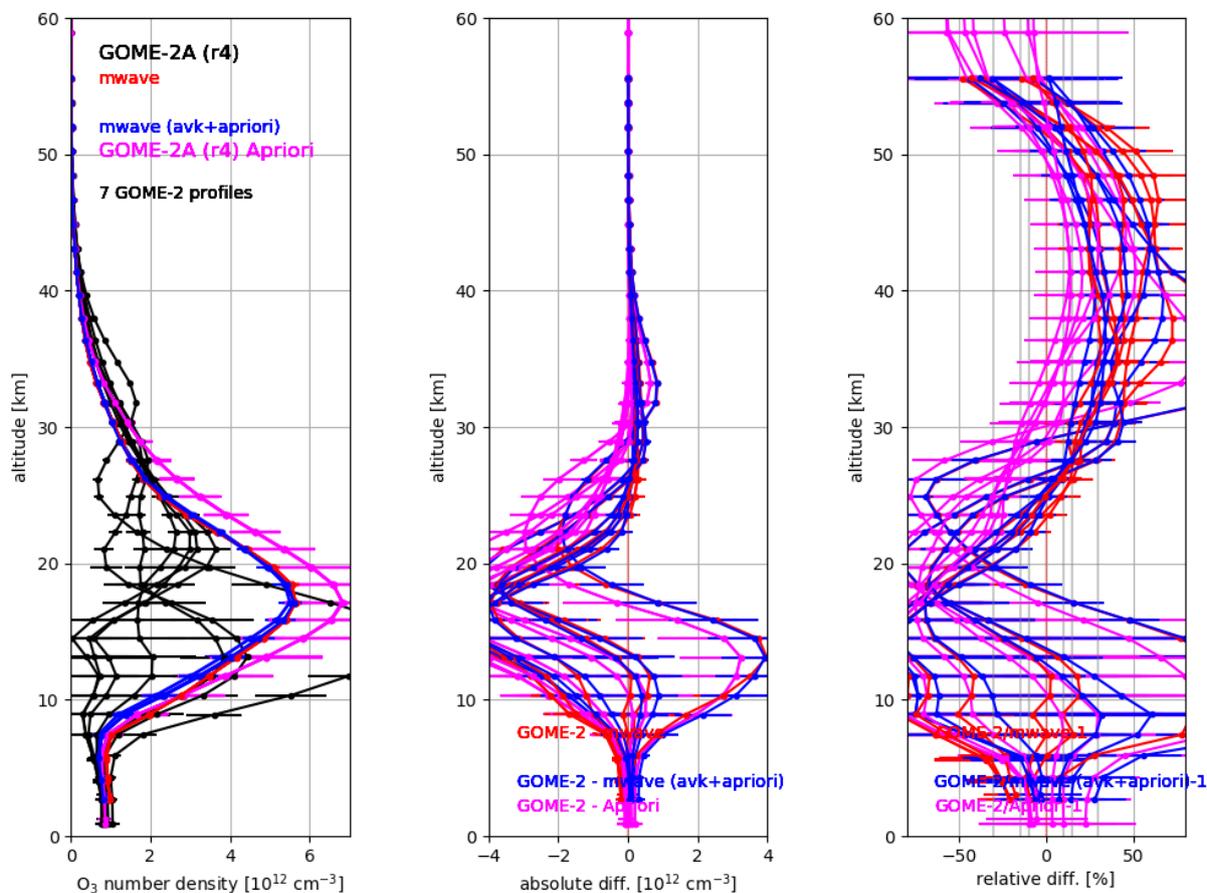
Figure 3.5: Example for the comparison of a microwave radiometer profile at Ny Alesund (20.2.2016), (red Xref, purple Xavk, blue Xavk, apriori) with the matching GOME-2 MetOp-A profiles (black). Left panel: Profiles. Middle panel: Absolute differences. Right panel: Relative differences. Note that the GOME-2 layer altitudes and averaging kernels vary slightly from profile to profile. This results in small differences in the smoothed microwave radiometer profiles. Error bars (1σ) are from the reported measurement uncertainties for GOME-2 and microwave radiometer. The vertical lines at ±30%, ±15%, and ±10% in the right panel are the threshold, target, and optimum accuracies specified for the GOME-2 product.

The GOME-2B time series in Figures Figure 3.6 and Figure 3.7 demonstrate that the reprocessed ozone data from GOME-2B also improved the systematic difference between GOME-2B and the reference ground-based MWR and lidars, especially in the upper stratosphere above 40 km. Again, the reprocessing removed the more than -40% per decade drift of the operational NRT GOME2 data. The effect of wrong ECMWF meteorological data (swapped hemisphere from January 2015 until March 2016) in the NRT processed data was also corrected. Overall, the quality of GOME2 MetOp A and B profiles in the reprocessed data record has improved substantially compared to the operational NRT data. The reprocessed data fulfil the target accuracy in the upper stratosphere (better than 15%) at most stations throughout the years, and have long-term stability better than 5% per decade.
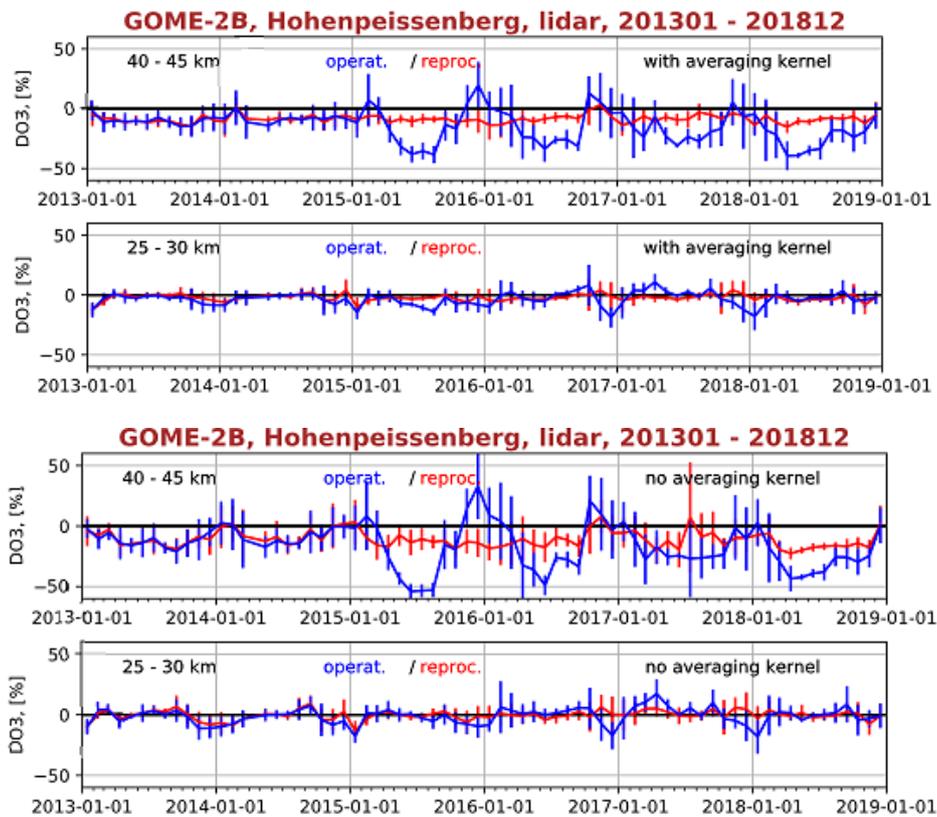
Figure 3.6: Time series of monthly mean difference with averaging (upper panel) and without averaging kernel (lower panel) between GOME-2 MetOp-B operational (blue) as well as reprocessed data (red) and NDACC lidar ground-based ozone measurements at Hohenpeissenberg.
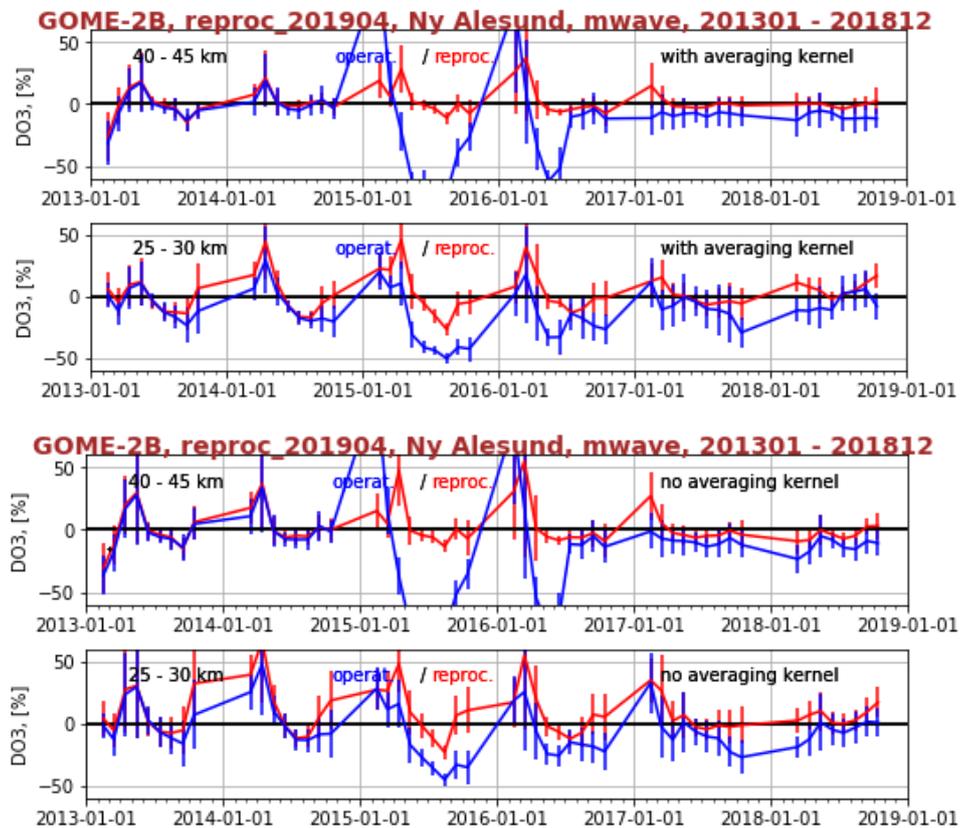
Figure 3.7: Time series of monthly mean difference with averaging between GOME-2 MetOp-B operational (blue) as well as reprocessed data (red) and NDACC microwave radiometer measurements at Ny Alesund.

