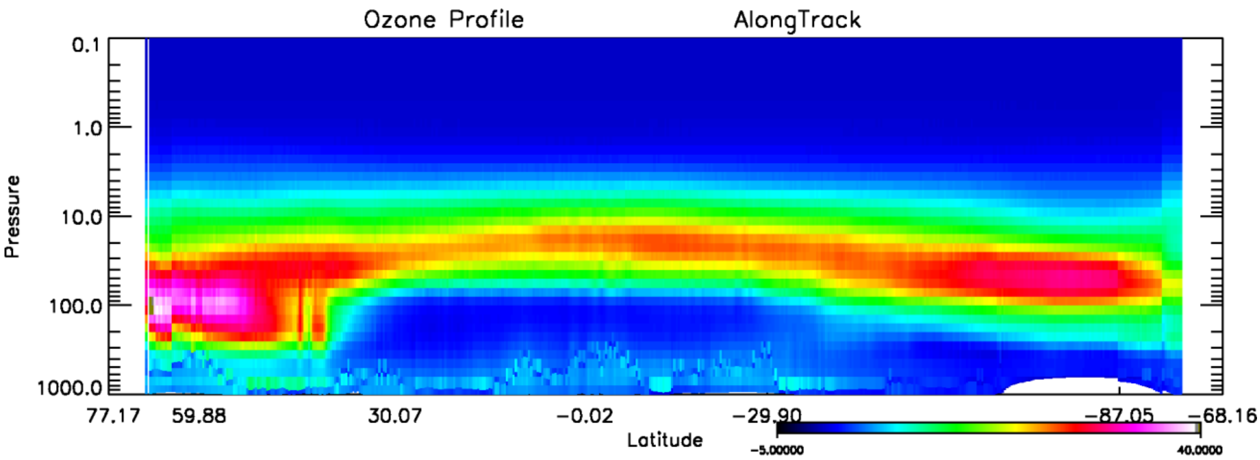
O3M SAF VALIDATION REPORT

Validated products:

Identifier	Name	Acronym
O3M-03	Near-Real-Time Ozone Profile	NOP
O3M-13	Offline Ozone Profile	OOP
O3M-38	Near-Real-Time High Resolution Ozone Profile	NOP/HR
O3M-39	Offline High Resolution Ozone Profile	OOP/HR



Authors:

Name	Institute
Andy Delcloo	Royal Meteorological Institute of Belgium
Karin Kreher	German Weather Service

Reporting period:	December 2012 – April 2013
Validation methods:	Lidars and microwave radiometers (altitude range 15 – 60 km) Balloon soundings (altitude range 0 – 34 km)
Input data versions:	Base Algorithm Version: 5.3
Data processor Versions:	Product Algorithm Version: 1.11 Product Software Version: 1.28

Table of Contents

1. General Introduction ... 3
2. Validation of ozone profiles with ozonesondes ... 3
   2.1 Introduction ... 3
   2.2 Dataset description ... 4
   2.3 Co-location criteria ... 6
   2.4 Ozone sounding pre-processing ... 6
   2.5 Results ... 7
   2.6 Solar Zenith Angle dependency ... 15
   2.7 Seasonal dependency ... 16
   2.8 Information content ... 18
   2.9 Cloud Cover ... 21
   2.10 Retrieval quality indicators ... 21
3. Validation of ozone profiles with lidar and microwave instruments ... 29
   3.1 Instruments ... 29
   3.2 Dataset description ... 30
   3.3 Comparison procedure ... 30
   3.4 Co-location criteria in time and space ... 30
   3.5 Pre-processing of the ozone profiles. ... 31
   3.6 Results ... 33
   3.7 Mean differences between GOME-2/MetopB and co-located lidar and microwave ozone profiles ... 33
   3.8 Ozone variations as seen by GOME-2/MetopB and by lidar and microwave ozone profiles ... 36
   3.9 GOME-2 MetopB coarse resolution ozone profiles. ... 45
   3.10 Comparing GOME-2 on MetopB with GOME-2 on MetopA ... 49
4. Summary and conclusions ... 52
5. Acknowledgement ... 53
6. References ... 53

1. General Introduction

This report contains validation results of the GOME-2/Metop-B ozone profile product retrieved by the Ozone Profile Retrieval Algorithm (OPERA) at KNMI. It covers the time period December 2012 until April 2013. Ozone profiles retrieved from processed level-1b data were used in coarse resolation (CR) and in high resolution (HR).
Since this work was carried out in two different institutions, this document is split up into two separate parts. The first part contains the validation of the retrieved GOME-2 ozone profiles using balloon ozonesondes (chapter 2). This part will mainly validate the retrieved ozone profiles in the troposphere and the lower stratosphere. The second part shows the validation with lidars and microwave radiometers (chapter 3) describing the performance of GOME-2 retrieved ozone profiles primarily in the stratosphere, from 20 to 60 km altitude. The outcome of both validation parts is summarized in the summary and conclusions section at the end of the report.

2. Validation of ozone profiles with ozonesondes

2.1 Introduction

This report presents validation results for the O3M SAF GOME-2 ozone profile product. The validation was carried out using ozone profile measurements with balloon sounding data.
Ozonesondes are lightweight balloon-borne instruments which are able to make ozone measurements from the surface up to about 30 km, with much better vertical resolution than satellite data. In general also the precision and accuracy will be better, at least in the lower stratosphere and the troposphere. Another advantage is that ozone soundings can be performed at any time and at any meteorological condition.

The precision of ozonesondes varies with altitude and depends on the type of sonde used. Table 2.1 below shows indicative precision (in percent) of the Electrochemical Concentration Cell (ECC), Brewer-Mast (B-M) and the Japanese KC79 ozonesondes, at different pressure levels of the sounding (taken from the O3MSAF Science Plan).

Table 2.1: Precision 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

It is shown from Table 2.1 that the 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 around 31 km altitude. For this validation report, only the station of Hohenpeissenberg is using B-M sondes, all the other stations under consideration (Table 2.3) use ECC sondes, while KC-79 sondes are not launched anymore.

2.2 Dataset description

GOME-2 ozone data used in this validation report is from the beginning of December 2012 up to the end of April 2013. GOME-2 ozone data was made available by KNMI at pre-selected sites. These sites correspond to sites where ozone soundings are performed on a regular basis. 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/). However, since this report is focusing on very preliminary results of the GOME-2 retrieved ozone profiles, we don't have as much ozonesonde data available.

Latitude belts from North to South:

1. Polar stations North: green (67N – 90 N)
2. Mid-Latitude stations North: black (30 N – 67 N)
3. Tropical stations: Red (30 N – 30 S)
4. Mid-Latitude stations South: grey (30 S – 70 S)
5. Polar stations South: blue (70 S – 90 S)

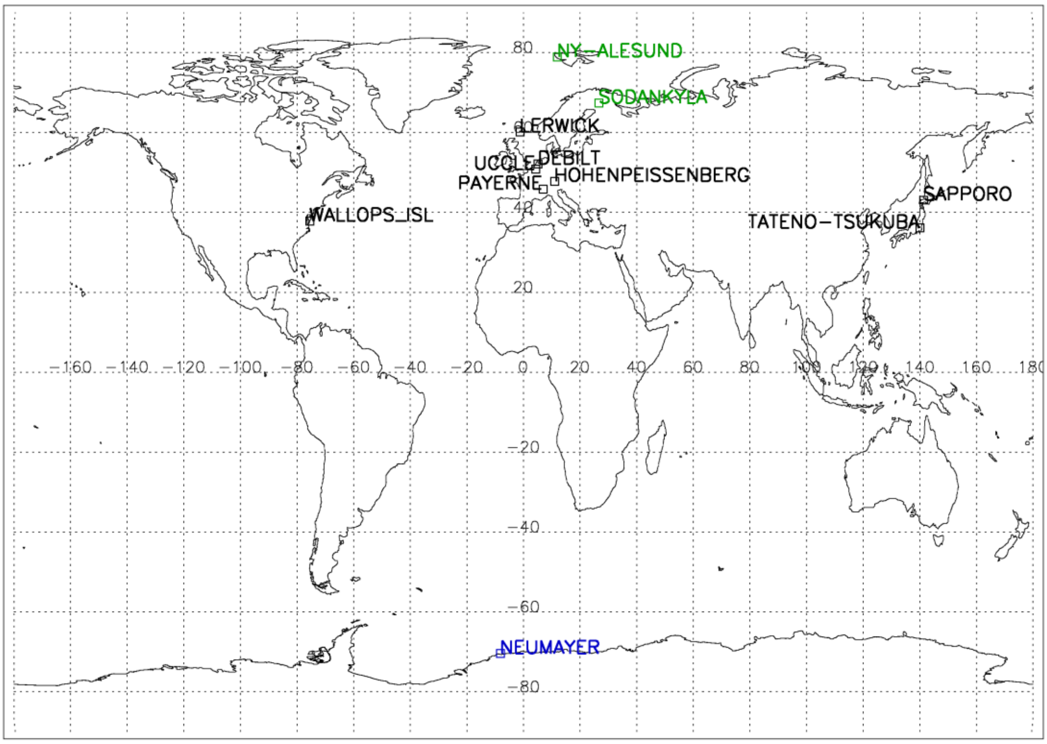


Figure 2.1: Stations used in the validation report

The algorithm versions used for the GOME-2/METOP-B ozone profile retrievals is **5.3**.

Ozonesonde data are generally made available by the organization carrying out observations after a delay in order to leave time for necessary verification and correction of the data quality. Nevertheless, some organizations make their ozone profile data readily available for validation purposes.

The time period we consider here for validation is:

- December 2012 - April 2013

A validation is performed in function of latitude belts. The number of coincidences is summarized in the Table 2.2. The GOME-2 ozone profile data taken into consideration are the ones which have received the quality processing status of “Overall convergence, successful retrieval”. More details about the quality flags can be found in the PUM document (Product User Manual for the Near Real Time and Offline Ozone Profile), pages 22-23. Figure 2.1 and Table 2.3 show an overview of the stations used in this validation report. Please, notify that only 11 stations are considered at this stage, covering three latitude belts.

Table 2.2: Overview of number of coincidences at different latitude belts for the time period December 2012 – April 2013 for coarse resolution and high resolution pixels

Latitude band	Nr of coincidences CR-pixels Metop-B	Nr of coincidences HR-pixels Metop-B	Nr of concidences CR-pixels Metop-A
Polar North (90-67)	453	409	6873
Mid-Latitudes North (30-67)	1823	2336	5185
Polar South (–70 –90)	419	390	446

Table 2.3: Overview of the stations taken into account with the numbers of sondes used in the analysis and the last day, a sonde was available for the intercomparison

STATION	Long	Lat	Nr of sondes	Last day available ozonesonde
DEBILT	52,10	5,18	20	25/04/2013
HOHENPEISSENBERG	47,80	11,02	55	26/04/2013
LERWICK	60,14	-1,19	18	26/04/2013
NEUMAYER	-70,39	-8,15	19	19/04/2013
NY-ALESUND	78,93	11,95	33	24/04/2013
PAYERNE	46,82	6,95	8	31/12/2012
SAPPORO	43,06	141,33	9	27/02/2013
SODANKYLA	67,37	26,63	19	25/04/2013
TATENO-TSUKUBA	36,10	140,10	11	27/02/2013
UCCLE	50,80	4,35	52	24/04/2013
WALLOPS_ISL	37,84	-75,48	16	26/03/2013

2.3 Co-location criteria

The selection criteria, taken into account are two fold:

- The geographic distance between the GOME-2 pixel center and the sounding station location for the coarse resolution (CR) pixels is 300 km, for the high resolution (HR) pixels, this distance is reduced towards 100 km.
- The time difference between the pixel sensing time and the sounding launch time is the second criterion and is fixed at ten hours of time difference. Each sounding that is correlated with a GOME-2 overpass is generally correlated with several GOME-2 pixels if the orbit falls within the 300 km (resp 100 km for the HR pixels) 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 the ozone concentration along the ascent with a much higher vertical resolution than GOME-2. Ozonesondes have a typical vertical resolution of 100m while GOME-2 profiles consist 40 layers between the ground and 0.001hPa. Ozonesondes give the ozone concentration in partial pressure. The integration requires some interpolation, as GOME-2 levels never match exactly ozonesonde layers. This interpolation causes negligible errors given the high vertical resolution of ozonesonde profiles.

For the comparison, ozonesonde profiles are integrated between the GOME-2 pressure levels of the GOME-2 profile being compared. This means 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. This integrated ozonesondes data will be further referred in this report as X_{sonde}.

However, GOME-2 layers are relatively thick and GOME-2 layers boundaries show small variations compared to the layer thickness. So, the same layer always falls 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 for making graphs.

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, to look at the GOME-2 profiles with “the eyes” from the satellite. Equation (1) shows how the kernels have been applied.

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

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

2.5 Results

To calculate the relative difference profiles, we applied the following formulas:

For comparing the GOME-2 profile with the ozone sounding we apply:

$$(X_{GOME\text{-}2} - X_{SONDE})/X_{SONDE} \qquad (2)$$

Notify that X_{sonde}.is integrated from ozonesondes measurements.

The line in the comparison figures (Fig.2.2a), corresponding to these comparisons is colored black.

For comparing the GOME-2 profile with the smoothed ozonesondes, further referred in this report as AVK ozonesondes, by applying the averaging kernels (1), we use the following equation:

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

The line in the comparison figures (Fig.2.2a), corresponding to these comparisons is colored blue.

Figure 2.2a shows relative difference profiles between GOME-2 ozone profiles at the one hand and on the other hand ozonesonde-, and AVK ozonesonde profiles for the latitude belts, listed in Table 2.2 for respectively the coarse resolution pixels and the high resolution pixels. The error bars represent one standard deviation on the mean error.
Please, notify that we don’t have a full year for this validation result. In this subsection, we will first describe the statistics for the considered time period December 2012 –April 2013 and in the next sections we will discuss the seasonal behavior and other possible influences on the quality of the ozone profile product.