Figure 3.8: Scatter plot of ozone differences GOME-2 MetOp-A – gound-based versus solar zenith angle (upper left) and total ozone column (upper right), versus scan angle (bottom left) as well as versus cloud fraction (bottom right) for different altitudes at Hohenpeißenberg.
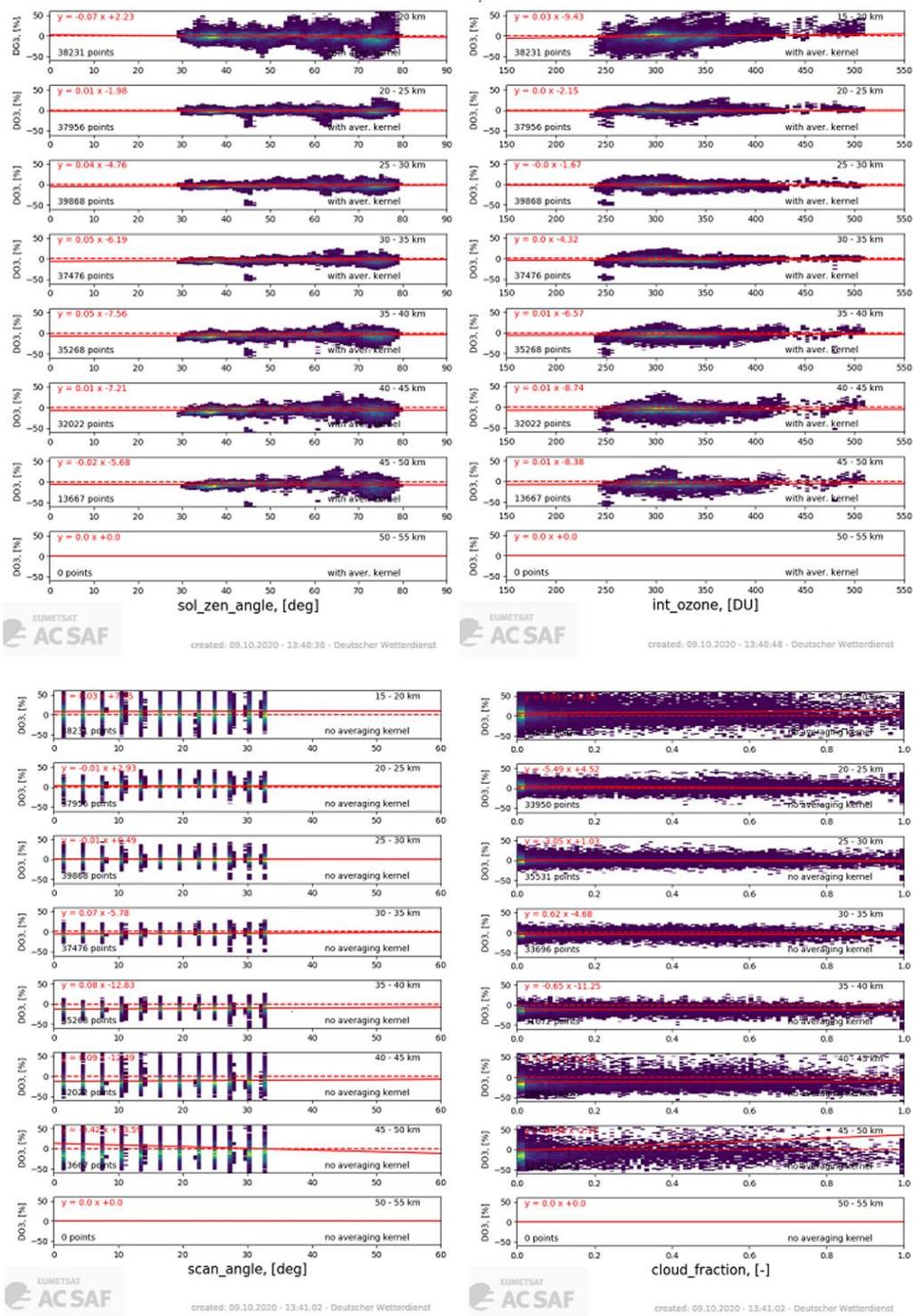
Exemplary results on the (lack of) correlation of differences between GOME-2A reprocessed ozone and ground-based ozone with important geometrical or geophysical parameters of the satellite measurement are presented in Figure 3.8. The reprocessed data do not vary significantly with solar zenith angle, cloud fraction, or scan angle. At most altitudes there is also no indication for a significant variation with total ozone column. The only exception is the lower stratosphere at higher latitudes (not shown), where data from Ny-Alesund and Lauder indicate that below 25 km GOME-2 tends to underestimate ozone when the total ozone column is low, and tends to overestimate ozone when the total ozone column is high. Figure 3.5 shows an example for the problem, and this seasonally varying bias remains for GOME-2A and B reprocessed ozone profile data at high latitudes. At mid-latitudes the seasonal bias variation is reduced in the reprocessed data (Figures Figure 3.3 and Figure 3.6).

Overall, these validation results show that GOME-2A and B reprocessed ozone profiles are of good quality and are much improved compared to the operational NRT processing. In the stratosphere, the reprocessed ozone data fullfil the ±10% optimal accuracy goal over a wide range of conditions, and the ±15% target accuracy under almost all conditions. An avenue for further improvement would be to improve the accuracy and uncertainty of the retrieval at high latitudes in early spring.

4. Integrated profiles validation using ground-based measurements

4.1 *Dataset description*

4.1.1 **GOME-2/MetOpA and GOME-2/MetOpB data**

The GOME-2/MetOpA and GOME-2/MetOpB (hereafter GOME-2A and GOME-2B) integrated ozone profiles were retrieved by the same algorithm and methodology that is described in the "Vertical Ozone Profile and Tropospheric Ozone Column Products" ATBD (Tuinder, 2019). The GOME-2A and GOME-2B integrated ozone profile datasets used in this validation report cover the following time periods:

- GOME-2A: 24 October 2007 – 31 December 2019

- GOME-2B: 13 December 2012 – 31 December 2019

The reprocessing of the ozone profiles was applied to GOME-2A and GOME-2B data until December 31, 2018. The operational algorithm was applied to the data since 2019, which are also used in this analysis, to demonstrate the continuity of the time-series.

To further establish the validation results of the GOME-2A and GOME-2B integrated ozone profiles, they are also compared to the respective operational total ozone column (TOC) products retrieved by the GDP v.4.8 algorithm.

4.1.2 **Ground-based data**

The ground-based measurements database used for this validation report consists of archived Brewer and Dobson total ozone data that are downloaded from the World Ozone and Ultraviolet Radiation Data Centre (http://www.woudc.org). WOUDC is one of the World Data Centers which are part of the Global Atmosphere Watch (GAW) program of the World Meteorological Organization (WMO). These data are quality controlled, first by each station and secondly by WOUDC.

For the quality of the reference ground-based data used for the validation of the GOME-2A and GOME-2B integrated ozone profiles products, updated information were extracted from recent inter-comparisons and calibration records. This continuously updated selection of ground-based measurements has already been used numerous times in the validation and analysis of global total ozone records such as the inter-comparison between the OMI/Aura TOMS and OMI/Aura DOAS algorithms (Balis et al., 2007a), the validation of ten years of GOME/ERS-2 ozone record (Balis et al., 2007b), the validation of the updated version of the

OMI/Aura TOMS algorithm (Antón et al., 2009), the GOME-2/MetOp-A validation (Loyola et al., 2011; Koukouli et al., 2012), the GOME-2/MetOp-B validation (Hao et al., 2014), the evaluation of the European Space Agency's Ozone Climate Change Initiative project (O3-CCI) TOCs (Koukouli et al., 2015, Garane et al., 2018) and the validation of the TROPOMI/S5P total ozone products (Garane et al., 2019). In all the aforementioned publications, LAP/AUTH assumes the leading role in the validation efforts.

In this report, archived data for the period January 2007 to December 2019 are used for the comparisons, depending on the availability of data for each individual station. The Brewer and Dobson WOUDC stations considered for the comparisons are listed in Tables A.1 and A.2 (Appendix 1) and their geographical distribution is depicted in Figure 4.1. In Figure 4.2, the distribution of the co-locations of the ground-based measurements in space in time are shown.
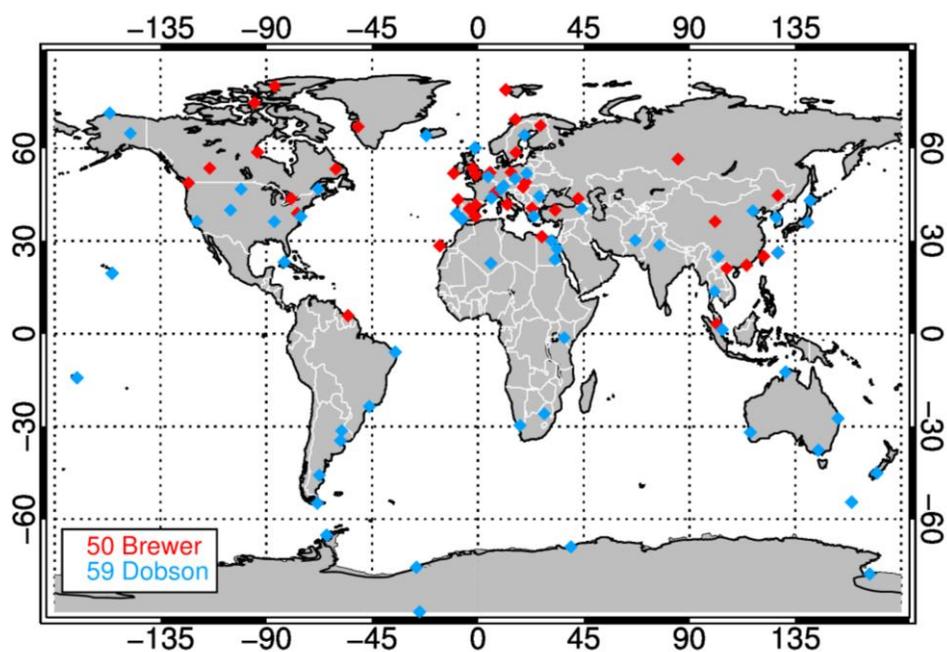

Figure 4.1: Spatial distribution of the Brewer and Dobson ground-based stations used for the comparisons.
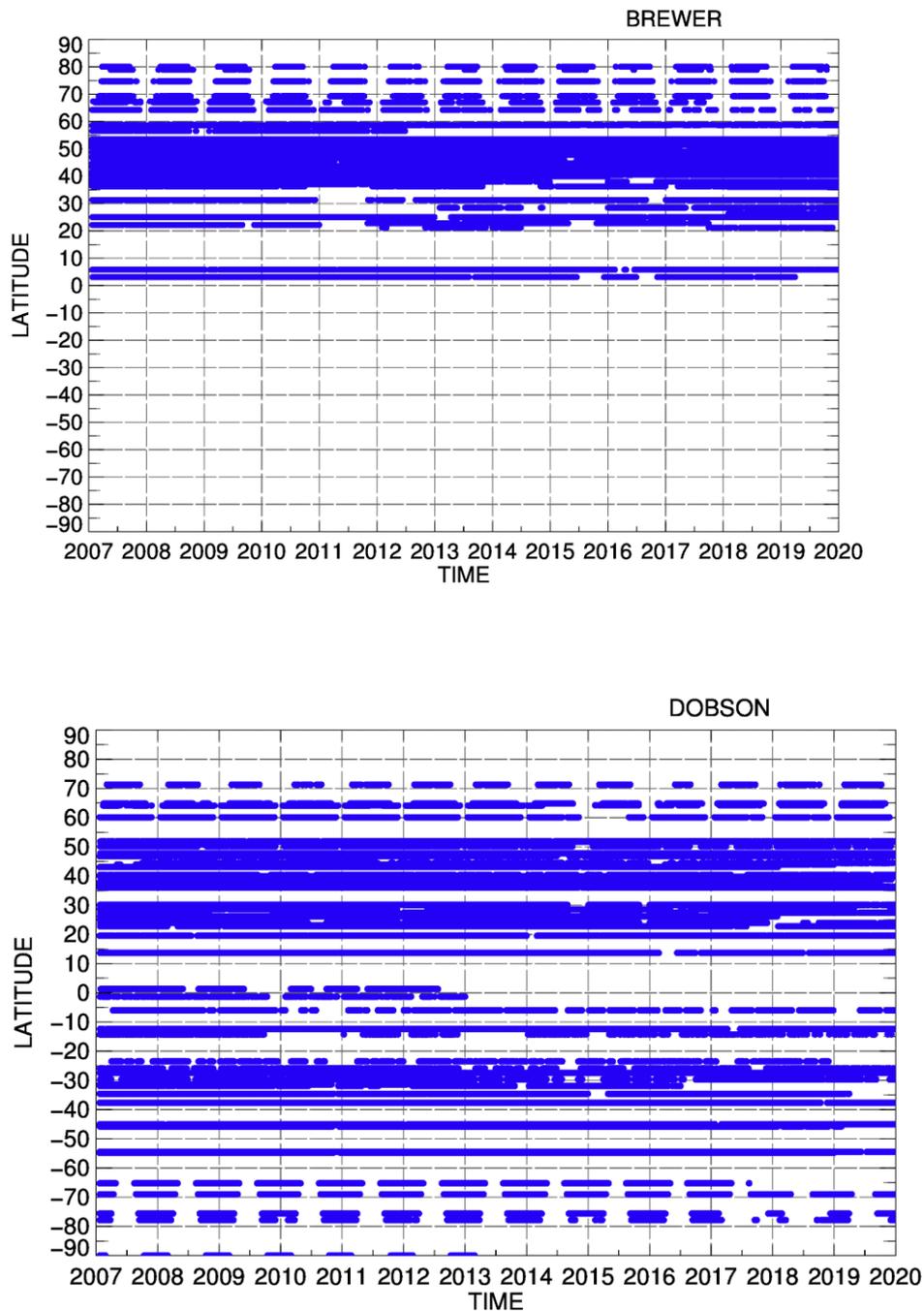
Figure 4.2 : Spatial and temporal representation of the co-location data used for the validation with ground-based measurements (upper panel: Brewer, lower panel: Dobson) for the time period of the GOME-2A operation (until December 2019).

In the comparison plots and statistics presented in this report, only the direct sun observations provided by the Brewers and Dobsons are utilized for the computation of the percentage differences between satellite and co-located (in space and in time) ground-based measurements, since they are considered of higher accuracy than all the other types of ground-based observations. Finally, only northern hemisphere Brewer ground-based stations are considered, because the number of stations in the southern hemisphere is very limited and they are mainly located in Antarctica.

4.2 *Validation of GOME-2A and GOME-2B integrated ozone profiles*

In this section, the archived and quality-controlled Dobson and Brewer daily total ozone measurements downloaded from WOUDC for the period January 2007 to December 2019, are used as ground-truth for the validation of GOME-2A and GOME-2B reprocessed integrated ozone profiles. The datasets of the two satellite sensors are temporally and spatially co-located to ground-based measurements using the following co-location criteria:

- the satellite and daily ground-based total ozone measurements must correspond to the same day, and

- the maximum search radius between the ground-based stations and the centre coordinates of the satellite pixel is set to 150 km. The spatially closest satellite observation is paired with the ground-based station's daily-mean measurement.

This is a rather different approach compared to the validation methodology used in Sections 2 and 3, where each ground-based profile is compared with more than one co-located satellite profiles, but it is an established methodology followed in many total ozone validation reports in the past (either for integrated ozone profiles or for the operational products), as well as in numerous published papers for total ozone column validation (for example, Koukouli et al., 2015; Garane et. al. 2018; Garane et. al, 2019; Garane et al., 2020). Namely, in the work of Garane et al, 2019, different sampling strategies have been compared for the validation of TROPOMI/S5P total ozone columns, where individual measurements from the EUBREWNET Brewers were used along with the WOUDC dataset of daily mean ground-based measurements, with no significant differences in the validation results.

The pairs of co-located satellite and daily-mean ground-based measurements are used to calculate their percentage difference by the simple formula:

$$Diff\,(\%) = \frac{(satellite - ground)}{ground}\%$$

The datasets of percentage differences are then filtered:

- for solar zenith angle (SZA), which is limited up to 83°, because the mean percentage differences of the co-locations with SZA above 83° were higher than -10 %. The number of co-locations affected by this filtering criterion is ~ 1.5 % of the total.

- for latitude, which is limited up to 85° S because the mean percentage differences of the co-locations with latitude above 85° S were higher than + 20 %. The number of co-locations affected by this filtering criterion is below 0.7% of the total.

The monthly means that are shown in the respective time-series plots are calculated by averaging the total number of available co-locations per month. Furthermore, the error bars in the following plots (where they are shown) stand for the 1σ standard deviation of the means.
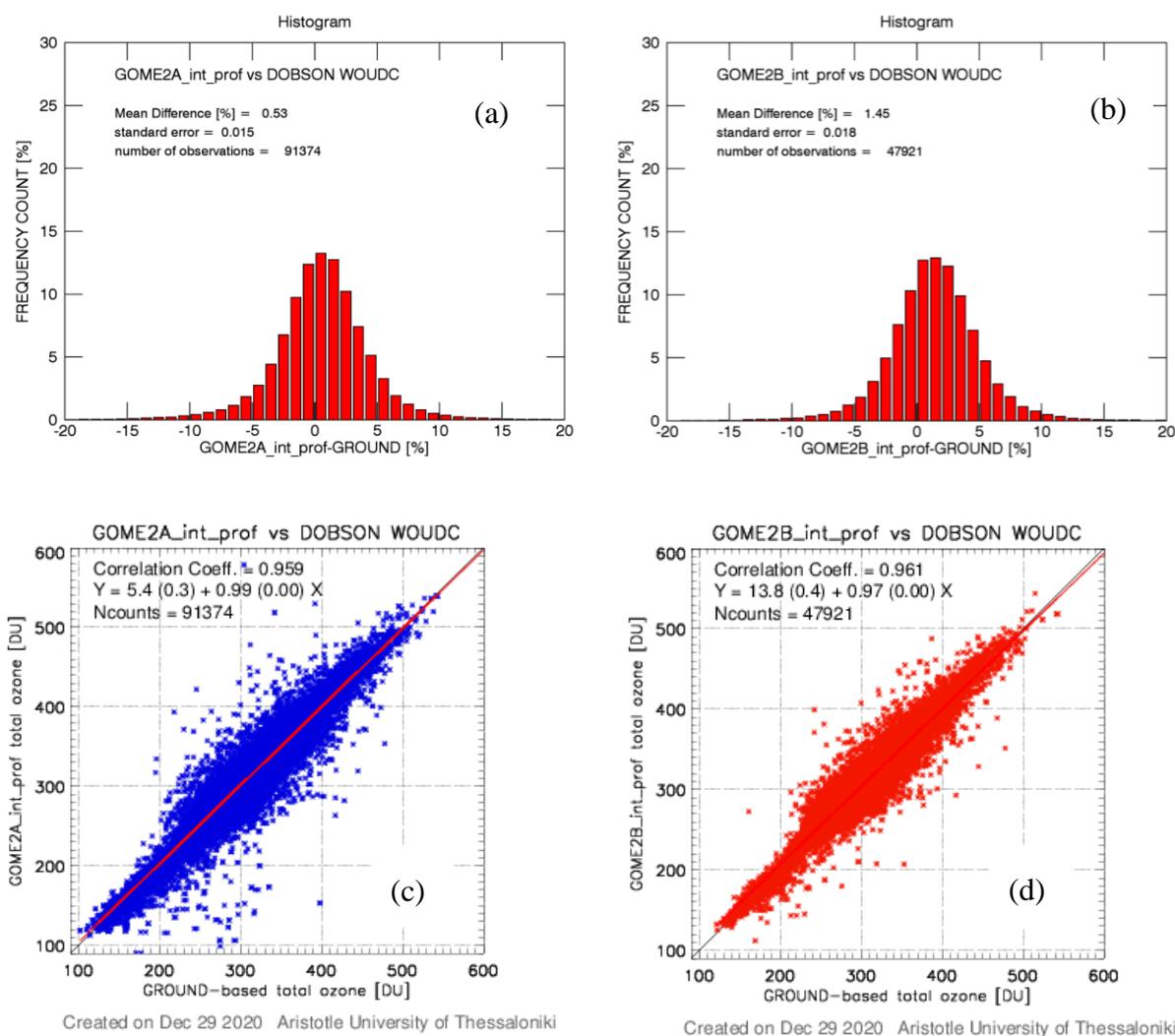

Figure 4.3 : Panels (a) and (b): Histograms showing the distributions of the percentage differences of the GOME-2A (panel a) and GOME-2B (panel b) co-locations to Dobson ground-based total ozone measurements. Panels (c) and (d): Scatter plots of the co-located GOME-2A (panel c) and GOME-2B (panel d) integrated ozone profile to ground-based TOC measurements from Dobson instruments.