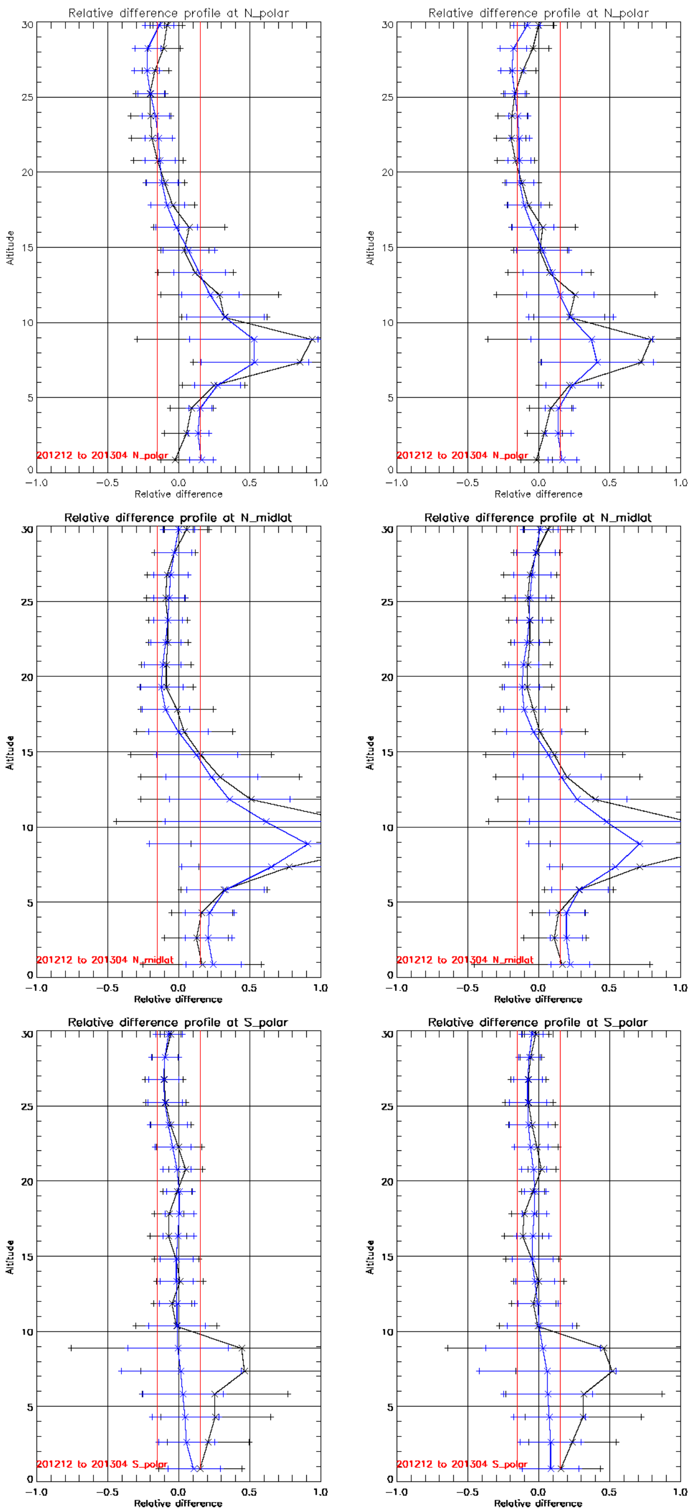


Figure 2.2a: Relative difference profiles (%) between GOME-2 ozone profiles, ozonesondes (black) and smoothed ozonesondes (blue), according to equations 2 and 3 for different latitude belts for the time period December 2012-April 2013, CR (left) and HR (right)

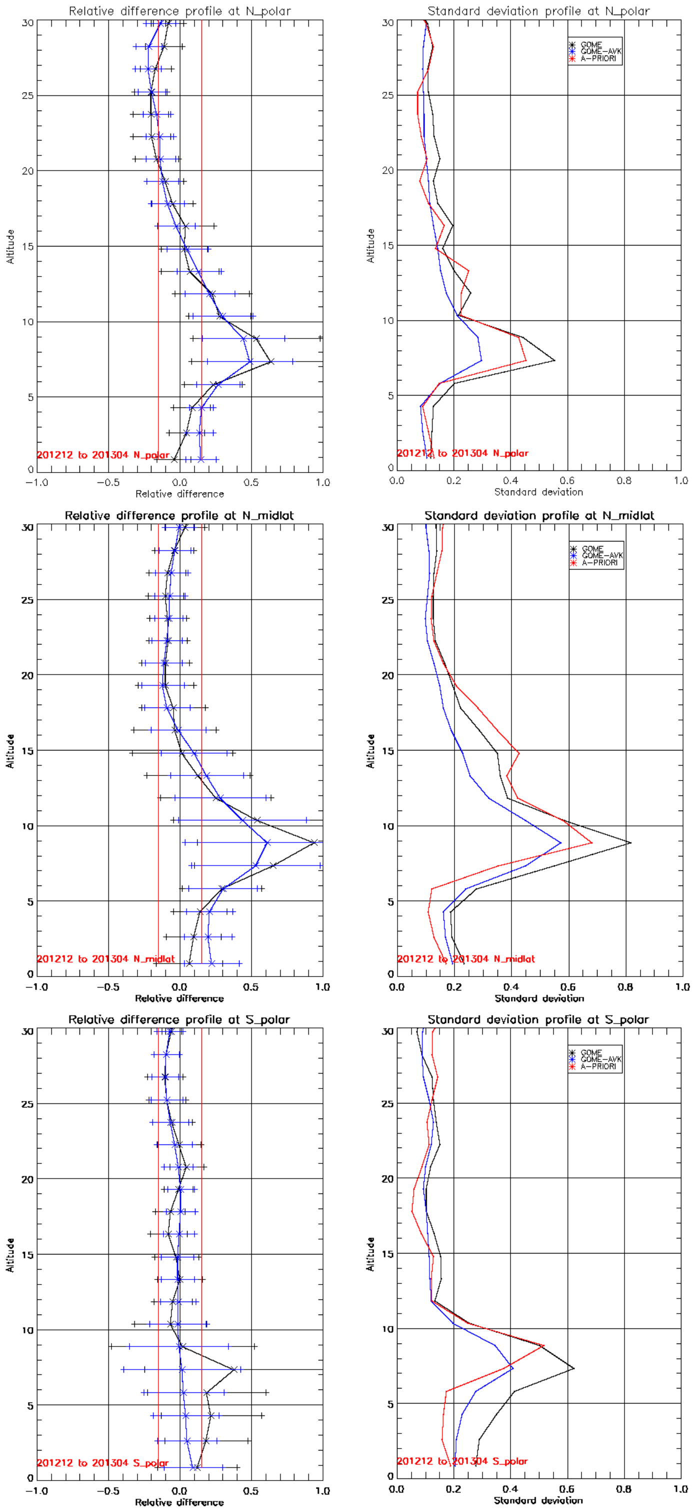


Figure 2.2b: Relative difference profiles (%) between GOME-2 ozone profiles, ozonesondes (black) and smoothed ozonesondes (blue), in function of the normalized mean for different latitude belts for the time period December 2012-April 2013, CR (left) and a plot of the standard deviation as difference between sondes and a-priori (red line), retrieval (black line) and retrieval accounting for avk (blue line) (right)

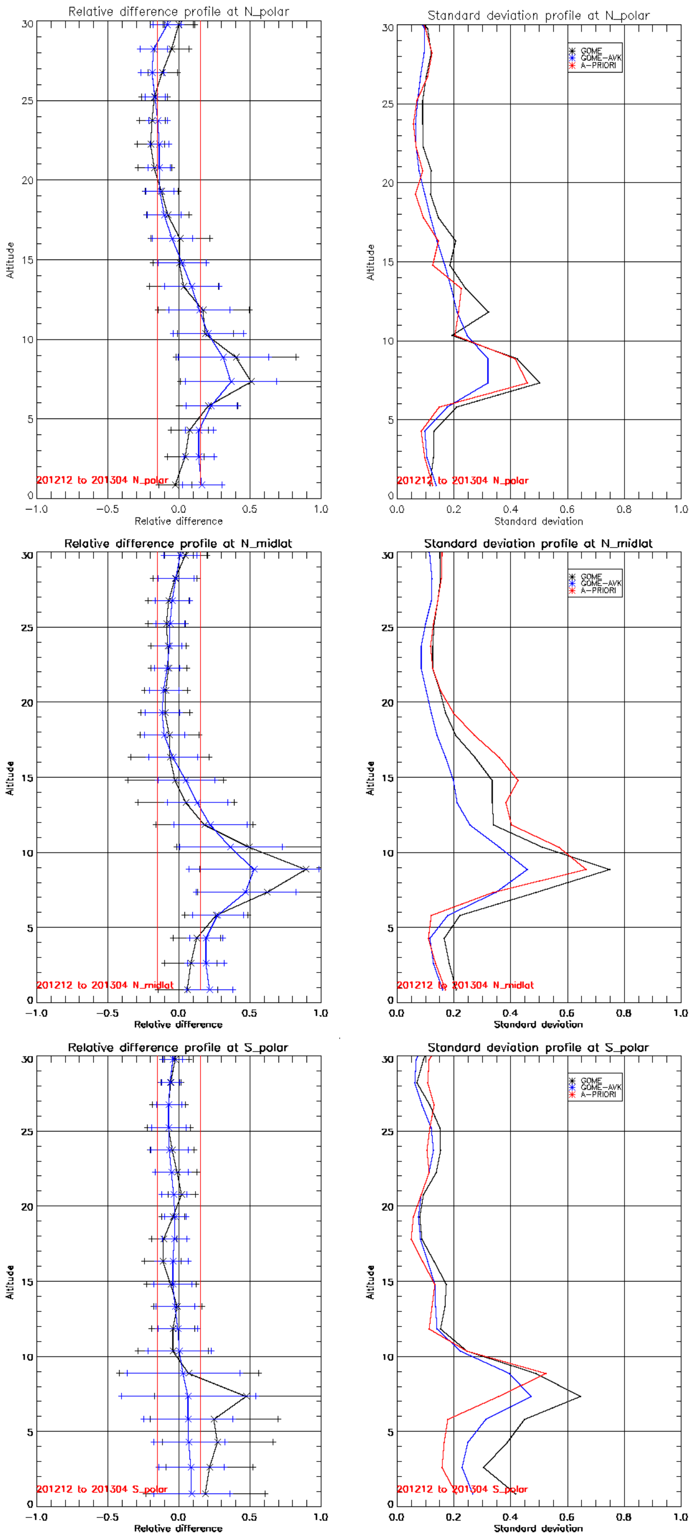


Figure 2.2c: Relative difference profiles (%) between GOME-2 ozone profiles, ozonesondes (black) and smoothed ozonesondes (blue), in function of the normalized mean for different latitude belts for the time period December 2012-April 2013, HR (left) and a plot of the standard deviation as difference between sondes and a-priori (red line), retrieval (black line) and retrieval accounting for avk (blue line) (right)

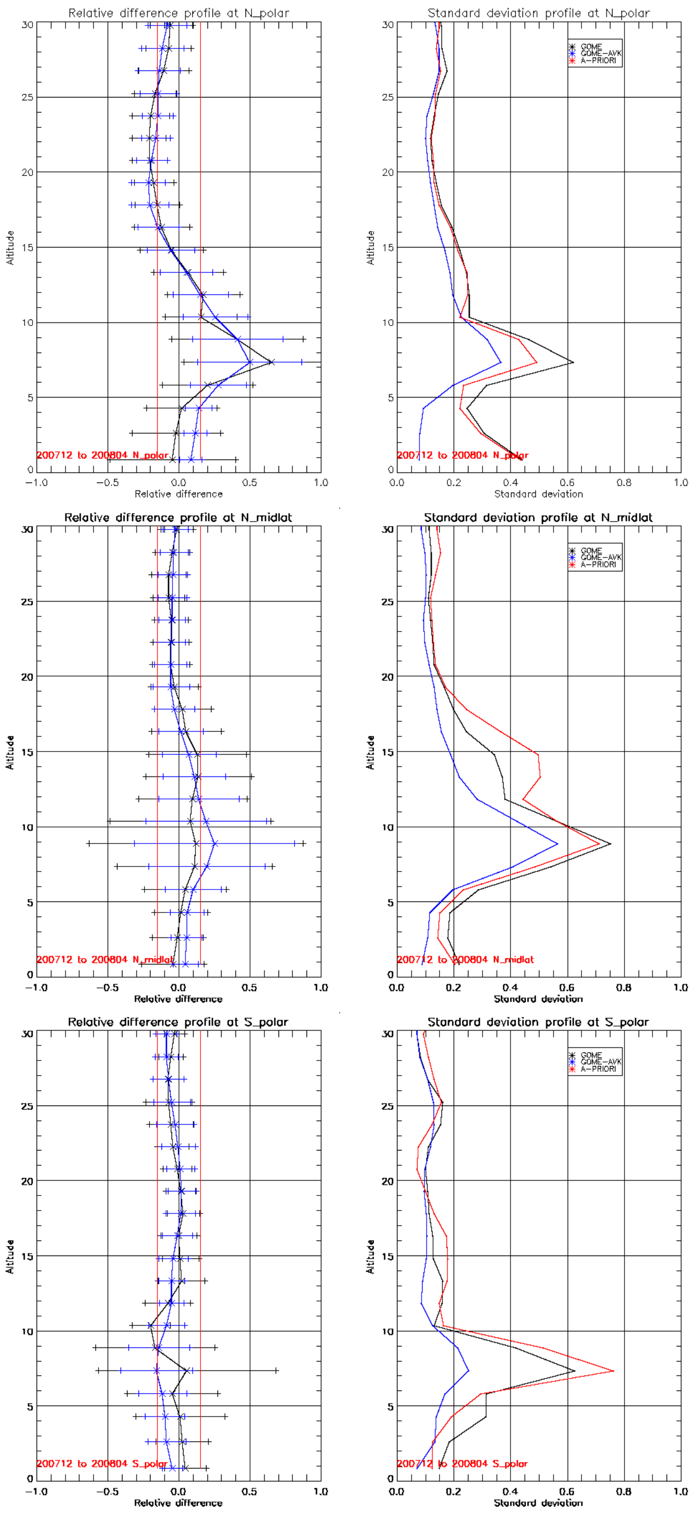


Figure 2.2d: Relative difference profiles (%) between GOME-2 ozone profiles, ozonesondes (black) and smoothed ozonesondes (blue), in function of the normalized mean for different latitude belts for the time period December 2007-April 2008, HR (left) and a plot of the standard deviation as difference between sondes and a-priori (red line), retrieval (black line) and retrieval accounting for avk (blue line) (right) for Metop-A.

For the Northern Polar stations, the difference plots show an underestimation up to –16 % in the stratosphere and an overestimation up to +90 % in the upper troposphere, lower stratosphere zone, further referred as the UTLS-zone in this report. For the lower troposphere, the relative difference is within 15 %. Applying the averaging kernels improves the comparison significantly (40 % in the UTLS-zone). For the high resolution pixels, the relative differences at the Northern Polar stations improved. Relative differences are on average for the stratosphere within the target values when validated against the $X_{AVK\text{-}sonde}$. The statistics in the stratosphere show a lower standard deviation for the HR pixels.

For the Northern Mid-Latitude stations, the GOME-2 ozone profile retrievals are underestimating the ozone profile (-10%) for altitudes between 20 km and 30 km. There is an overestimation present in the UTLS-zone up to +90 % for the coarse resolution pixels and up to +70 % for the high resolution pixels, validated against the ozonesondes, applying the averaging kernels (blue line). For the lower part of the troposphere, the retrieved ozone profile is overestimating the ozone concentrations with 20 %.

For the Southern Polar belt, the ozonesondes at the station Neumayer are used. The relative difference is within 30 % for the troposphere. When applying the averaging kernels, it is shown from figure 2.2a that the relative differences are within 15 %, with, in the stratosphere, an averaged underestimation of about 5 % when using the averaging kernels.

It is shown from these preliminary results that the statistics for the coarse resolution ozone profile product and for the high resolution ozone profile product show similar behaviour, with some better statistics for the high resolution ozone profile product, translated in general in reduced relative differences and standard deviations.

To report the validation results of GOME-2 ozone profile product in a more condensed way, the statistics for the different output levels of GOME-2 are reduced to three layers: Troposphere, UTLS-zone (Upper Troposphere, Lower Stratosphere) and Stratosphere (up to an altitude of 30 km). Since it is impossible to address just one altitude to the tropopause height for the different belts, Table 2.4 gives an overview on how we define the ranges in height for the different belts for troposphere, UTLS-zone and stratosphere:

Table 2.4: Definition of the ranges in km for Troposphere, UTLS and Stratosphere for the different belts.

	Troposphere	UTLS	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 2.5 shows an overview of the obtained results for the time period December 2012 – April 2013:

Table 2.5: Relative Differences (RD) and standard deviation (STDEV) are shown (in percent) on the accuracy of GOME-2 ozone profiles product for the troposphere, UTLS-zone and the stratosphere for three different latitude belts for the time period December 2012 – April 2013, coarse resolution (CR) and the time period December 2012 – April 2013 ,high resolution (HR) validated against $X_{AVK\text{-}sonde}$.

	Troposphere		UTLS		Stratosphere	
December 2012-April 2013 (CR)	RD (%)	STDEV (%)	RD (%)	STDEV (%)	RD (%)	STDEV (%)
Northern Polar Regions	18.03	9.93	40.30	32.65	-9.98	11.77
Northern Mid-Latitudes	32.98	28.64	52.95	64.43	-4.64	14.62
Southern Polar Regions	6.05	22.48	-0.33	27.45	-4.24	10.92
December 2012-April 2013 (HR)	RD (%)	STDEV (%)	RD (%)	STDEV (%)	RD (%)	STDEV (%)
Northern Polar Regions	17.13	11.80	29.20	33.95	-9.81	10.90
Northern Mid-Latitudes	28.94	20.82	40.65	48.68	-4.90	13.64
Southern Polar Regions	7.70	24.53	2.30	31.50	-4.89	10.75

It can be concluded that for the high resolution ozone profile product, the target values are met in the troposphere (30 %) and the stratosphere (15 %), not taking into account the UTLS zone, which shows more elevated relative differences which cannot be appointed to the troposphere or the stratosphere. For the coarse resolution ozone profile product, this target value is almost met.
When calculating the relative difference according to equations 2 and 3 in function of the normalized mean, we obtain a comparable but different plot, showing a more conservative overestimation in the UTLS zone (Fig. 2.2b for CR and Fig. 2.2c for the HR). On the right hand side of figures 2.2b and 2.2c, the standard deviation (%) is shown as a difference between a-priori and sondes (red line), retrieval (black line) and retrieval accounting for avk (blue line) in function of the altitude (km). This shows that especially in the stratosphere, the algorithm is able to retrieve information on the vertical distribution of ozone in those specific layers. For the troposphere, it is shown from these plots that the sensitivity is rather low for the specific time period under consideration.

To make a comparison possible with the Metop-A coarse resolution ozone profile product, figure 2.2d is added. They show the relative difference profiles (%) between GOME-2 ozone profiles, ozonesondes (black) and smoothed ozonesondes (blue), in function of the normalized mean for different latitude belts for the time period December 2007- April 2008 and on the right side of these plots, the standard deviation (%) is shown as a difference between a-priori and sondes (red line), retrieval (black line) and retrieval accounting for avk (blue line) (right). The current Metop-B validation results are similar to those obtained during Metop-A for the time period under consideration.

Figure 2.3 gives an overview of the averaged ozone profiles for the GOME-2 retrieved ozone profiles, ozonesondes and AVK ozonesondes for three latitude belts. It shows that in general the GOME-2 ozone profile overestimates the ozone concentrations in the UTLS zone and underestimates the ozone concentrations in the upper stratosphere for the Northern Polar stations and the Northern Mid-Latitude stations. For the Neumayer stations there is a good match between the ozonesondes and the retrieved ozone profile.

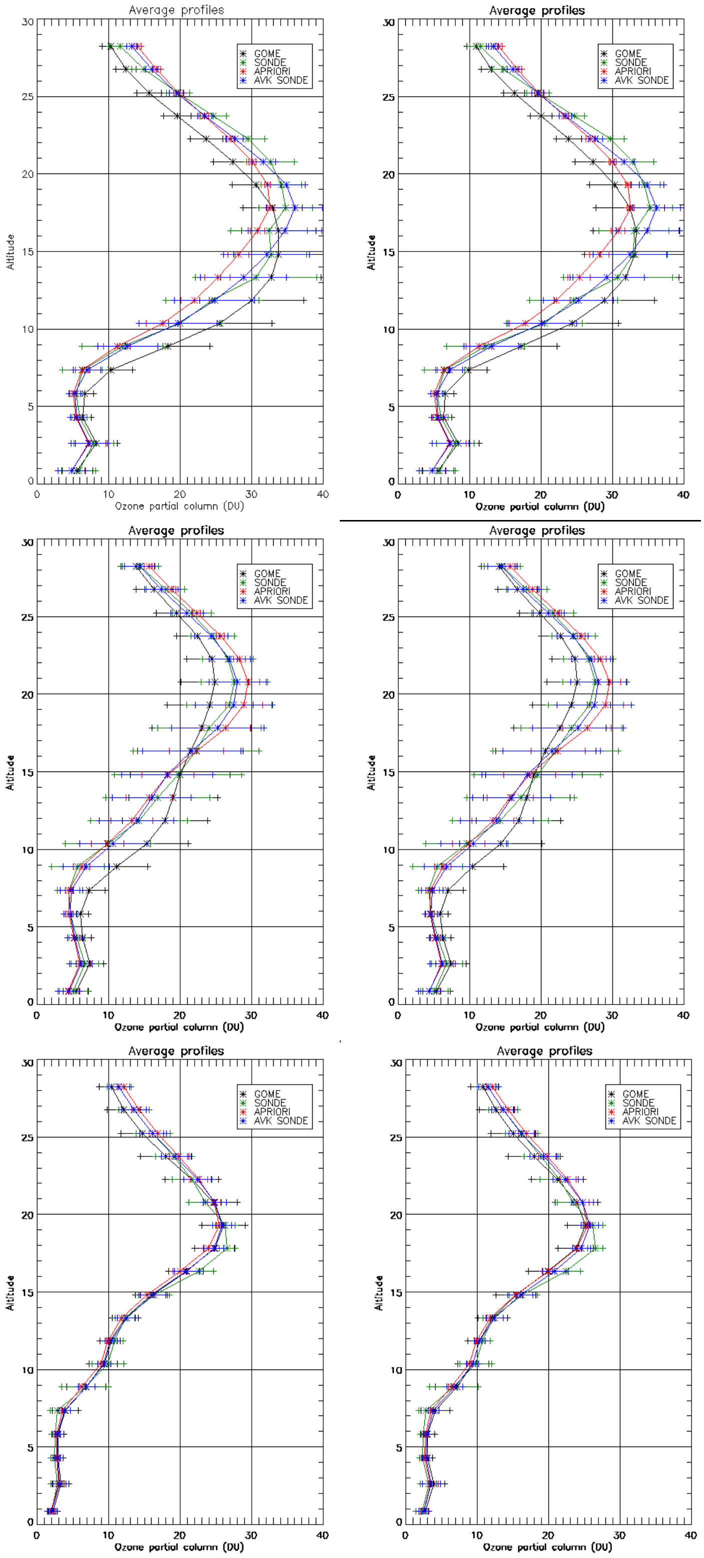


Figure 2.3: Averaged ozone profiles of GOME-2 profile (black), apriori profile (red), smoothed ozonesondes (blue) and ozonesondes (green) for the time period December 2012 – April 2013, CR (left) and for the time period December 2012 –April 2013, HR (right).

2.6 Solar Zenith Angle dependency

From previous studies with GOME-2/METOP-A data (Delcloo and Kins, 2009/2012), it is known that the GOME-2 ozone retrieval shows a seasonal dependency and is also influenced by the Solar Zenith Angle (SZA). The next figure shows the dependency on SZA for the Mid-Latitude stations (Fig. 2.4a) and the Northern Polar stations (Fig. 2.4b) for METOP-B.

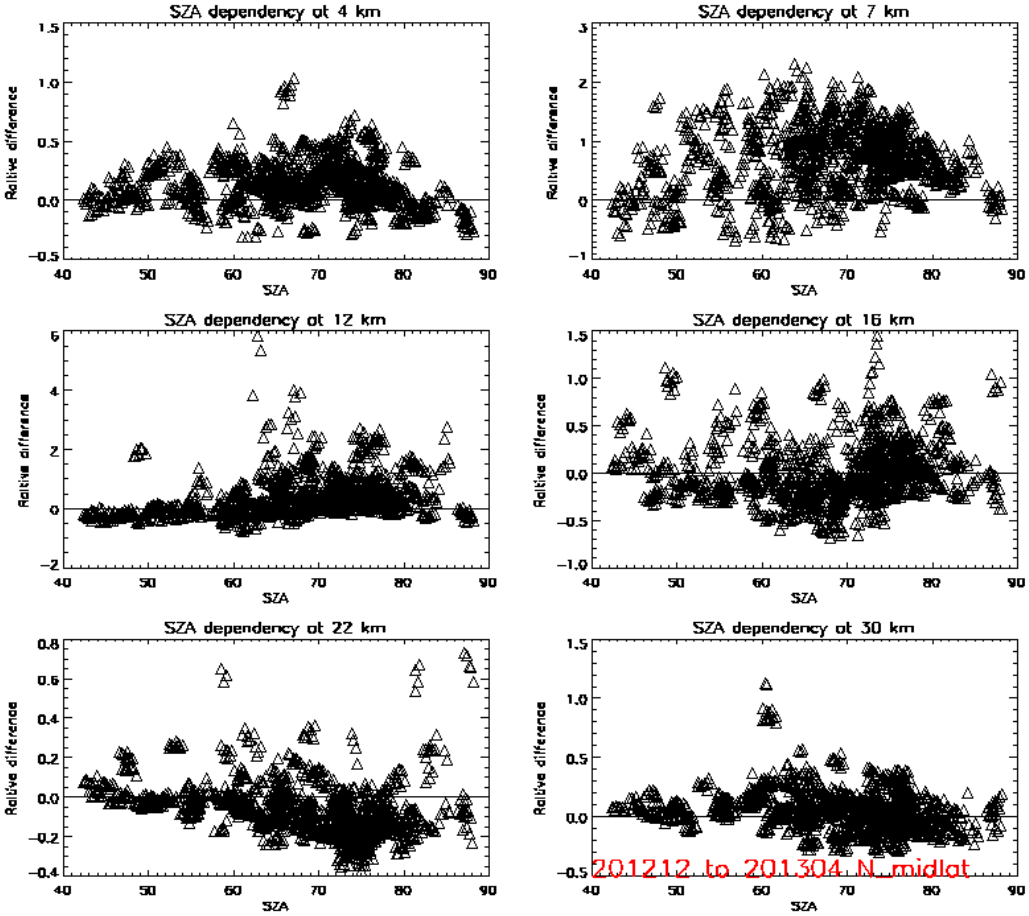


Figure 2.4a: Solar Zenith Angle dependency for six altitude levels for the HR-pixels, Northern Mid-Latitude stations, time period December 2012 –April 2013