For the purposes of this validation report, in Sections 0 and 4.2.1.3 only the temporally common co-locations to ground-based measurements between the GOME-2A and GOME-2B are used to achieve the comparability between the datasets.

4.2.1 Validation results of GOME- 2A and GOME-2B integrated ozone profiles with respect to ground-based measurements

Figure 4.3 shows the overall statistical analysis of the GOME-2A and GOME-2B reprocessed integrated ozone profiles' co-locations to Dobson ground-based total ozone measurements. In the upper part of the plot (panels a and b), the distributions of the percentage differences of the co-locations are shown to be normal around the centre value, which is 0.5% for the GOME-2A (panel a) and 1.5% for the GOME-2B (panel b) comparisons. The scatter plots, shown in the panels c and d, reflect the very good overall agreement (correlation coefficient = 0.96) of both sensors' integrated ozone profiles to the ground-based TOC measurements from Dobson instruments. The respective correlation coefficients for the Brewer comparisons (not shown here) is 0.98, resulting from nearly 132.500 (for GOME-2A) and 80.800 (for GOME-2B) co-locations.

4.2.1.1 Temporal evolution of the GOME-2A and GOME-2B comparisons to ground-based measurements

 Figure 4.4 shows the full time series of the monthly mean percentage differences of GOME-2A (blue line and symbols) and GOME-2B (red line and symbols) with respect to the co-located (in space and in time) ground-based measurements. The comparisons to Dobson measurements are shown in panels a (northern hemisphere) and b (southern hemisphere), while in panel c the Brewer comparisons are displayed (northern hemisphere only). The green vertical line denotes the end of the reprocessed dataset in December 2018.

The first noticeable feature of the GOME-2A time-series is the abrupt change of level in the comparisons during the second half of 2013, which is present in both Brewer and Dobson co-locations and in both hemispheres. On 15th of July 2013, the GOME-2A swath width changed from 1920 to 960 km. This means that the instrument now has more pixels in a more nadir direction than before. This can lead to a change in the averaged ozone profile in case there is a cross track bias. Also, there is a discontinuity in mid-2013 in the degradation correction, because this correction is cross-track-pixel / scan angle dependent. As a result, the GOME-2A time-series is studied separately, before and after mid-2013, in terms of mean bias of the differences to ground-based measurements.

The mean relative bias $\pm 1\sigma$ (in %) and mean standard deviation for GOME-2A (before and after mid-2013) and GOME-2B with respect to ground-based measurements from Brewer and Dobson instruments, are summarized in Table 4.1. Since the start of its operation until mid-2013, GOME-2A has a mean bias of ~ ± 0.3 %. For the second part of its time series, the mean bias increases to ~ +2 %, very close to the GOME-2B mean bias which is 1.5 – 2.0 %.

The mean standard deviation, which includes the sensors' and the ground-based measurements' variability, is almost the same for the two instruments, 3-3.5% for the NH and 4-4.5% for the SH, where the number of Dobson stations is smaller.

(a)

(c)

(b)
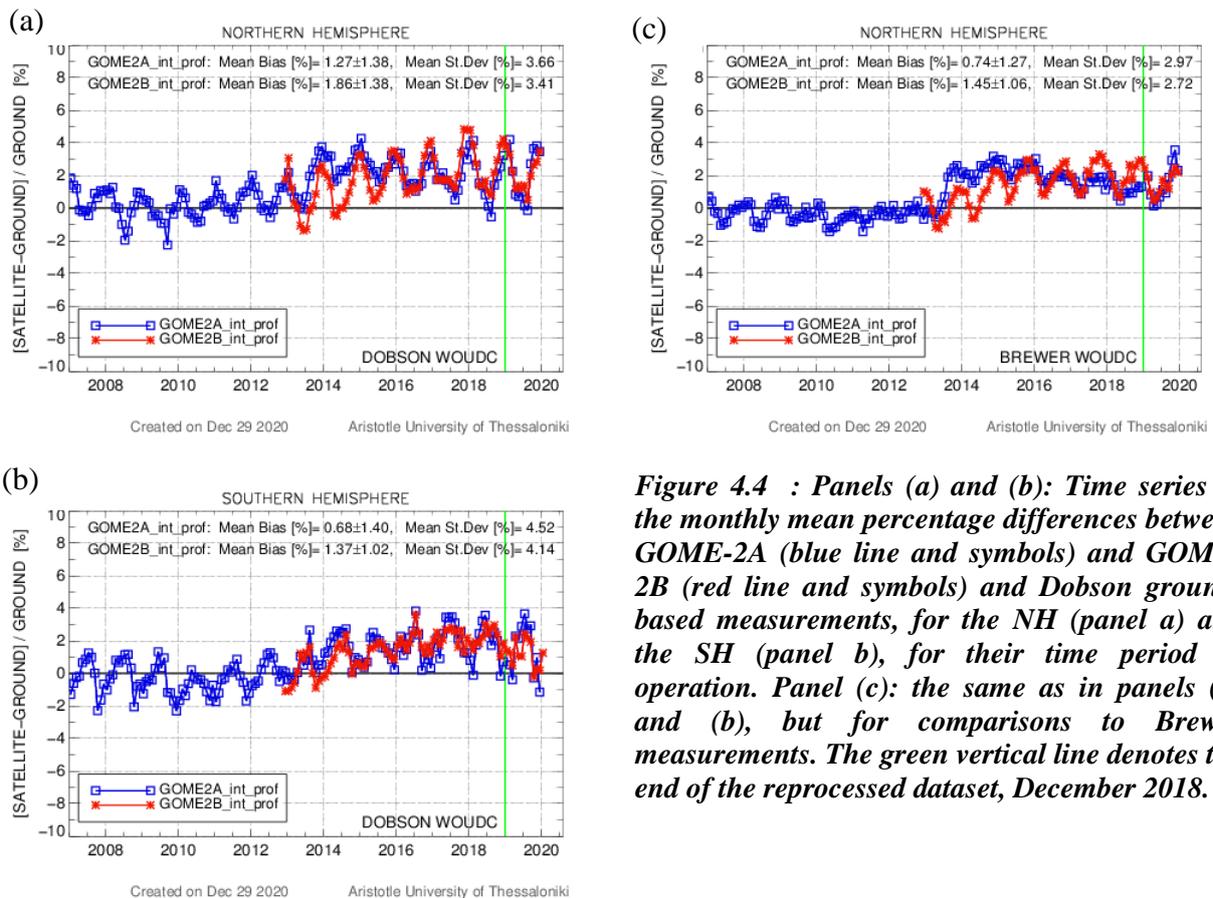
Figure 4.4 : Panels (a) and (b): Time series of the monthly mean percentage differences between GOME-2A (blue line and symbols) and GOME-2B (red line and symbols) and Dobson ground-based measurements, for the NH (panel a) and the SH (panel b), for their time period of operation. Panel (c): the same as in panels (a) and (b), but for comparisons to Brewer measurements. The green vertical line denotes the end of the reprocessed dataset, December 2018.

Table 4.1 : The statistical analysis (mean bias in % ± mean standard deviation in %) of the comparisons of GOME-2A (before and after 2013) and GOME-2B to Dobson and Brewer ground-based measurements, for their time period of operation until December 2019.

		GOME-2A (2007-2013)	GOME-2A (2013-2019)	GOME-2B (2013 – 2019)
DOBSON	Mean Bias± 1σ % (NH)	+0.3 ± 0.9	+2.3 ± 1.1	+1.9 ± 1.4
	Mean St. Dev. % (NH)	3.7	3.7	3.4
	Mean Bias ± 1σ % (SH)	-0.3 ± 0.9	+1.7 ± 1.1	+1.4 ± 1.0
	Mean St. Dev. % (SH)	4.6	4.4	4.1
BREWER	Mean Bias ± 1σ % (NH)	-0.3 ± 0.5	+1.9 ± 0.7	+1.5 ± 1.1
	Mean St. Dev. % (NH)	3.0	2.9	2.7

As for the consistency between the two sensors, it is very clear that, since the start of their common operation until mid-2015 they deviate by 1-2%, with GOME-2A reporting higher ozone values than GOME-2B. This is a feature seen in the NH mainly. Since mid-2015, their performances agree very well in terms of mean bias. Furthermore, the transition from the reprocessed datasets (until December 2018) to the operational data that span the year 2019, is very smooth for both sensors, not introducing any abrupt changes.

Another interesting feature seen in Figure 4.4 is the enhanced seasonality of the GOME-2B differences to Brewer ground-based measurements compared to the seasonality of GOME-2A comparisons. It should be noted here that the seasonality feature of the percentage differences to ground-based measurements is principally studied using Brewers, because the Dobson measurements are known to be highly affected by the effective temperature (see Koukouli et al., 2016). In Figure 4.5 the seasonality of the two sensors compared to Brewer measurements is more clearly seen. GOME-2A is split into two periods as before:

- the dataset of its co-locations to Brewers until mid-2013 is seen in blue dots and has a peak-to-peak seasonality of 1.2%.

- The respective peak-to-peak seasonality for the period after mid-2013 is very similar, 1.3%, proving that the main difference between the two time periods is the absolute level of the sensors' integrated ozone profile.

GOME-2B, on the other hand, has a stronger seasonality compared to GOME-2A, by ~1% peak-to-peak. This is perfectly normal, since the two instruments have possible differences in their calibration and use different wavelength ranges for the retrieval of the ozone profiles, which can lead to different total ozone columns.

To further investigate the quality of the reprocessed integrated ozone profiles, the operational GOME-2A and GOME-2B total ozone column products retrieved with the algorithm GDP v.4.8 are indirectly compared to the respective integrated ozone profiles via their validation with ground-based measurements.
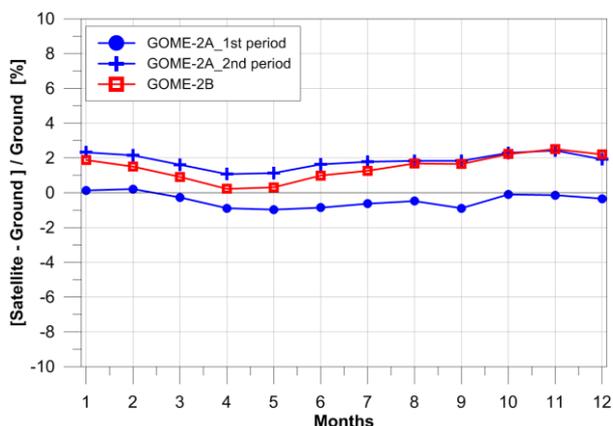

Figure 4.5 : Seasonality of the relative percentage differences for GOME-2A (dots: start of operation until mid-2013, crosses: mid-2013 until Dec. 2019) and GOME-2B, with respect to Brewer measurements.
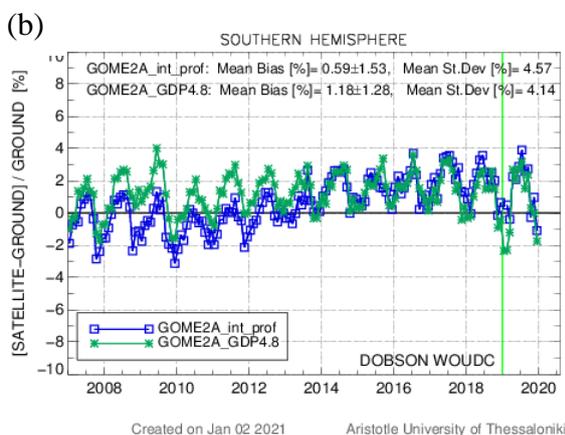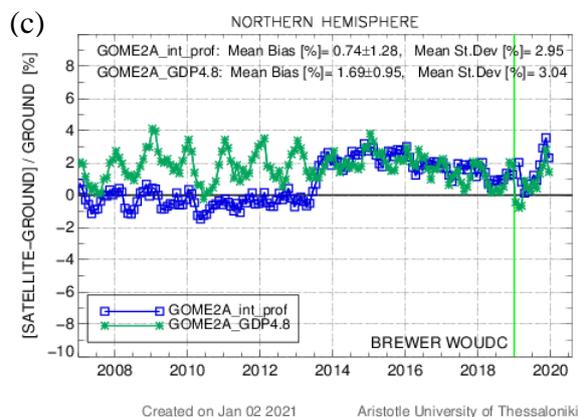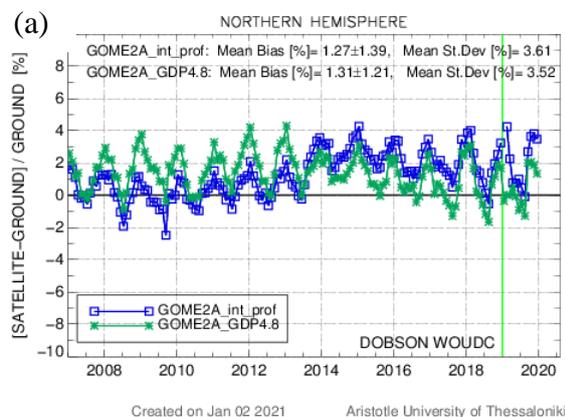
(a)

(c)

(b)

Figure 4.6 : The time series of the percentage differences between ground-based measurements and GOME-2A integrated ozone profiles (blue line and symbols) and GOME-2A total ozone column retrieved with GDP4.8 (green line and symbols), for the northern (panels a and c) and the southern (panel b) hemisphere .

In Figure 4.6 the time series of the percentage differences between ground-based measurements and GOME-2A integrated ozone profiles (blue line and symbols) and the operational total ozone column GOME-2A GDP4.8 (green line and symbols) are shown for the northern (panels a and c) and southern (panel b) hemisphere. The overall difference in mean bias between the two products is very small (up to 0.5%) for the Dobson comparisons, for which the seasonality is known to be enhanced. When the Brewer comparisons are examined, there is a very clear difference before and after mid-2013, in both the seasonality pattern of the two satellite products as well as in the mean bias of their differences to Brewer TOCs:

- before mid-2013, the GDP4.8 TOCs have a peak-to-peak seasonality of 2.5%, and a mean bias of +2%. The respective GOME-2A integrated ozone profiles' seasonality is 1.2% and the mean bias is +0.3%.

- after mid-2013, the GDP4.8 seasonality and mean bias w.r.t. Brewer ground-based measurements are very close to the integrated ozone profiles' comparison.
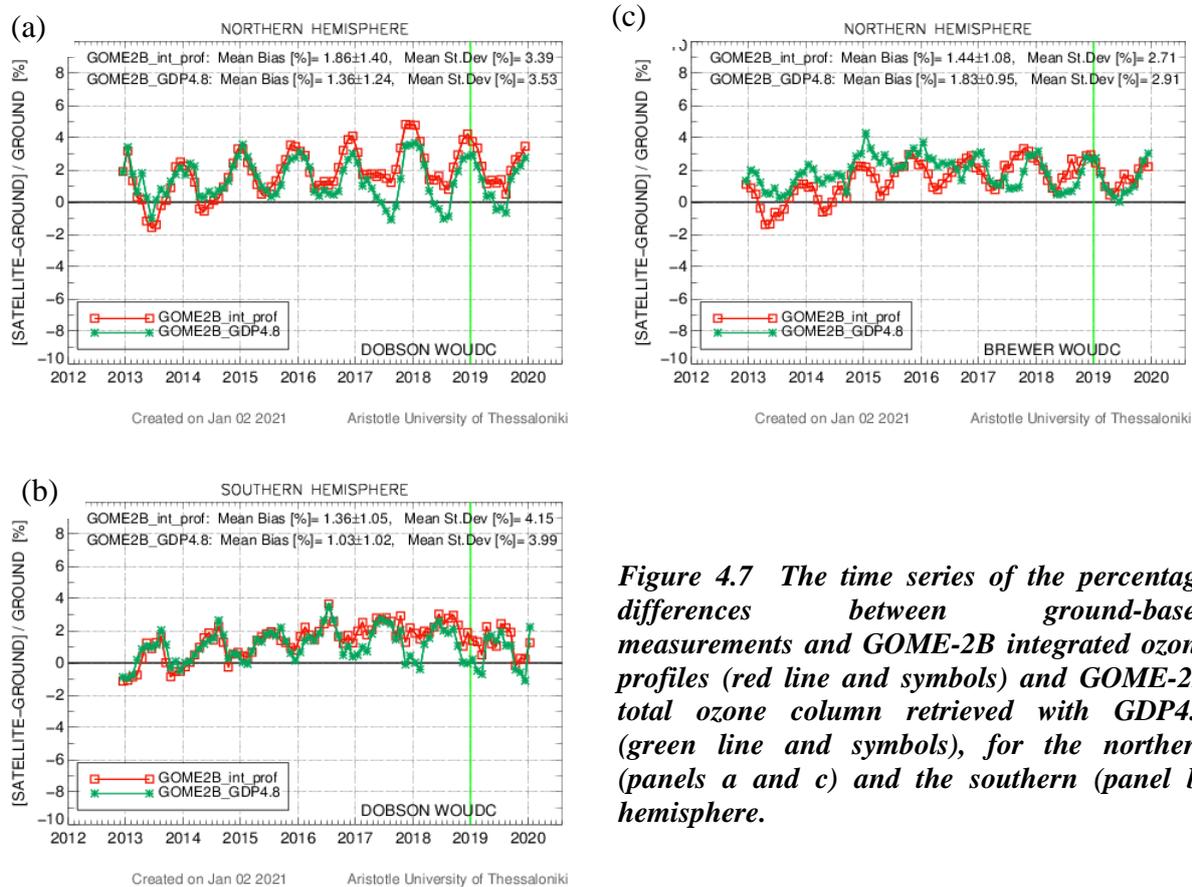
Figure 4.7 The time series of the percentage differences between ground-based measurements and GOME-2B integrated ozone profiles (red line and symbols) and GOME-2B total ozone column retrieved with GDP4.8 (green line and symbols), for the northern (panels a and c) and the southern (panel b) hemisphere.

Similarly, in Figure 4.7 the GOME-2B integrated ozone profile time series (red line and symbols) is indirectly compared to the operational GOME-2B GDP4.8 TOC (green line and symbols). Since mid-2016 the two algorithms start to deviate by ~1.5-2 % (Figure 4.7a and b), especially during summer months in both hemispheres. For the Brewer comparisons (Figure 4.7c) this is seen as a temporal shift with the GDP4.8 total ozone differences having a minimum during summer months, while the integrated ozone profiles have a minimum during spring months. During 2019 (operational data), no shift between the two algorithms is seen. The respective comparison of GOME-2A and GOME-2B operational total ozone products to ground-based measurements (see http://acsaf.physics.auth.gr/eumetsat/) does not show any change in their relative variation during 2016. The most possible reason causing this feature could be the continuous drift of the lower wavelengths in Band 2B (see https://d1qb6yzwaaq4he.cloudfront.net/acsaf/metop/metopb/SpectralDrift/current_SMR_Wav elength_drift_Ch2_full.png) that is used for the GOME-2B ozone profiles retrievals (299 - 412 nm & +/- 0.13 nm & 0.26 - 0.28 nm), resulting to a mis-alignment of the cross sections and the measured spectrum. Another (less likely) reason could be a possible change in the Level-1b data or a software version change of the ozone profiles retrieval algorithm, which was introduced at the time and is listed in Table 11.1 (PPF 6.1, Alg. version 1.12, Software

version 1.35 → 1.37) of the PUM (Tuinder, 2021). In any case, this issue has to be further investigated algorithm wise.