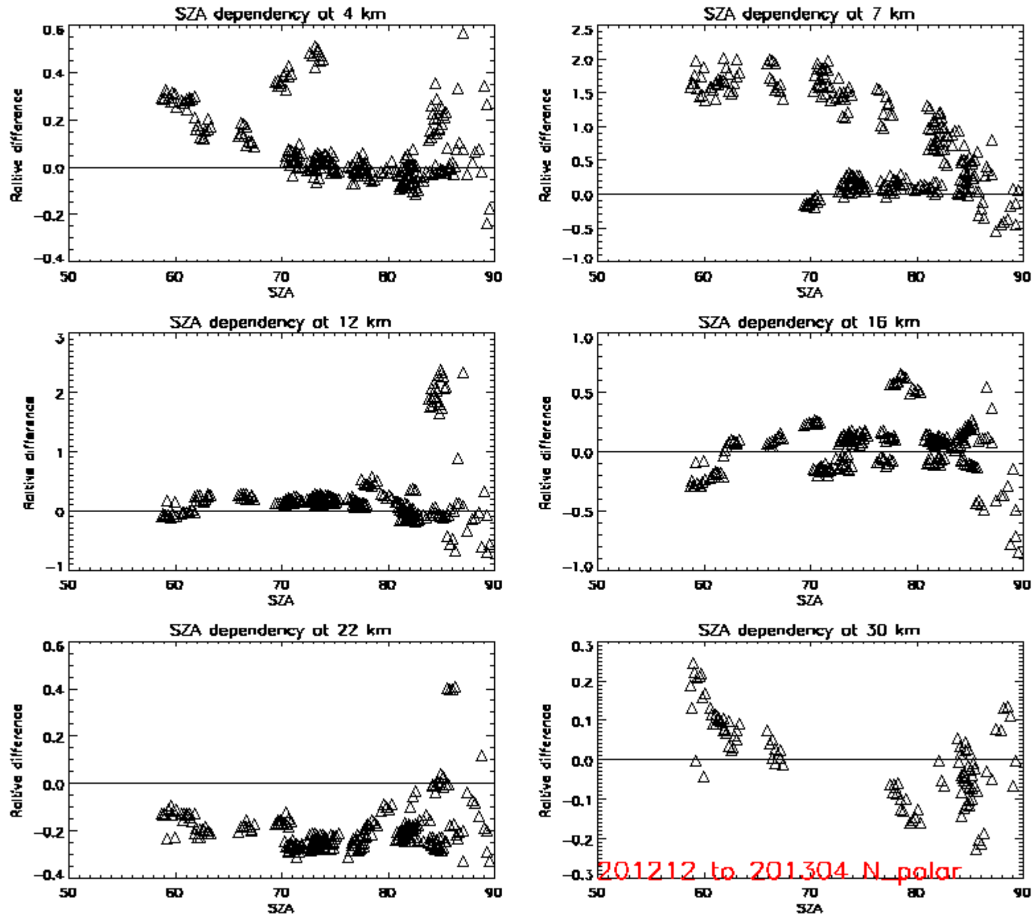


Figure 2.4b: Solar Zenith Angle dependency for six altitude levels for the HR-pixels, Northern Polar stations, time period December 2012 –April 2013

Both figures show that elevated solar zenith angles (SZA > 60°) are correlated with the poor comparisons for the Northern Mid-Latitude stations and Northern Polar stations at higher altitudes, especially around the ozone maximum (22 km). This statement should be taken with care, since for the Northern Polar stations there is only data available with solar zenith angles higher than 60°.

2.7 Seasonal dependency

Figure 2.5a shows relative differences between GOME-2 and ozonesonde data for 6 altitude levels at the Northern Mid-Latitude stations. This graph shows if there is any seasonal dependency present in the GOME-2 dataset. At the Northern Mid-Latitude and Northern Polar stations, there is during winter months and early spring an overestimation present in the UTLS-zone, but no clear seasonal behaviour can be observed yet for this altitude level because of the absence of a full year of retrieved GOME-2 ozone profile data. From Metop-A, it is known that there is a seasonal cycle present in the data (Delcloo and Kins, 2009/2012) and the results shown here, are comparable with the METOP-A GOME-2 ozone profile product.

For the stratosphere, a seasonal behaviour can already be identified at the 22 km level. Statistics improve significantly for March and April. This improvement can also be observed for the Northern Polar stations at the 30 km altitude level (Fig. 2.5b).

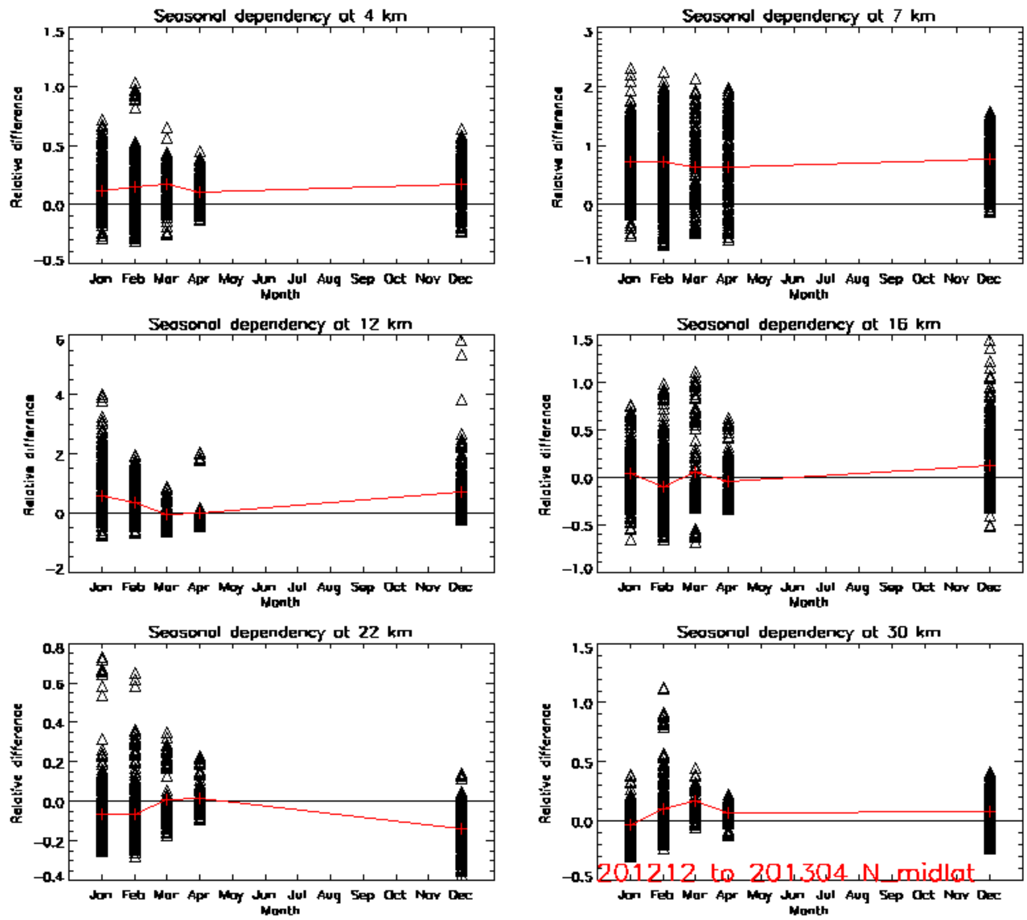


Figure 2.5a: Seasonal dependency at different altitude levels for the Northern Mid-Latitude stations, time period December2012-April 2013(HR) (relative difference between GOME-2 and ozonesondes)

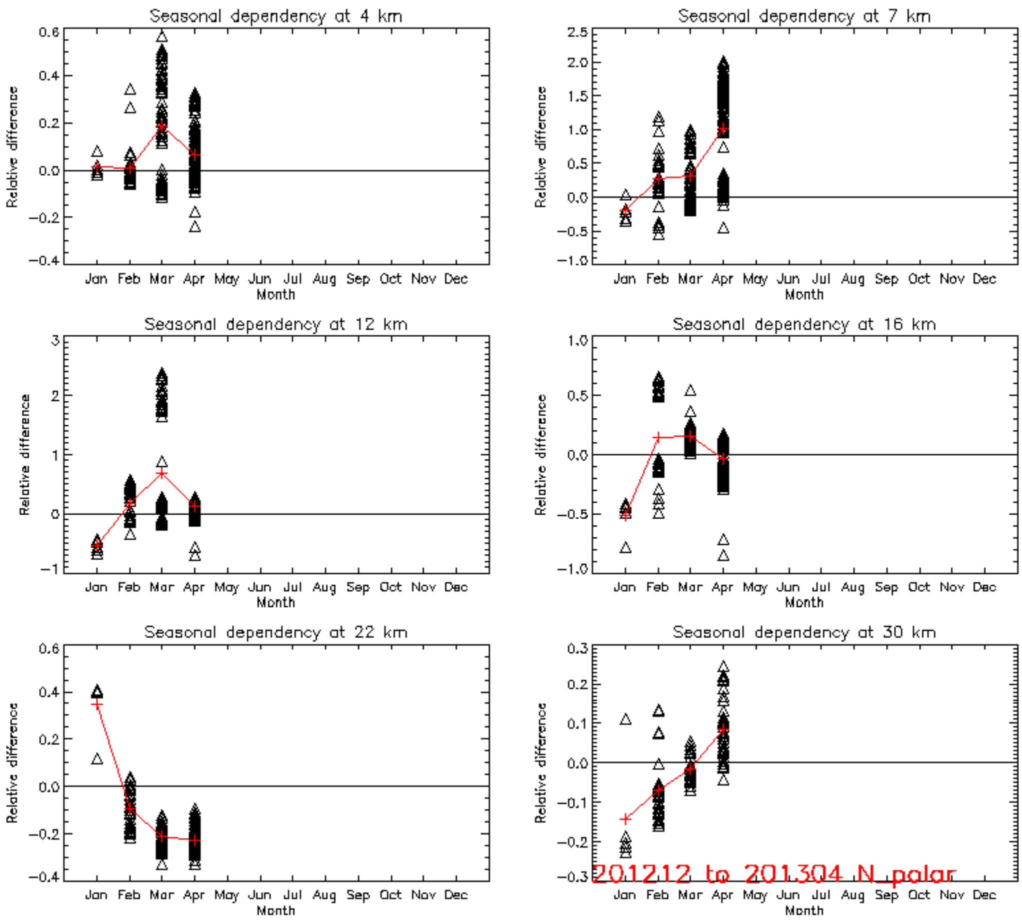


Figure 2.5b: Seasonal dependency at different altitude levels for the Northern Polar stations, time period January 2013-April 2013(HR) (relative difference between GOME-2 and ozonesondes)

2.8 Information content

Scatter plots showing the retrieved ozone partial columns as a function of the reference partial column measured by the ozonesonde give a measure of the amount of information actually present in the retrieved layer.

Scatter plots in figure 2.6a show the retrieved ozone partial columns as a function of the reference partial column measured by the ozonesondes. The slope of the regression line can be seen as a measure for the amount of information actually present in the retrieved layer. This is done for 6 altitude layers for the Northern Mid-Latitude stations. To show the influence of applying the averaging kernels it is shown from figure 2.6b that the slope values are improved (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 that GOME-2 loses sensitivity in the lower troposphere and around the UTLS-zone (Fig. 2.6a).

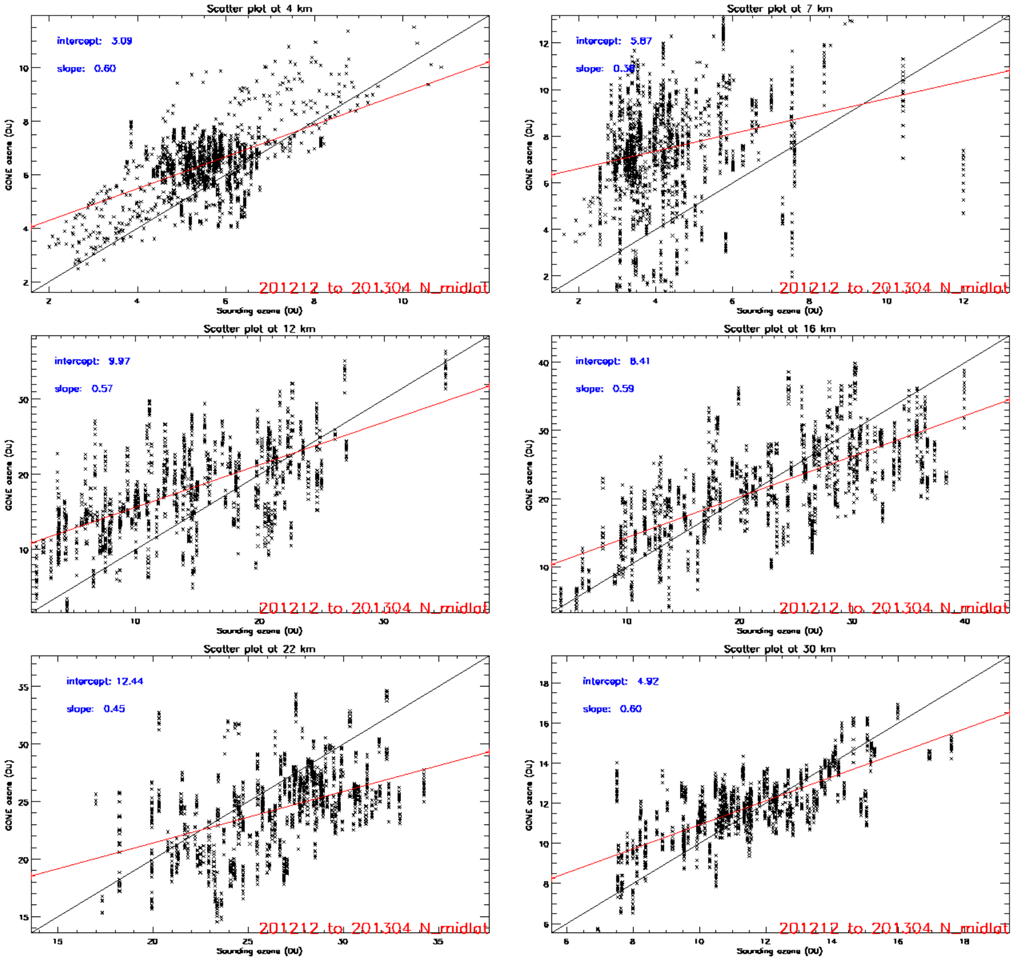


Figure 2.6a: Scatter plot at 6 different altitude levels for the stations at Northern Mid-Latitudes for the HR pixels (December 2012-April 2013).

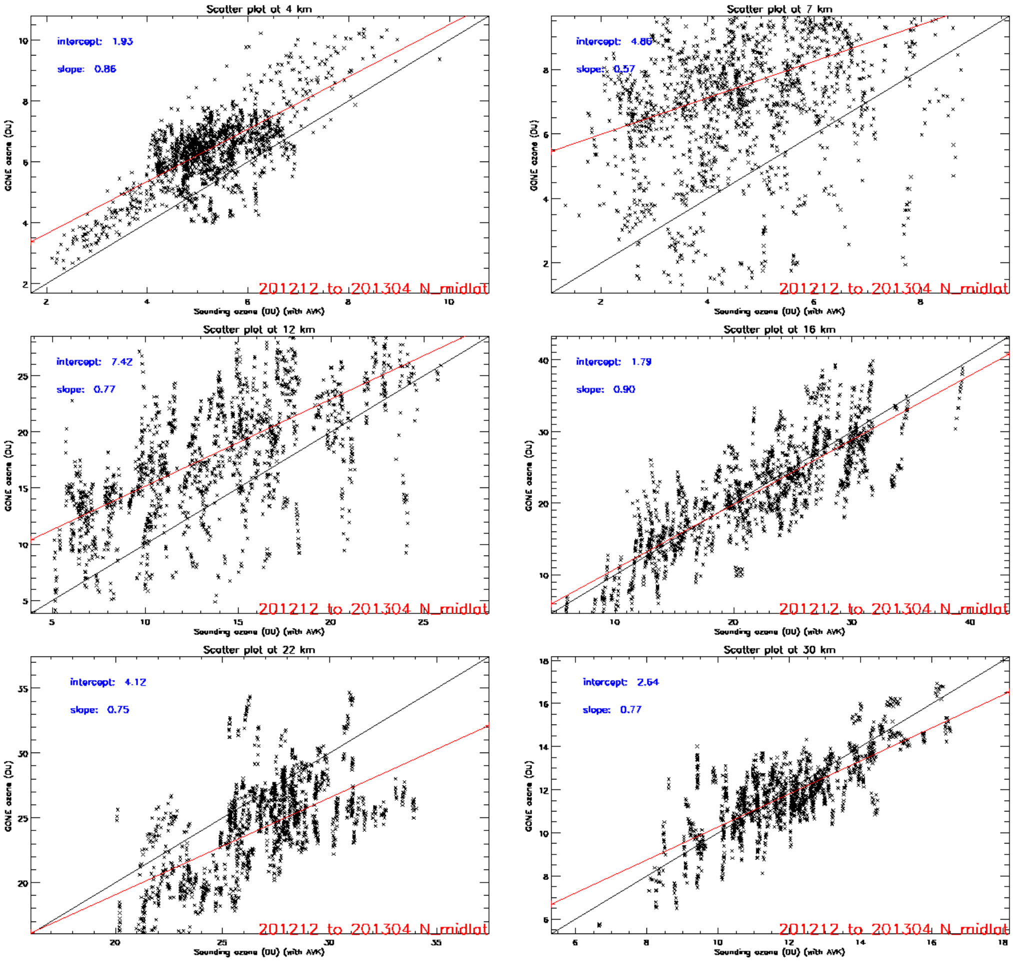


Figure 2.6b: Scatter plot at 6 different altitude levels for the stations at Northern Mid-Latitudes, applying the AVK for the HR pixels (December 2012-April 2013)

2.9 Cloud Cover

The influence of cloudcover on the retrieved ozone profiles only reveals a significant influence for the 30 km level for the Northern Polar stations. Figure 2.7 shows that with increasing cloud cover fraction (> 0.4), the ozone concentrations are more underestimated by GOME-2, while with a lower cloud cover fraction (<0.4) the ozone concentrations are overestimated. This could be related to the absence of a polarisation correction at higher SZA for higher latitudes.

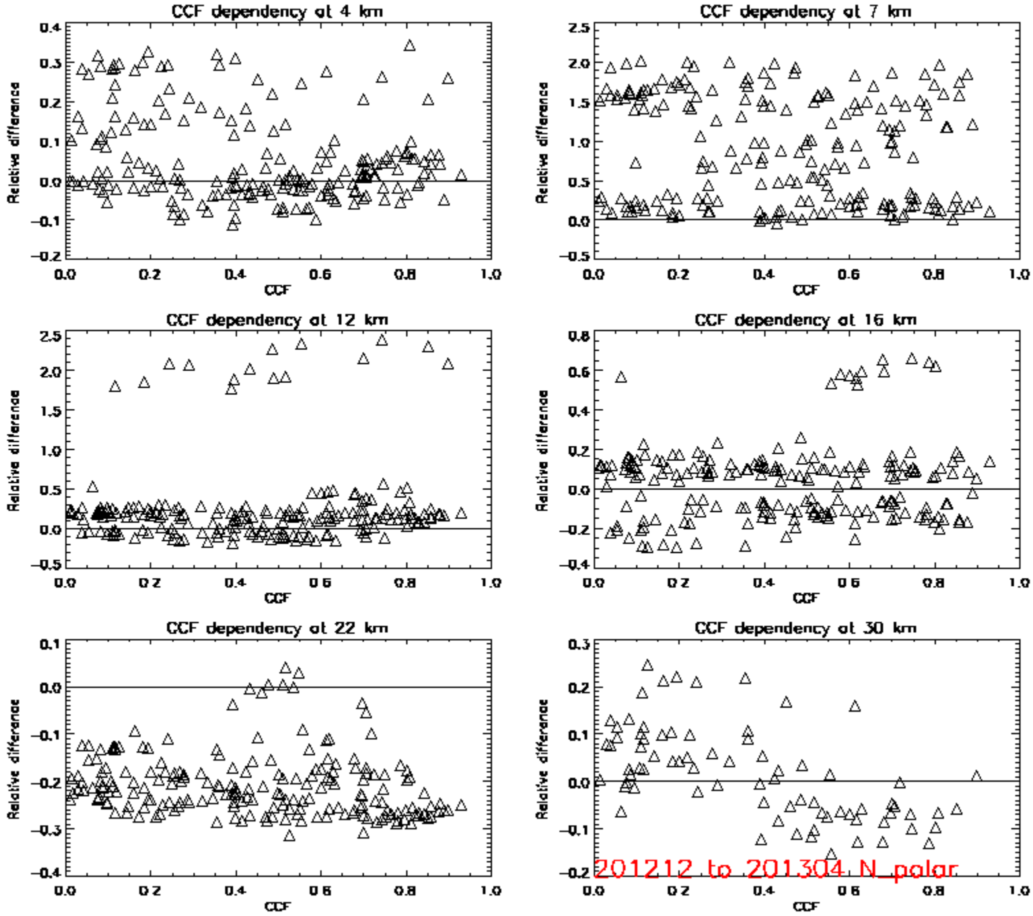


Figure 2.7: Cloud Cover Fraction dependency for six altitude levels for the HR-pixels, Northern Mid-Latitude stations, time period December 2012 –April 2013

2.10 Retrieval quality indicators

Retrieval quality indicators are important to give the user feedback on the quality of the ozone profile retrieval product. In the next 12 figures, a comparison is shown in different retrieval quality indicators between Metop-A and Metop-B for a comparable moment in the lifetime of the satellite, i.e. in the begining.

When comparing Metop-A and Metop-B, we took for Metop-A 2007-03-22 and for Metop-B 2013-03-29.

The first quality indicators shown in the figures 2.8, 2.9, 2.10 and 2.11 are respectively the Chi-square values for fitting windows 1, 2 and 3 and the RMS, which is also a measure for the fit-residual, based on the spectral measurements and the simulation.

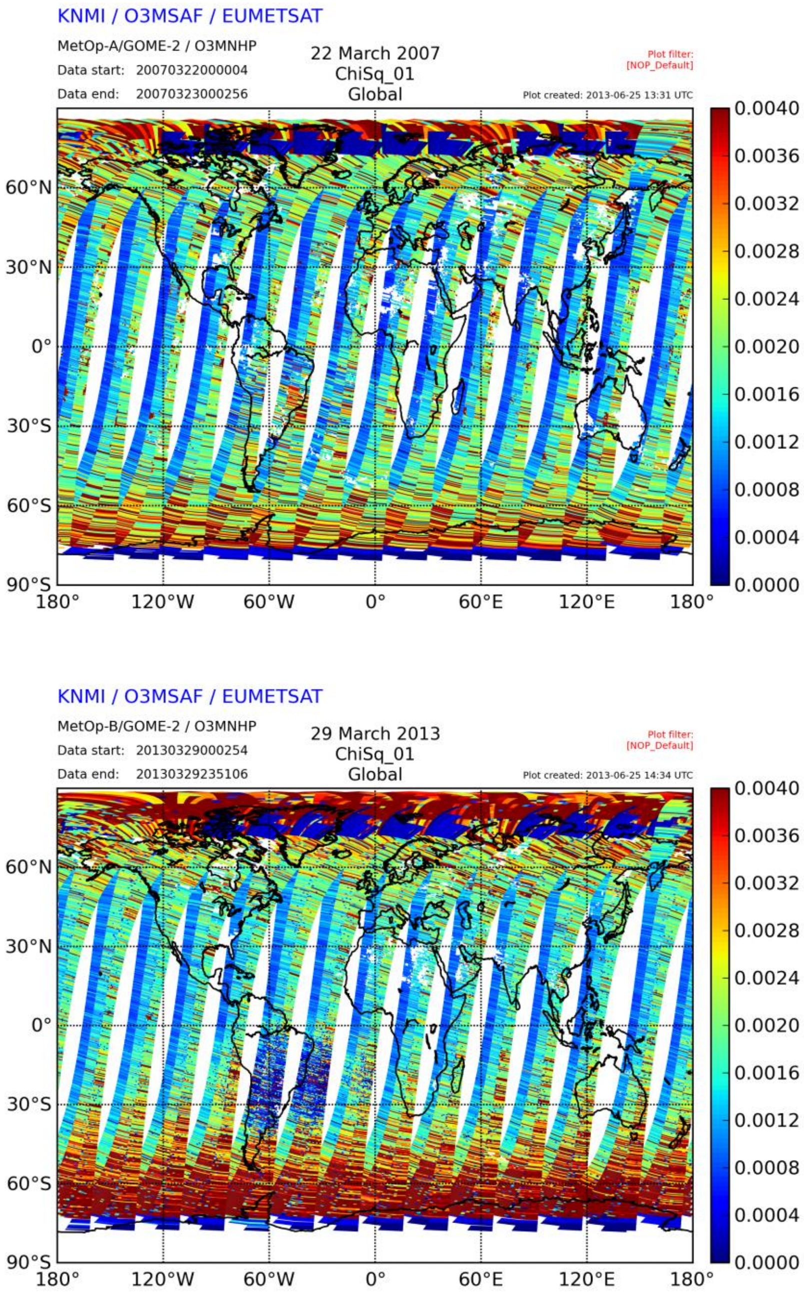
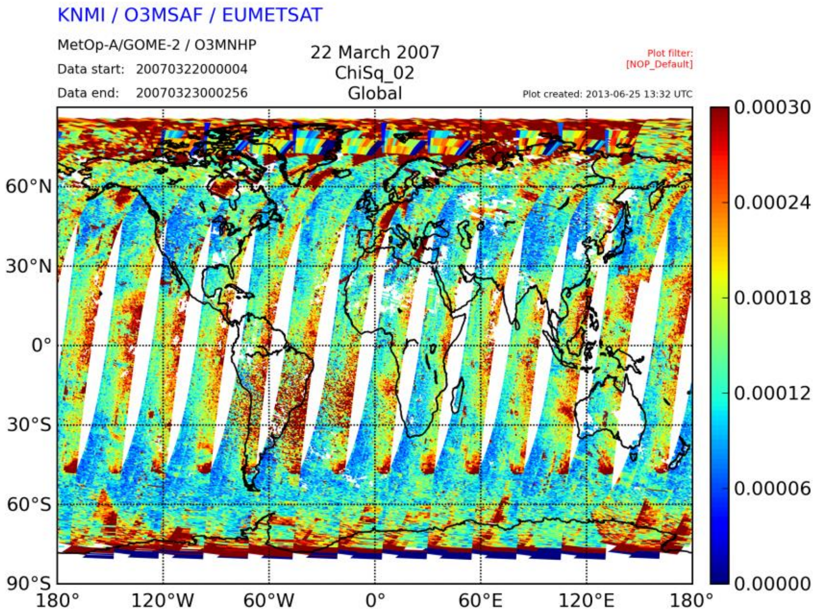


Figure 2.8: ChiSq value for fitting window 1 for data from March 22nd 2007 (upper) March 29th 2013(lower)



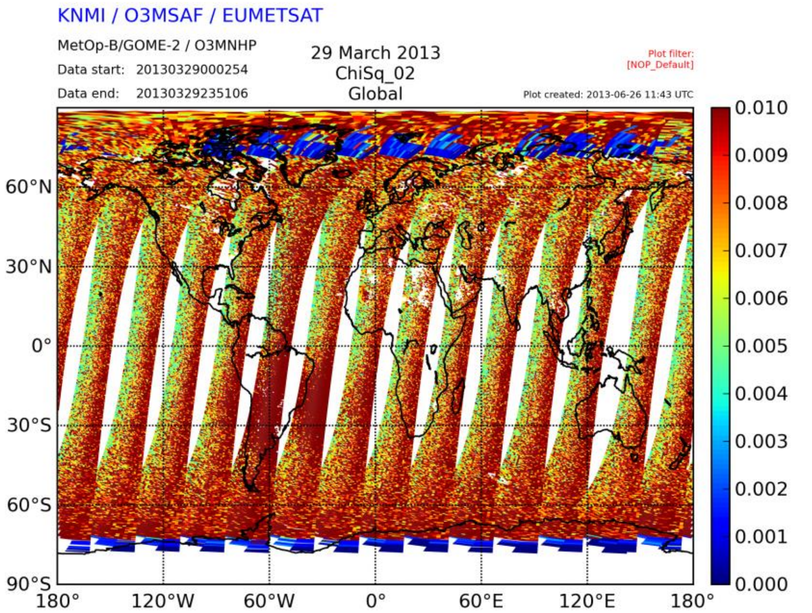


Figure 2.9: ChiSq value for fitting window 2 for data from March 22nd 2007 (upper) March 29th 2013(lower)