Nevertheless, the differences in the overall mean relative biases (statistics shown at the top of the figures) are very limited, up to ±0.5%, in both hemispheres with respect to both types of ground-based instruments, which is very positive considering the differences in the retrieval methodologies.

4.2.1.2 Latitudinal dependency of the comparisons

In this section and the ones that follow, only temporally common co-locations between GOME-2A and GOME-2B are used. This means that (a) GOME-2A is temporally restricted to mid-December 2012 - December 2019, to match the GOME-2B record and (b) it is verified that only days with available co-locations for both sensors are kept within the dataset. The reason for this restriction is to have two utterly comparable datasets, in terms of the influence variables that are studied onwards, like latitude, solar zenith angle, clouds, etc.

In Figure 4.8 the percentage differences between the reprocessed ozone integrated profiles retrieved by the two sensors and the TOC measurements performed by Dobson (left panel) and Brewer (right panel) ground-based instruments, are averaged in 10° latitude bins and displayed versus latitude. As it follows from the figures, both sensors have higher differences to ground-based measurements in the tropics and mid-latitudes. Close to the north high latitudes the bias decreases. Furthermore, GOME-2B reports lower ozone values compared to GOME-2A mainly in the tropics and the mid-latitudes of both hemispheres. Nevertheless, the differences between the two sensos are always less than 1% regardless of latitude of the co-locations.
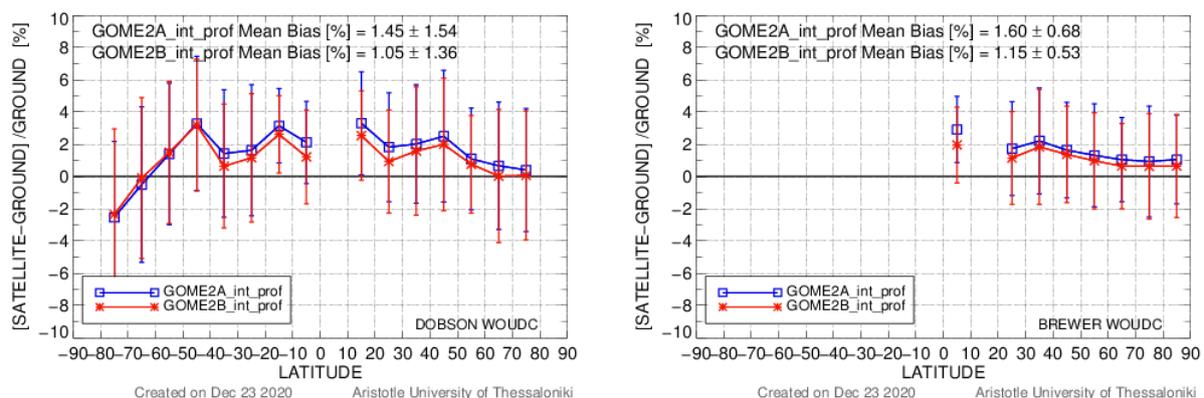

Figure 4.8 : The latitudinal dependency of the percentage differences between the ozone integrated profile retrieved by the three sensors (GOME-2C, GOME-2B and GOME-2A) and the TOC measurements performed by Dobson (left panel) and Brewer (right panel) ground-based instruments.

As for the dependence of the percentage differences on solar zenith angle (SZA), in Figure 4.9 it is seen that the Dobson comparisons below 80° have a small bias, up to ± 1.5%. Above 80° the dependence on SZA is enhanced to -2.5% but the respective number of co-locations is limited and come from stations in the latitude bin -70°S to -80°S. The dependence on SZA is less pronounced for the Brewer comparisons, which come from the northern hemisphere stations only and are not affected as much as Dobson ground-based measurements by seasonality. The GOME-2A overestimation of ~ 0.5 - 1 % compared to GOME-2B for measurements with SZAs that span 30°-60°, is noticed here as well, but the patterns of the dependency for the two sensors is very similar.

It is worth mentioning that according to the Product Requirements Document (Hovilla et. all, 2019), the accuracy requirements for the GOME-2 MetOp-A and MetOp-B Total Ozone product are 4% for SZAs < 80° and 6% for SZAs > 80°. As seen in Figure 4.9, the GOME-2A and -2B accuracy is well within these target values, which proves that the reprocessed integrated ozone profiles product is of similar quality to the operational GOME-2A and GOME-2B TOC products.
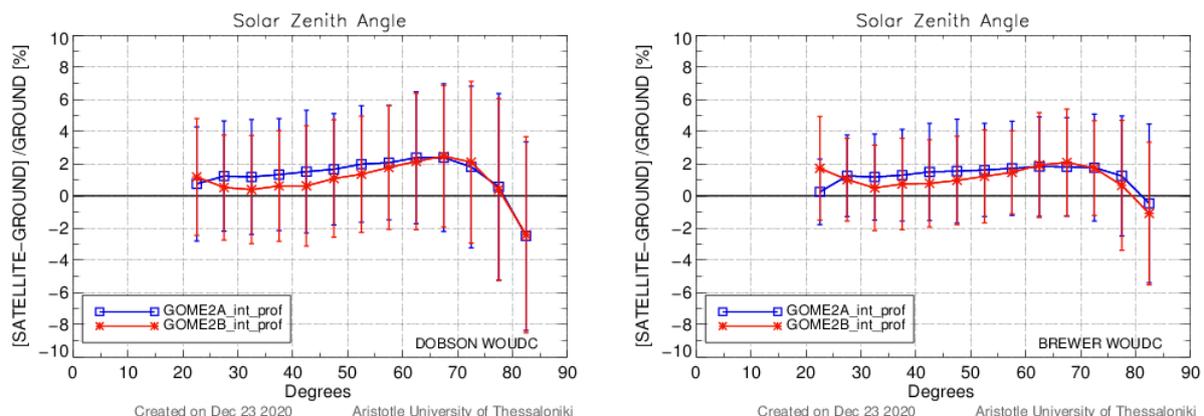
Figure 4.9 : The dependence of the percentage differences on solar zenith angle. Left panel: the Dobson comparisons, right panel: the Brewer comparisons.

4.2.1.3 Dependence on other influence quantities

For the purposes of this report, the effect of many influence quantities, such as cloud parameters, surface albedo etc., on the validation results was studied and no unexpected dependences were found.

An interesting feature that was seen during this validation exercise, is the dependence of the percentage differences on the integrated vertical profile error, seen in The dependencies of the integrated column on the integrated vertical profile error and the index in scan seen in these

figures is a mixture of numerous ozone profile parameters and is currently under investigation by the algorithm team., left panel (*Note: Numbers on the upper part of the figures appear only if the population of the co-locations in each averaging bin is below 5% of the total number of available co-locations.*). The majority of the co-locations correspond to measurements with error values spanning 0.5 - 3.5 DU. It is noticed that for error values below 3.5 DU there is an increase of the percentage differences by ~ 3% and after that point a decrease by up to 3-4% follows for co-locations with errors up to 7 DU. This is a common feature for the two sensors as well as for the co-locations to Dobson instruments (not shown here).
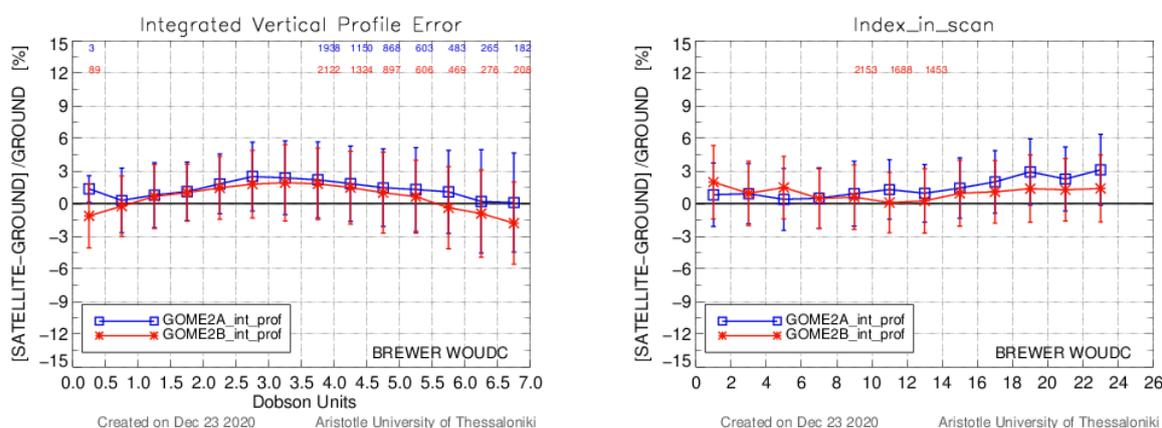
Figure 4.10: The dependence of the percentage differences on the integrated vertical profile error (left panel) and the Index In Scan of the measurements (right panel) with respect to Brewer TOC measurements.

Moreover, in the right panel of The dependencies of the integrated column on the integrated vertical profile error and the index in scan seen in these figures is a mixture of numerous ozone profile parameters and is currently under investigation by the algorithm team., a less pronounced dependence of the percentage differences on the index in scan of the measurements, is shown. Specifically, while GOME-2B does not depend on the index in scan, being almost stable at the 1% level, GOME-2A for index values above 16 increases by 1.5%. As above, the same feature results from the Dobson co-locations (not shown here).

The dependencies of the integrated column on the integrated vertical profile error and the index in scan seen in these figures is a mixture of numerous ozone profile parameters and is currently under investigation by the algorithm team.

4.3 *Conclusions from the GOME-2A and GOME-2B reprocessed integrated ozone profile validation*

The GOME-2A and GOME-2B reprocessed integrated ozone vertical profiles were validated using ground-based daily total ozone measurements from Dobson and Brewer instruments, downloaded from WOUDC. The products under validation were also compared to the temporally and spatially co-located operational total ozone products from GOME-2A and GOME-2B (retrieval algorithm GDP4.8), to further assess their consistency.

The validation results can be summarized to the following points:

- The comparisons of the two sensors had to be filtered for latitude. Their co-locations with Dobson ground-based measurements with latitude greater than 85° S had a mean percentage difference of ~ +20 %. This indicates that there is an issue with the products' retrieval algorithm in the Southern high latitudes that should be studied and resolved.

- Likewise, the comparisons with SZA > 83° had to be excluded, because their ~ -10 % mean bias introduced a lot of noise in the measurements and their statistics.

- The GOME-2A reprocessed integrated ozone profile time-series had to be studied separately before and after mid-2013, when the change in its swath and a discontinuity in the degradation correction occurred and affected the retrieval algorithm.

- The statistical analysis (mean bias in % ± mean standard deviation in %) of the GOME-2A and GOME-2B comparisons to co-located (in space and in time) Dobson and Brewer ground-based measurements is shown in Table 4.1 , where it can be seen that the GOME-2A integrated ozone profile product agrees very well (difference up to ± 0.3%) with the ground-based data before mid-2013. The bias increases to +2 % after that point. Overall, GOME-2B has a mean bias of ~+1.5 to +2% and reports slightly lower integrated ozone vertical profile values (by up to 0.5%) compared to GOME-2A (after mid-2013).

- To further support this conclusion, the comparison of the GOME-2A and GOME-2B integrated ozone profile products to the respective operational total ozone products processed with the GDP4.8 algorithm, is seen in Figure 4.6 and Figure 4.7. The underestimation of GOME-2A integrated profile w.r.t. GDP4.8 by 2 % before mid-2013 in both hemispheres, is obvious. GOME-2B agrees well with the GDP4.8 TOC product. Their differences as they result from the Brewer comparisons are up to 1-2%.

- The peak-to-peak seasonality of the GOME-2A and GOME-2B co-locations to ground-based measurements is 1.3 and 2.3%, respectively.

- The latitudinal analysis of the comparisons showed that both sensors have increased bias with respect to ground-based measurements in the tropics and mid-latitudes, up to 2.5 – 3%. GOME-2B reports lower values of integrated ozone profile compared to GOME-2A by ~0.5%, mainly in the tropics and the middle latitudes of both hemispheres. The consistency between the two sensors is better towards the poles.

- The dependency of the comparisons on SZA showed very similar features for the two sensors.

- Other influence parameters and their effect on the comparisons were also studied, but no alarming dependencies were found.

In conclusion, the validation of the GOME-2A and GOME-2B reprocessed integrated ozone profiles with respect to Dobson and Brewer ground-based total ozone measurements, shows that they are products of very good quality. The step function in the GOME-2A time-series during mid-2013 does not make it suitable for long term studies, but the overall the good agreement of both sensors to ground-based measurements, to the respective operational products, and the luck of severe dependencies on various influence quantities, is an additional proof of the good quality of the reprocessed GOME-2A and GOME-2B vertical ozone profiles products.

5. General conclusions

The GOME-2A and GOME-2B reprocessed vertical ozone profile products were validated against data from measurements with ozonesonde, microwave and lidar. Additionally, Dobson and Brewer measurements were used to validate the quality of the integrated ozone profile product. Both products are also compared with the current operational ozone profile products, derived from GOME-2A and GOME-2B.

It is shown that the optimal goal (10% accuracy) stated in the GOME-2 ozone profile ATBD (Tuinder, 2019) is met in both lower and upper stratosphere for nearly all belts under consideration for both sensors.

The median sensitivity plots confirm in a different way that the applied algorithms on the reprocessed datasets result in a more consistent retrieval behaviour for both sensors (median sensitivity plots between both sensors are comparable now).

The validation results for the GOME-2A/2B integrated ozone profile confirm that this product is of very good quality. It is in excellent agreement with the co-located ground-based measurements.

6. References

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

Balis, D., Kroon, M., Koukouli, M. E., et al.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, J. Geophys. Res., 112, D24S46, doi:10.1029/2007JD008796, 2007a

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

Boyd, I. S., A. D. Parrish, L. Froidevaux, T. von Clarmann, E. Kyrölä, J. M. Russell III, and J. M. Zawodny (2007), Ground-based microwave ozone radiometer measurements compared with Aura-MLS v2.2 and other instruments at two Network for Detection of Atmospheric Composition Change sites, J. Geophys. Res., 112, D24S33, doi:10.1029/2007JD008720.

Delcloo A., and L. Kins (2009, 2012): O3M SAF Validation report on GOME-2 near real-time and offline ozone profiles.

Delcloo A. and K. Kreher (2013): O3M SAF Validation report on GOME-2 near real-time and offline ozone profiles.

Delcloo A. (2020): AC SAF Validation report on GOME-2 near real-time and offline global tropospheric ozone

Brewer, A. and J. Milford, The Oxford Kew ozone sonde, Proc. Roy. Soc. London, Ser. A, 256, 470, 1960.

Kobayashi, J., and Y. Toyama, On various methods of measuring the vertical distribution of atmospheric ozone (III) - Carbon iodine type chemical ozone sonde-. Pap. Met. Geophys., 17, 113-126, 1966

Komhyr, W.D., Electrochemical concentration cells for gas analysis, Ann. Geoph., 25, 203-210, 1969.

Komhyr, W.D., Development of an ECC-Ozonesonde, NOAA Techn. Rep. ERL 200-APCL 18ARL-149, 1971.

Koukouli, M. E., Zyrichidou, I., Balis, D. S., Valks, P., Hao, N., and Valks, P.: GOME-2/MetopA & GOME-2/MetopB GDP 4.8 total ozone data validation, Validation Report SAF/O3M/AUTH/VRR/O3, Issue 1/0, available at: https://acsaf.org/docs/vr/Validation_Report_NTO_OTO_DR_O3_GDP48_Dec_2015.pdf (last access: 22 April 2021), 2015

Loyola, D., Koukouli, M., Valks, P., Balis, D., Hao, N., Van Roozendael, M., Spurr, R., Zimmer, W., Kiemle, S., Lerot, C., and Lambert, J-C.: The GOME-2 Total Column Ozone Product: Retrieval Algorithm and Ground-Based Validation, J. Geophys. Res., 116, D07302, doi:10.1029/2010JD014675, 2011.

Garane, K., Lerot, C., Coldewey-Egbers, M., Verhoelst, T., Koukouli, M. E., Zyrichidou, I., Balis, D. S., Danckaert, T., Goutail, F., Granville, J., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., Pommereau, J.-P., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 1: Ground-based validation of total ozone column data products, Atmos. Meas. Tech., 11, 1385–1402, https://doi.org/10.5194/amt-11-1385-2018, 2018.

Garane, K., Koukouli, M.-E., Verhoelst, T., Lerot, C., Heue, K.-P., Fioletov, V., Balis, D., Bais, A., Bazureau, A., Dehn, A., Goutail, F., Granville, J., Griffin, D., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., McLinden, C., Pazmino, A., Pommereau, J.-P., Redondas, A., Romahn, F., Valks, P., Van Roozendael, M., Xu, J., Zehner, C., Zerefos, C., and Zimmer, W.: TROPOMI/S5P total ozone column data: global ground-based validation and consistency with other satellite missions, Atmos. Meas. Tech., 12, 5263–5287, https://doi.org/10.5194/amt-12-5263-2019, 2019.

Garane, K., Koukouli, M. E., D. Balis, K.P. Heue, D. Loyola: GOME-2/MetopC GDP 4.9 total ozone data validation, Validation Report SAF/AC/AUTH/VR/O3, Issue 1/2020,

---

https://acsaf.org/docs/vr/Validation_Report_NTO_OTO_O3_May_2020.pdf, (last access: 22 April 2021), 2020

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, https://doi.org/10.5194/amt-7-2937-2014, 2014

Hovila, D., S. Hassinen, P. Valks, J., S. Kiemle, O. Tuinder, H. Joench-Soerensen , Product Requirements Document, Issue 1.5, SAF/AC/FMI/RQ/PRD/001, Issue 1.5, June 2019

Keppens, A., Lambert, J.-C., Granville, J., Miles, G., Siddans, R., van Peet, J. C. A., van der A, R. J., Hubert, D., Verhoelst, T., Delcloo, A., Godin-Beekmann, S., Kivi, R., Stübi, R., and Zehner, C.: Round-robin evaluation of nadir ozone profile retrievals: methodology and application to MetOp-A GOME-2, Atmos. Meas. Tech., 8, 2093–2120, https://doi.org/10.5194/amt-8-2093-2015, 2015.

Koukouli, M. E., Balis, D. S., Loyola, D., Valks, P., Zimmer, W., Hao, N., Lambert, J.-C., Van Roozendael, M., Lerot, C., and Spurr, R. J. D.: Geophysical validation and long-term consistency between GOME-2/MetOp-A total ozone column and measurements from the sensors GOME/ERS-2, SCIAMACHY/ENVISAT and OMI/Aura, Atmos. Meas. Tech., 5, 2169-2181, https://doi.org/10.5194/amt-5-2169-2012, 2012.

Koukouli, M. E., Lerot, C., Granville, J., Goutail, F., Lambert, J.-C., Pommereau, J.-P., Balis, D., Zyrichidou, I., Van Roozendael, M., Coldewey-Egbers, M., Loyola, D., Labow, G., Frith, S., Spurr, R., and Zehner, C.: Evaluating a new homogeneous total ozone climate data record from GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, J. Geophys. Res., 120, 12296–12312, https://doi.org/10.1002/2015JD023699, 2015

Lee, G. Mégie, H. Nakane and R. Neuber, Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change, J. Environ. Monit., 6, 721 – 733, doi:10.1039/B404256E, 2004.

Leblanc, T., and I. S. McDermid, Stratospheric Ozone Climatology From Lidar Measurements at Table Mountain (34.4°N, 117.7°W) and Mauna Loa (19.5°N, 155.6°W), J. Geophysical Research, 105, 14,613-14,623, 2000.

Lobsiger E., K.F. Künzi and H.U. Dütsch, Comparison of stratospheric ozone profiles retrieved from microwave-radiometer and Dobson-spectrometer data, J. Atm. and Terr. Phys., 46, 799-806, 1984.

Parrish, A., R.L. de Zafra, P.M. Solomon, and J.W. Barrett, A ground-based technique for millimeter wave spectroscopic observations of stratospheric trace constituents, Radio Sci., 23, 106-118, 1988.

Rodgers C.D., Characterization and Error Analysis of Profiles Retrieved from Remote Sounding Measurements, J. Geophys. Res., 95, 5587-5595, 1990.