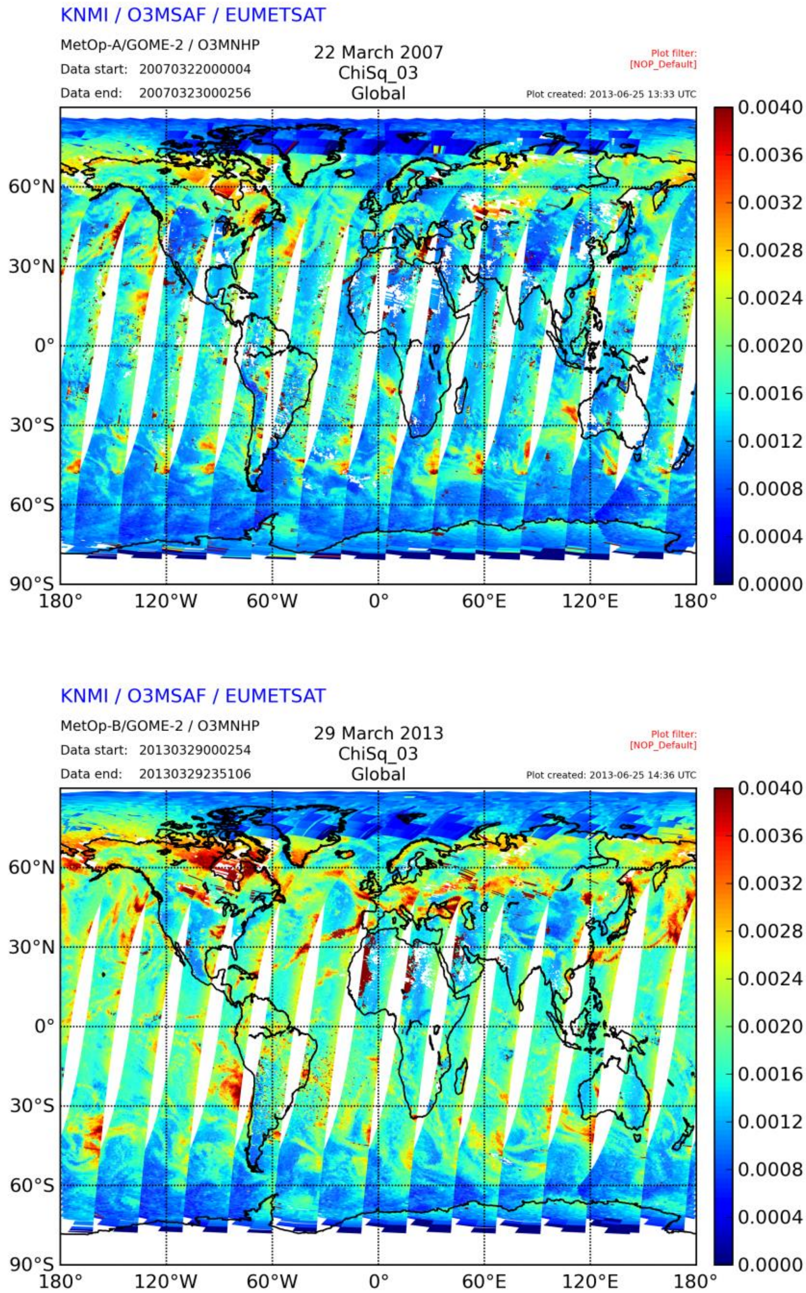
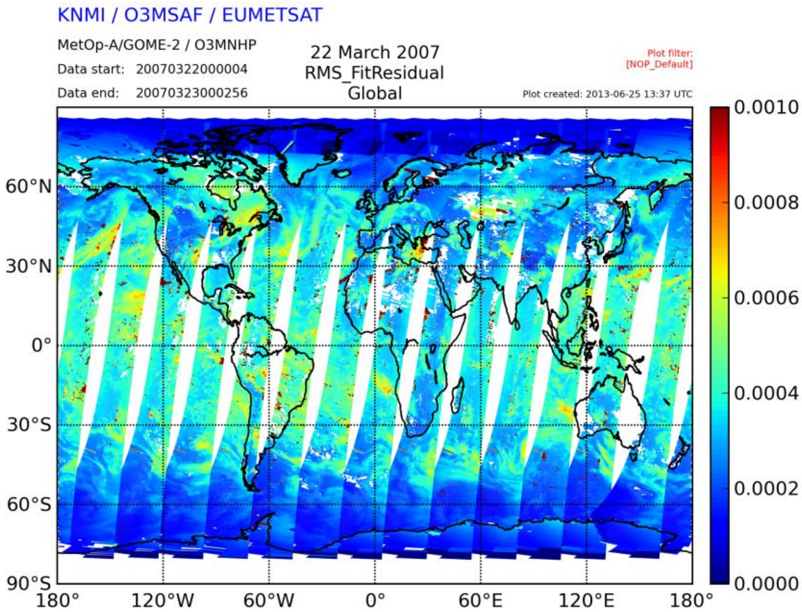


Figure 2.10: ChiSq value for fitting window 3 for data from March 22nd 2007 (upper) March 29th 2013 (lower)



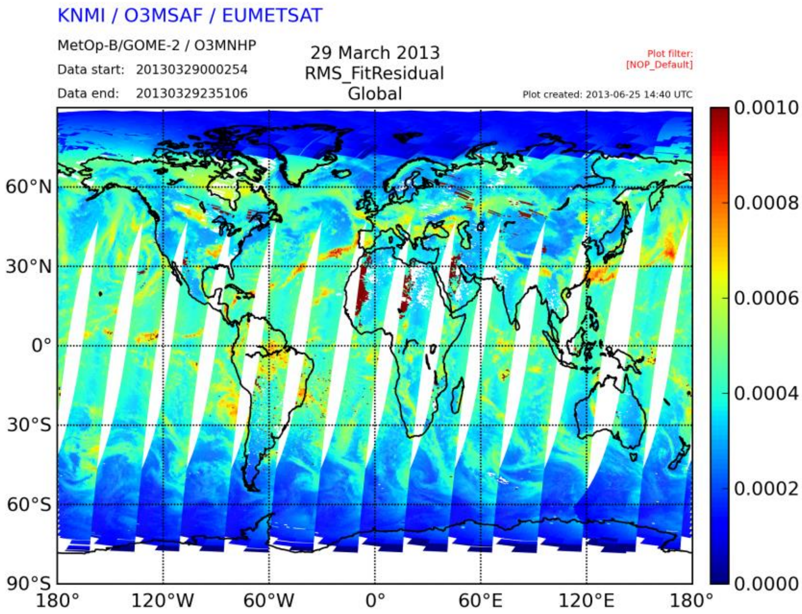
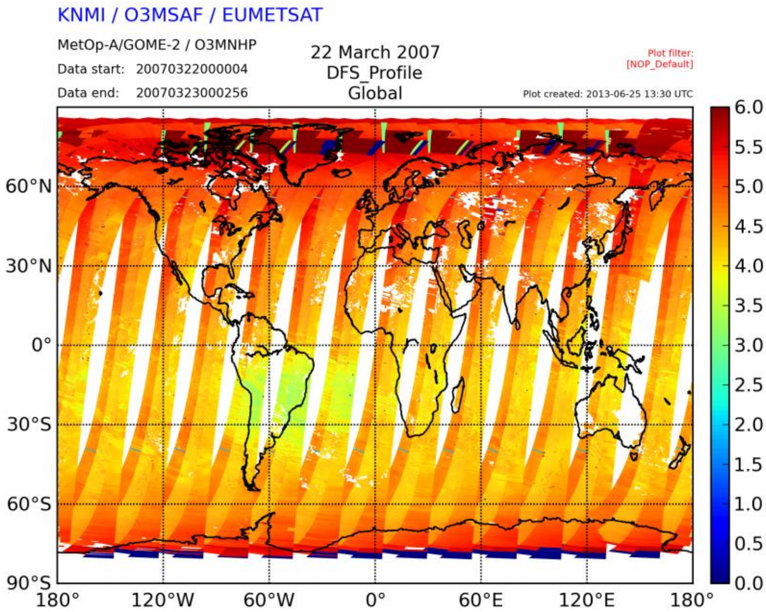


Figure 2.11: RMS Fit Residual values for data from March 22nd 2007 (upper) March 29th 2013 (lower)

As shown from figure 2.9, not the same scale has been applied for Metop-A and Metop-B, much lower values are achieved in 2013, which results from the number of spectral measurements in fitting window 2 (307 –311 nm). These values are much smaller than those measured in 2013 on Metop-B (283 – 311 nm). This makes it more difficult to compare both Chi-Square values. The RMS (Root Mean Square) is calculated afterwards.
The Chi-2 is calculated as one of the normal performance parameters in the Opera retrieval, the RMS is currently not part of the output product and was calculated in post-processing.



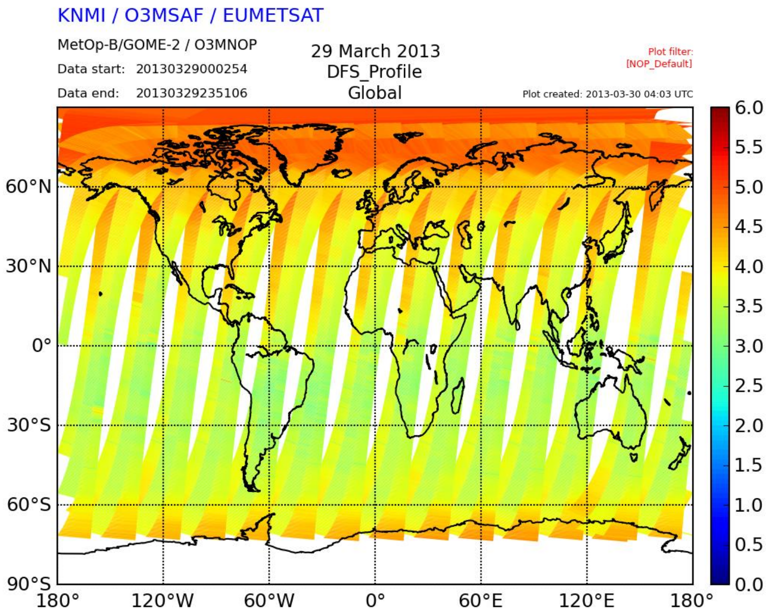


Figure 2.12: Degrees of Freedom for Signal for data from March 22nd 2007 (upper) March 29th 2013(lower)

Figure 2.12 shows the degrees of freedom for Signal (DFS). The DFS is an indicator of how much the retrieval has learned from the measurement. The DFS value is calculated as the sum of the diagonal of the averaging kernel matrix. A low DFS means that almost all information has been deduced from the a-priori and none from the spectral measurement. The DFS is seasonal, latitudinal and line of sight angle dependent due to the changes in light path and a-priori profile information.

Fig. 2.12 shows that Metop-B contains fewer DFS values when compared with the Metop-A values.

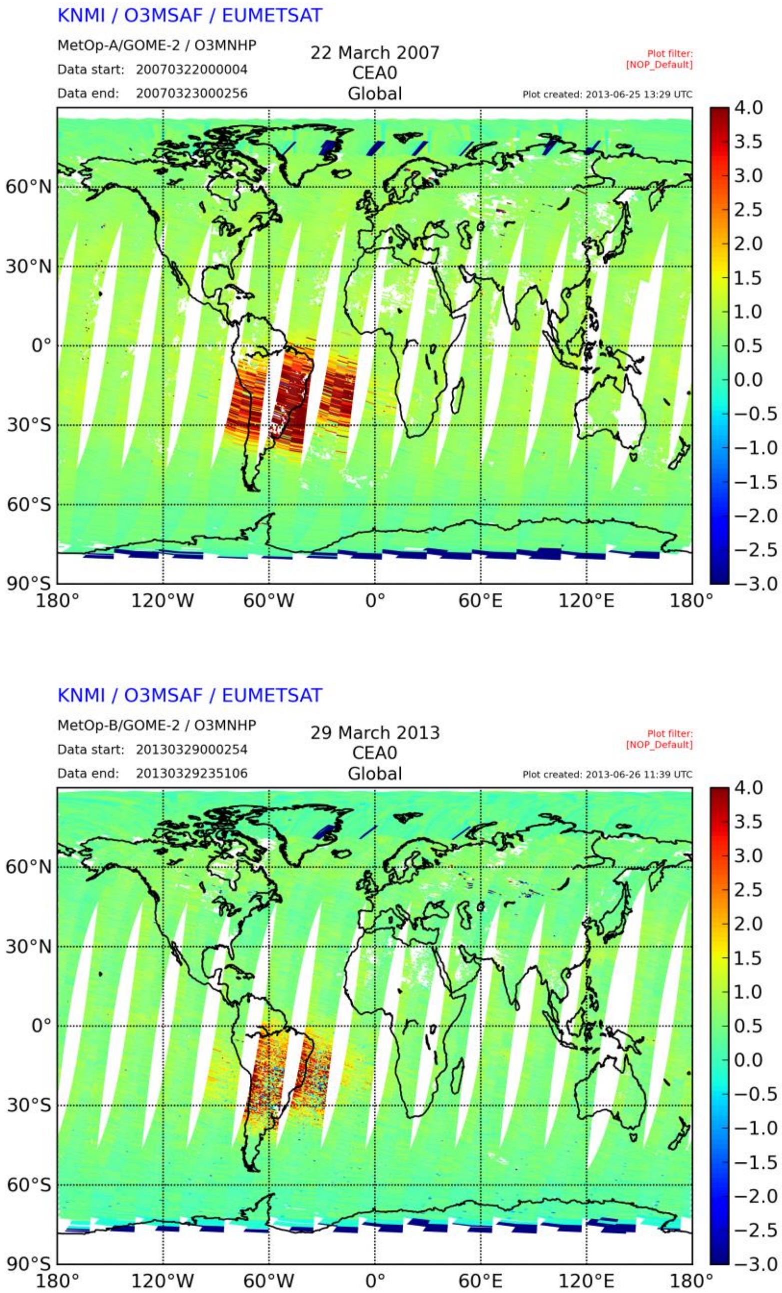


Figure 2.13: Additional offset to fit windows of channel 1 for data from March 22nd 2007 (upper) March 29th 2013(lower)

Figure 2.13 shows an additional offset to fit the windows of channel 1, which includes Band 1-a and Band 1-b. Optionally, an additional offset can be applied to the fitting windows (and this value can be fitted as part of the normal state vector). The additional offset CEA0 is an additive value applied to the radiance. The value fitted can be X * 1.0E+9 photons (shown in

Fig. 2.13) or a factor X * the radiance value at the lowest wavelength in the first window. The offset can be coupled between fitting windows (e.g.: fitted in window 1 and applied to window 2 as well). Tests have shown that this offset is required in at least Channel-1 to compensate/correct between the measured and simulated spectrum.

It is shown that both values for Metop-A and Metop-B are comparable at the beginning of the mission. Currently, the values for Metop-A are much more elevated.

3. Validation of ozone profiles with lidar and microwave instruments

3.1 Instruments

Lidars and microwave radiometers 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. This significantly extends the range covered by ozone sondes upwards to higher altitudes. It also provides a good overlap from 15 to 30 km altitude. Note that there are only about 10 operational lidar and microwave stations world-wide. These stations do provide regular data, but not as rapidly and operationally as the ozone sonde stations. Typically, ozone profiles to 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 15km to 50km, depending on the individual lidar system (Godin et al., 1989). Clouds and daylight conditions do not allow 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 lasts several hours. For the lidars, number density versus geometric altitude is the natural coordinate system of the measurement.

Microwave radiometers measure 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 retrieving the ozone profile. Similar to many satellite measurements, optimal estimation retrieval (Rodgers, 1990) provides ozone profiles in various coordinate systems, including number density versus altitude for the NDACC microwave profiles. Microwave ozone profiles typically cover 20 to 60 km altitude. In contrast to lidars, microwave radiometers have little weather dependence, and measure during daylight hours as well. On average, microwave profiles are measured on 20 days per month, or more. The integration time of one microwave profile varies from 30 minutes to 4-5 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 microwaves (Steinbrecht et al., 2006).

	lidar		microwave radiometer	
Height [km]	Precision [%]	height resol.[km]	Precision [%]	height resol.[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

3.2 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.
Table 3.2 lists all stations used here. They have profiles available for more than one month in the validation period.

Table 3.2: Reference ozone profiles available for comparison

Station (instrument)	Lat/lon	availability	profiles
Ny Alesund (μwave)	78.9/11.9	Dec 2012 – April 2013	282
Hohenpeissenberg (lidar)	47.8/11.0	Dec 2012 – April 2013	20
Bern (μwave)	47.0/7.6	Dec 2012 – Mar 2013	719
Haute Provence (lidar)	43.9/5.8	Dec 2012 – April 2013	44
Mauna Loa (lidar)	19.5/-155.6	Dec 2012 – Feb 2013	14
MaunaLoa (μwave)	19.5/-155.6	Dec 2012 – Feb 2013	57
Lauder (μwave)	-45.0/169.7	Dec 2012 – Feb 2013	46

3.3 Comparison procedure

Generally, the comparison procedure is the same as for the ozone sondes, 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.

3.4 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.
Microwave radiometers measure around the clock, typically one profile every hour. So usually microwave 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 quality_processing_status of "Overall convergence, successful retrieval".(see Product User Manual pages 22-23) and with ground pixels centers closer than 200 km to the validation stations were considered. A 200 km radius typically gives about 50 co-located GOME-2 high resolution profiles per station and per day. The number of coarse resolution profiles is lower, but still provides a good statistical basis. Larger co-location radii result in larger geophysical differences, smaller radii result in too few comparisons cases.

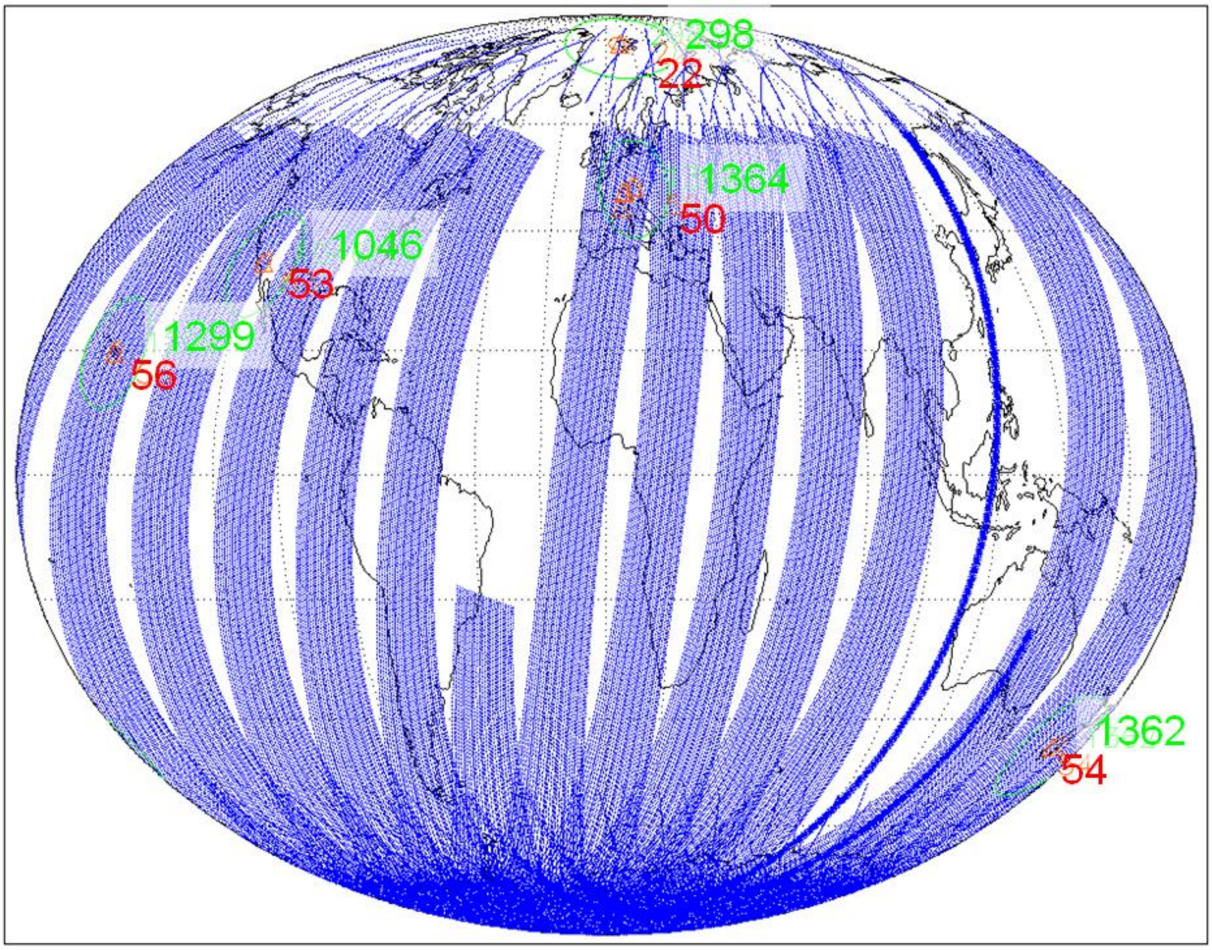


Fig 3.1: World map with GOME-2 high resolution ground pixels for Dec.17th, 2012. Red triangles show NDACC lidar and microwave stations. Green circles are for 1000 km radius around the station. Green numbers give the number of GOME-2 ground pixels within the 1000 km circle. Red circle and red numbers are for a 200 km radius. GOME-2 swaths start in the North near the date-line, and then move westwards around the globe. All measurements are taken in the local morning. Note that in December there is no sunlight near the North Pole, so there are no measurements there.

3.5 Pre-processing of the ozone profiles.

Same as the ozone sonde data, the lidar and microwave 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 $X_{ref.}$ They give the average number density for the GOME-2 retrieval layers. Since the upper and lower altitude and pressure of the GOME-2 layers are given in the GOME data files, average number density (from lidar or microwave) and partial column (in DU, from GOME-2 retrieval) are easily converted and are equivalent. The resulting $X_{ref.}$ profiles have the same layer altitudes as the GOME-2 retrieved profiles, and can be directly compared.

In the next step, the X_{ref} lidar and microwave profiles are further smoothed over altitude by applying the GOME-2 averaging kernels (with proper scaling, see Eq. 14 and Fig. 11 of the

Algorithm Theoretical Basis Document). These smoothed profiles X_{AVK} have altitude resolution comparable to the GOME-2 profiles (or coarser).

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% from the measurement. At other altitudes (usually the troposphere and mesosphere), the GOME-2 profile comes to 80 or 90% from an a-priori profile (i.e. from climatology). For the validation of the retrieval process, it makes sense to also consider reference profiles 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, and also Eq. 13 of the Algorithm Theoretical Basis Document) describes the underlying mathematics. The resulting profiles are called $X_{AVK\ apriori}$ in the following.

Fortunately, in nearly all cases the validation of GOME-2 profiles gives almost the same results for all three "flavors" of smoothed reference profiles (X_{ref}, X_{AVK}, $X_{AVK\ apriori}$).

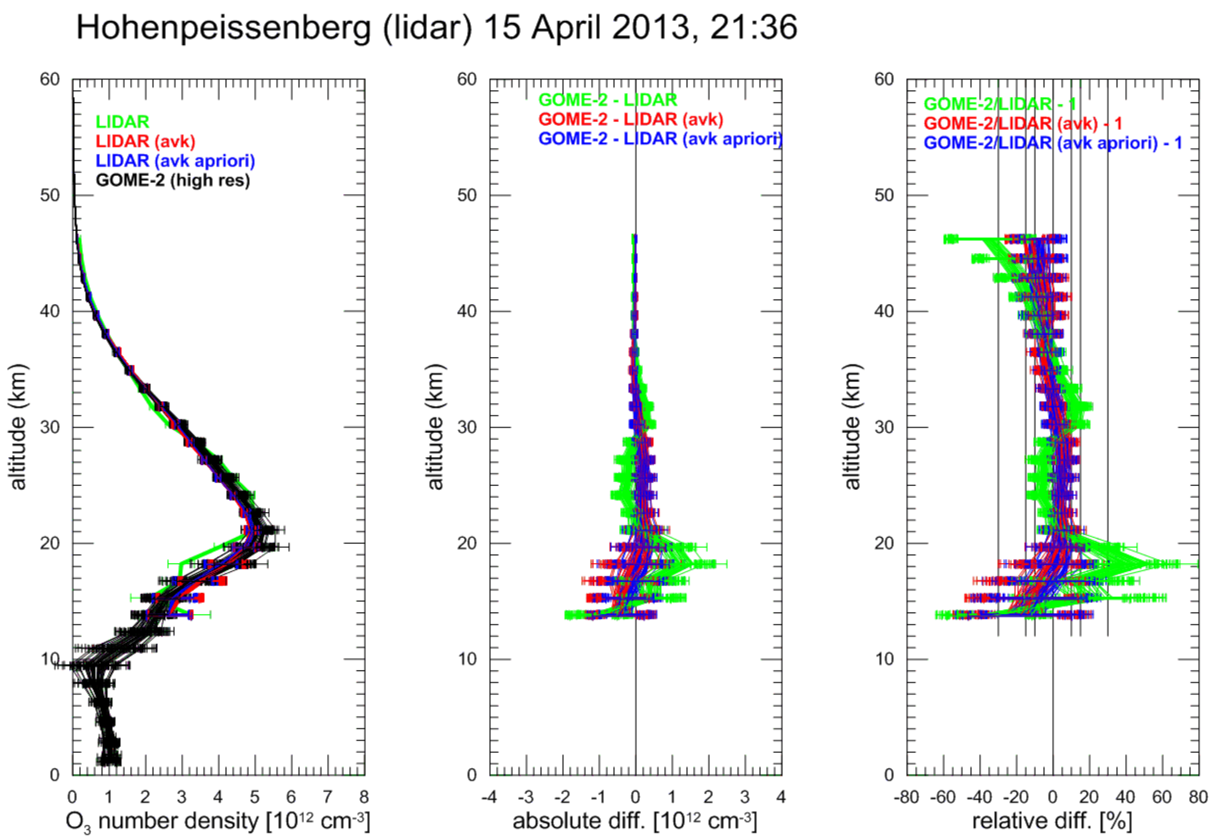


Fig 3.2: Example for the comparison of a lidar profile at Hohenpeissenberg, Germany, (green X_{ref}, red X_{avk}, blue $X_{avk,\ apriori}$) with the matching GOME-2 MetopB high resolution profiles (black). Left panel: Profiles. Middle panel: Absolute differences. Right panel: Relative differences. Note that the GOME-2 layer altitudes and averaging kernels vary slightly from profile to profile. This results in small differences in the smoothed 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.