Sterling, C. W., Johnson, B. J., Oltmans, S. J., Smit, H. G. J., Jordan, A. F., Cullis, P. D., Hall, E. G., Thompson, A. M., and Witte, J. C.: Homogenizing and estimating the uncertainty in

NOAA's long-term vertical ozone profile records measured with the electrochemical concentration cell ozonesonde, Atmos. Meas. Tech., 11, 3661–3687, https://doi.org/10.5194/amt-11-3661-2018, 2018.

Steinbrecht W., et al. (2006), Long-term evolution of upper stratospheric ozone at selected stations of the Network for the Detection of Stratospheric Change (NDSC), J. Geophys. Res., 111, D10308, doi:10.1029/2005J

Thompson, A. M., Witte, J. C., Sterling, C., Jordan, A., Johnson, B. J., Oltmans, S. J., … Thiongo, K. (2017). First reprocessing of Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone profiles (1998–2016): 2. Comparisons with satellites and ground-based instruments. Journal of Geophysical Research: Atmospheres, 122. https://doi.org/10.1002/2017JD027406.

Tuinder, O., R. van Oss, J. de Haanlaf,, A. Delcloo, ATBD of the Vertical Ozone Profile algorithm, ACSAF/KNMI/ATBD/001, Issue 2.0.2, June 2019

Tuinder, O., Product User Manual (PUM) of the Vertical Ozone Profile Products, ACSAF/KNMI/PUM/001, Issue 2.3.0, February 2021

APPENDIX I

Table A. 1: List of Dobson ground-based stations used for the comparisons

STATION ID	NAME	COUNTRY	LONGITUDE (degrees)	LATITUDE (degrees)	Last day of available measurement
2	Tamanrasset	Algeria	5.51	22.78	31-DEC-2019
10	New Delhi	India	77.17	28.63	30-APR-2019
14	Tateno	Japan	140.13	36.05	27-DEC-2019
19	Bismarck	USA	-100.75	46.76	31-JUL-2019
27	Brisbane	Australia	153.08	-27.42	31-OCT-2019
29	Macquarie island	Australia	158.93	-54.49	31-OCT-2019
31	Mauna Loa	USA	-155.57	19.54	31-JUL-2019
43	Lerwick	UK	-1.18	60.13	25-NOV-2019
57	Halley Bay	Antarctica	-26.18	-75.62	30-MAR-2019
67	Boulder	USA	-105.26	39.99	28-JUL-2019
68	Belsk	Poland	20.79	51.84	31-DEC-2019
82	Lisbon	Portugal	-9.13	38.76	17-DEC-2019
84	Darwin	Australia	130.88	-12.42	31-OCT-2019
91	Buenos-aires	Argentina	-58.48	-34.59	31-MAR-2019
96	Hradec Kralove	Czech_Republic	15.83	50.18	11-DEC-2019
99	Hohenpeissenberg	Germany	11.01	47.80	30-DEC-2019
101	Syowa	Antarctica	39.58	-69.00	31-DEC-2019
105	Fairbanks	USA	-147.87	64.82	31-JUL-2019
107	Wallops island	USA	-75.46	37.94	31-JUL-2019
111	Amundsen-Scott	Antarctica	-24.80	-89.99	25-FEB-2019
152	Cairo	Egypt	31.28	30.08	31-DEC-2019
199	Barrow	USA	-156.61	71.32	31-JUL-2019
208	Shiangher	China	116.96	39.75	31-DEC-2019
216	Bangkok	Thailand	100.62	13.67	31-DEC-2019
219	Natal	Brazil	-35.20	-6.00	30-DEC-2019
226	Bucharest	Romania	26.13	44.48	26-NOV-2019
245	Aswan	Egypt	32.783	23.96	31-DEC-2019
253	Melbourne	Australia	144.83	-37.66	31-OCT-2019
268	Arrival Heights	Antarctica	166.66	-77.83	24-MAR-2019
284	Vindeln	Sweden	19.77	64.23	27-SEP-2019
293	Athens	Greece	23.73	37.98	30-SEP-2019
341	Hanford	USA	-119.63	36.32	31-JUL-2019
342	Comodoro Rivadavia	Argentina	-67.50	-45.78	14-FEB-2019
409	Hurghada	EGU	33.75	27.42	31-DEC-2019
410	Amberd	ARM	44.25	40.38	30-DEC-2019

Table A. 2: List of Brewer ground-based stations used for the comparisons.

STATION ID	NAME	COUNTRY	LONGITUTE (degrees)	LATITUDE (degrees)	Last day of available measurement
53	Uccle	Belgium	4.35	50.79	31-DEC-2019
89	Ny Alesund	Norway	11.92	78.92	18-OCT-2019
95	Taipei	Taiwan	121.48	25.02	31-DEC-2019
96	Hradec Kralove	Czech Republic	15.83	50.18	31-DEC-2019
99	Hohenpeissenberg	Germany	11.01	47.80	31-DEC-2019
213	El Arenosillo	Spain	-6.73	37.10	30-NOV-2019
261	Thessaloniki	Greece	22.96	40.63	31-MAY-2019
279	Norkoping	Sweden	16.15	58.58	31-DEC-2019
284	Vindeln	Sweden	19.76	64.23	15-NOV-2019
308	Madrid	Spain	-3.72	40.45	29-DEC-2019
316	Debilt	Netherlands	5.18	52.10	31-DEC-2019
318	Valentia	Ireland	-10.25	51.94	29-DEC-2019
322	Petaling Jaya	Malaysia	101.65	3.10	31-MAR-2019
330	Hanoi	Vietnam	105.80	21.20	23-NOV-2019
331	Poprad-Ganovce	Slovakia	20.32	49.03	31-DEC-2019
346	Murcia	Spain	-1.17	38.00	31-DEC-2019
352	Manchester	GBR	-2.23	53.47	31-DEC-2019
353	Reading	GBR	-0.94	51.44	31-DEC-2019
376	Mrsa_mtrouh	Egypt	27.22	31.33	31-DEC-2019
401	Santa Cruz	Spain	-16.25	28.47	31-DEC-2019
405	La Coruna	Spain	-8.47	43.33	27-DEC-2019
411	Zaragoza	ESP	-0.91	41.63	31-DEC-2019
476	Andoya	NOR	16.01	69.28	11-OCT-2019
479	Aosta	ITA	7.36	45.74	31-DEC-2019

Table A. 3 List of all ozonesonde stations used for the comparisons

Station	Lat	long	nr of profiles	Last day measurement
ALAJUELA	9.98	-84.21	466	14-DEC-18
ALERT	82.5	-62.33	441	05-DEC-18
ASCENSION	-7.98	-14.42	260	19-DEC-18
BRATTS_LAKE	50.2	-104.7	218	28-SEP-11
BROADMEADOWS	-37.69	144.95	567	19-DEC-18
CHURCHIL	58.74	-94.07	203	12-MAR-14
DAVIS	-68.577	77.973	207	27-NOV-13
DEBILT	52.1	5.18	608	27-DEC-18
EDMONTON	53.55	-114.1	264	09-APR-14
EGBERT	44.23	-79.78	211	31-AUG-11
EUREKA	80	-85.56	442	01-OCT-14
FIJI	-18.1	178.4	174	30-DEC-18
GOOSE_BAY	53.3	-60.36	271	17-JAN-13
HILO	19.717	-155.083	517	26-DEC-18

HOHENPEISSENBERG	47.8	11.02	1431	28-DEC-18
IRENE	-25.9	28.22	130	07-NOV-18
JAVA	-7.5	112.6	90	30-OCT-13
KELOWNA	49.67	-119.4	291	09-APR-14
KUALA_LUMPUR	2.73	101.7	252	21-DEC-18
LAUDER	-45.045	169.684	559	23-DEC-18
LA_REUNION	-20.99	55.48	250	22-JAN-18
LERWICK	60.14	-1.19	580	26-DEC-18
MACQUARIE_ISL	-54.5	158.94	542	25-DEC-18
NAHA	26.2	127.683	401	30-JAN-18
NAIROBI	-1.27	36.8	482	19-DEC-18
NATAL	-5.42	-35.38	313	11-DEC-18
NEUMAYER	-70.39	-8.15	822	26-DEC-18
NY-ALESUND	78.93	11.95	881	28-DEC-18
PARAMARIBO	5.81	-55.21	363	31-DEC-18
PAYERNE	46.817	6.95	1621	27-DEC-18
RESOLUTE	74.71	-94.97	226	02-APR-14
SAMOA	-14.23	-170.56	369	19-DEC-18
SAN_CRISTOBAL	-0.92	-89.6	119	07-JAN-16
SAPPORO	43.06	141.3315	462	29-JAN-18
SODANKYLA	67.3666	26.6297	638	20-DEC-18
SOUTH_POLE	-89.99	-24.8	168	30-DEC-18
TATENO-TSUKUBA	36.1	140.1	532	26-DEC-18
TORONTO	43.78	-79.47	207	26-DEC-12
UCCLE	50.8	4.35	1689	28-DEC-18
USHUAIA	-54.85	-68.308	158	26-OCT-16
VALENTIA	51.93	-10.25	290	28-DEC-18
WALLOPS_ISL	37.84	-75.48	476	19-JUL-16
YARMOUTH	43.87	-66.11	286	02-APR-14

Table A. 4: List of all lidar and MWR stations used for the comparisons

STATION	Latitude	Longitude	No. of profiles	Last measurement used here
Lidar:				
HOHENPEISSENBERG, Germany	47.8	11.02	870	28-Dec-2018
OBS. HAUTE PROVENCE, France	43.94	5,71	1062	17-Dec-2018
TABLE MOUNTAIN, Ca., USA	34.4	117.7	462	30-Dec-2018
MAUNA LOA, Hawaii, USA	19.54	155.58	636	28-Dec-2018
LAUDER, New Zealand	-45.04	169.68	294	06-Nov-2018
Microwave:				
NY-ALESUND, Spitzbergen, Norway	78.93	11.95	16670	24-Oct-2018
BERN, Switzerland	46.95	7.45	18920	29-Dec-2018
PAYERNE, Switzerland	46.82	6.95	6492	29-Dec-2018
MAUNALOA, Hawaii, USA	19.54	155.58	1512	28-Dec-2018
LAUDER, New Zealand	-45.04	169.68	470	09-Oct-2016