3.6 Results

Figures 3.2 and 3.3 give typical examples for the comparison between ground-based and the matching GOME-2 profiles. In most cases, GOME-2 profiles and the ground-based profiles agree reasonably well. Most comparisons show relative differences smaller than the ±30% threshold accuracy specified for GOME-2. However, many relative differences, especially around 40 km, also fall outside of the ±10% optimum accuracy specified for GOME-2. As with the sondes, the largest differences are often seen below 20 km. Larger differences below 20 km are common features of nearly all ozone profile comparisons: The lower stratosphere is a region with large vertical, horizontal and temporal ozone gradients. The resulting large spatial and temporal ozone variations make profile comparisons difficult and usually result in larger differences and larger uncertainties below 20 km (see also Section 2).

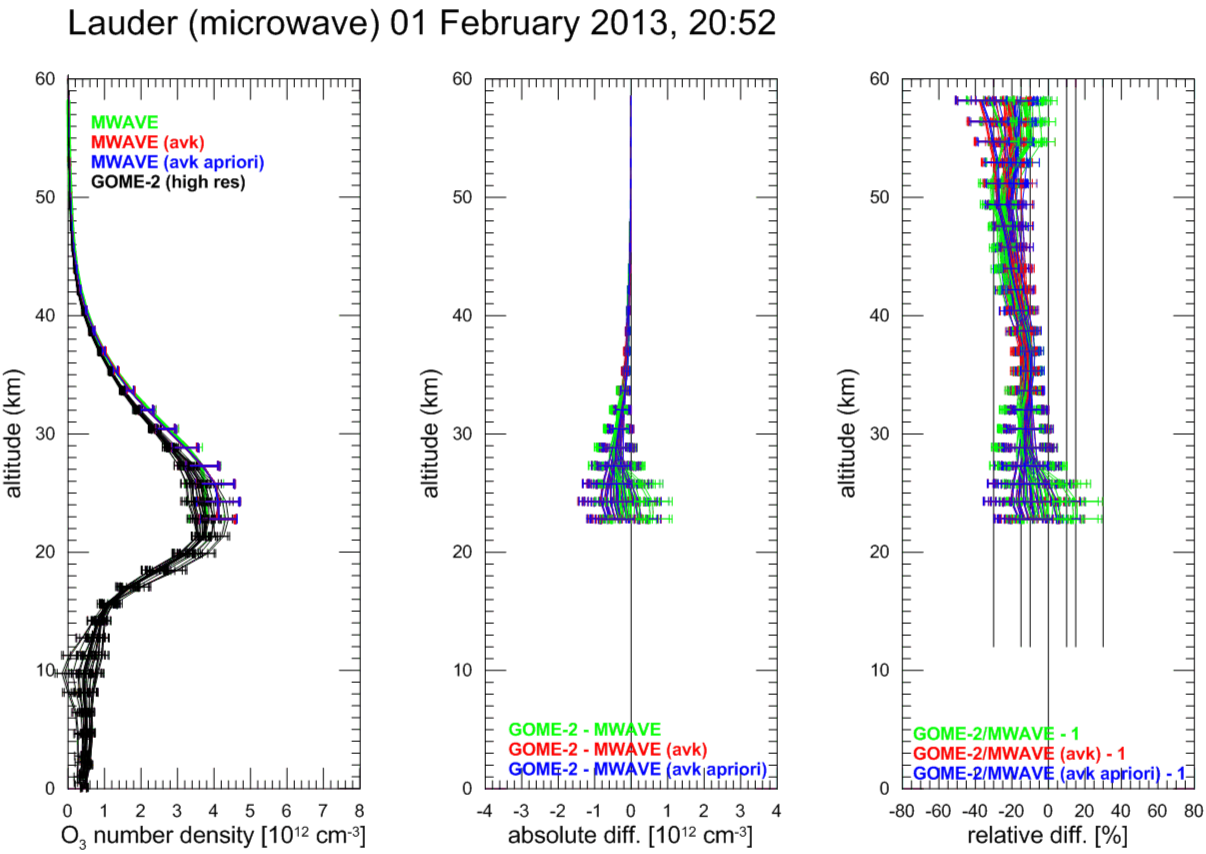


Fig 3.3: Same as Fig. 3.2, but for a comparison with the microwave radiometer at Lauder, New Zealand.

3.7 Mean differences between GOME-2/MetopB and co-located lidar and microwave ozone profiles

Figures 3.4 and 3.5 summarize all the individual comparisons with GOME-2/MetopB high resolution profiles for the microwave radiometer at Bern, and for the lidar at Mauna Loa. Results cover the December 2012 to April 2013 time frame. For clarity, only the results for X_{AVK} are shown. The dots are individual profile differences, the lines and error bars give the average difference profile, and the standard deviation. Generally, GOME-2 and ground based profiles agree in magnitude and shape. Most relative differences lie within the ±30%

threshold accuracy. The average differences indicate a GOME-2 low bias by about 10% above 40 km, both at Bern and at Mauna Loa. Between 30 and 40 km, the average difference is close to 0%. Below 30 km the average differences increase again, to -10% near 20 km for Bern, and to +30% near 20 km for Mauna Loa. The ±30% threshold accuracy requirement for GOME-2 is fulfilled at most altitudes between 20 and 60 km cases. Between 25 and 40 km, ±15% target accuracy and ±10% threshold accuracy are reached in many cases.

Although not shown here, very similar results are obtained for the less smoothed X_{REF} profiles, or for the $X_{AVK\ apriori}$ profiles, where the a-priori contribution has been included.

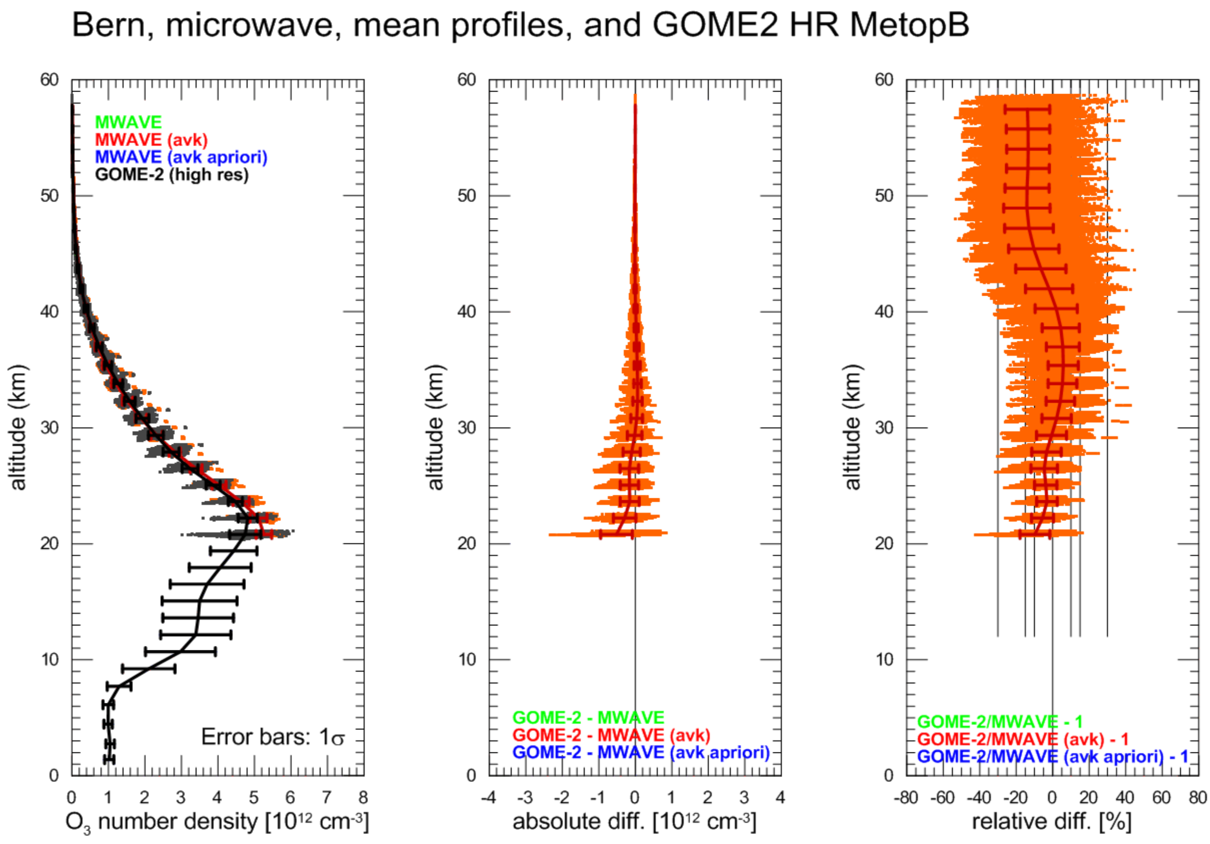


Fig 3.4: Summary of all individual comparisons between GOME-2 high resolution and microwave ozone profiles at Bern, Switzerland, Left panel: Profiles. Middle panel: Absolute differences. Right panel: Relative differences. Only X_{avk}, microwave profiles are shown here. Red lines give the average over all profiles, error bars are 1 standard deviation. 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.

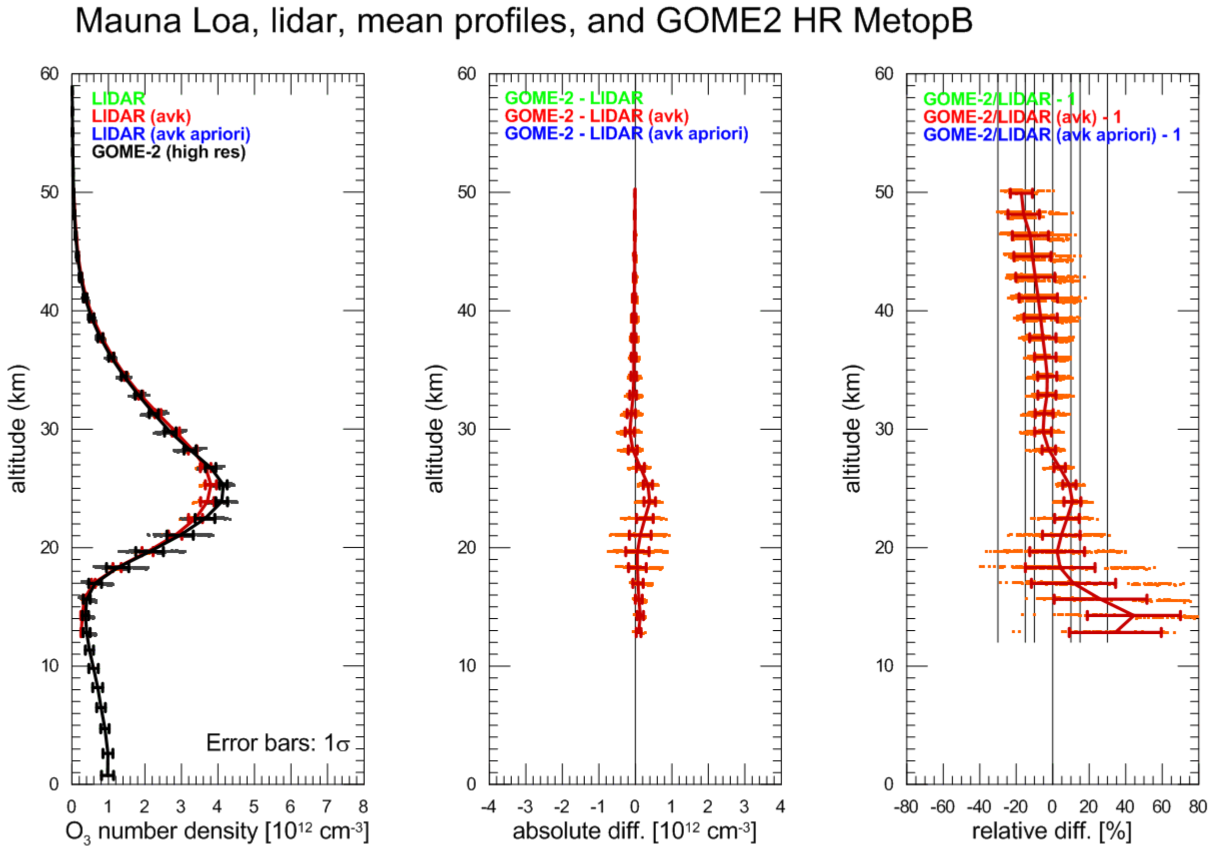


Fig 3.5: Same as Fig. 3.4, but for the lidar at Mauna Loa.

The average relative differences between GOME-2 and the ground-based instruments for all stations are given in Fig. 3.6. They confirm the findings of the previous examples: Generally, the bias between GOME-2 profiles and NDACC reference profiles is smaller than 10 to 15% between 20 and 35 km altitude. Around 40 km, GOME-2 shows a negative bias, about -10%. This bias increases to -15% around 50 km altitude. Below 20 km, uncertainties are larger, as mentioned, and results are much less conclusive. There is no indication for a large bias of the GOME-2 profiles below 20 km. Above this altitude, the GOME-2 bias is usually smaller than the specified target accuracy of ±15%. While all lidars show very similar differences to GOME-2, the microwave results seem to fall into two groups: One for the microwave radiometers operated by the University of Massachusetts (Mauna Loa and Lauder), the other (Ny Alesund, and Bern) for instruments originally developed at the University of Bern twenty years ago. It remains to be seen, if additional data in the future confirm this.

Overall, the GOME-2 MetopB bias results in Fig. 3.6 are very similar to previous validation results for GOME-2 on MetopA. Using data from 2007 to 2011, the March 2012 MetopA validation report, for example, found a GOME-2 MetopA bias against the NDACC instruments of 0%±4% near 20 km, -5%±4% near 30km, -15%±3% near 40km, and -10%±3% near 50km (1σ uncertainties) – quite comparable to the current findings for GOME-2 MetopB. Based on these findings, it appears that GOME-2 MetopB produces ozone profiles that are very comparable, on average, to the GOME-2 MetopA ozone profiles.

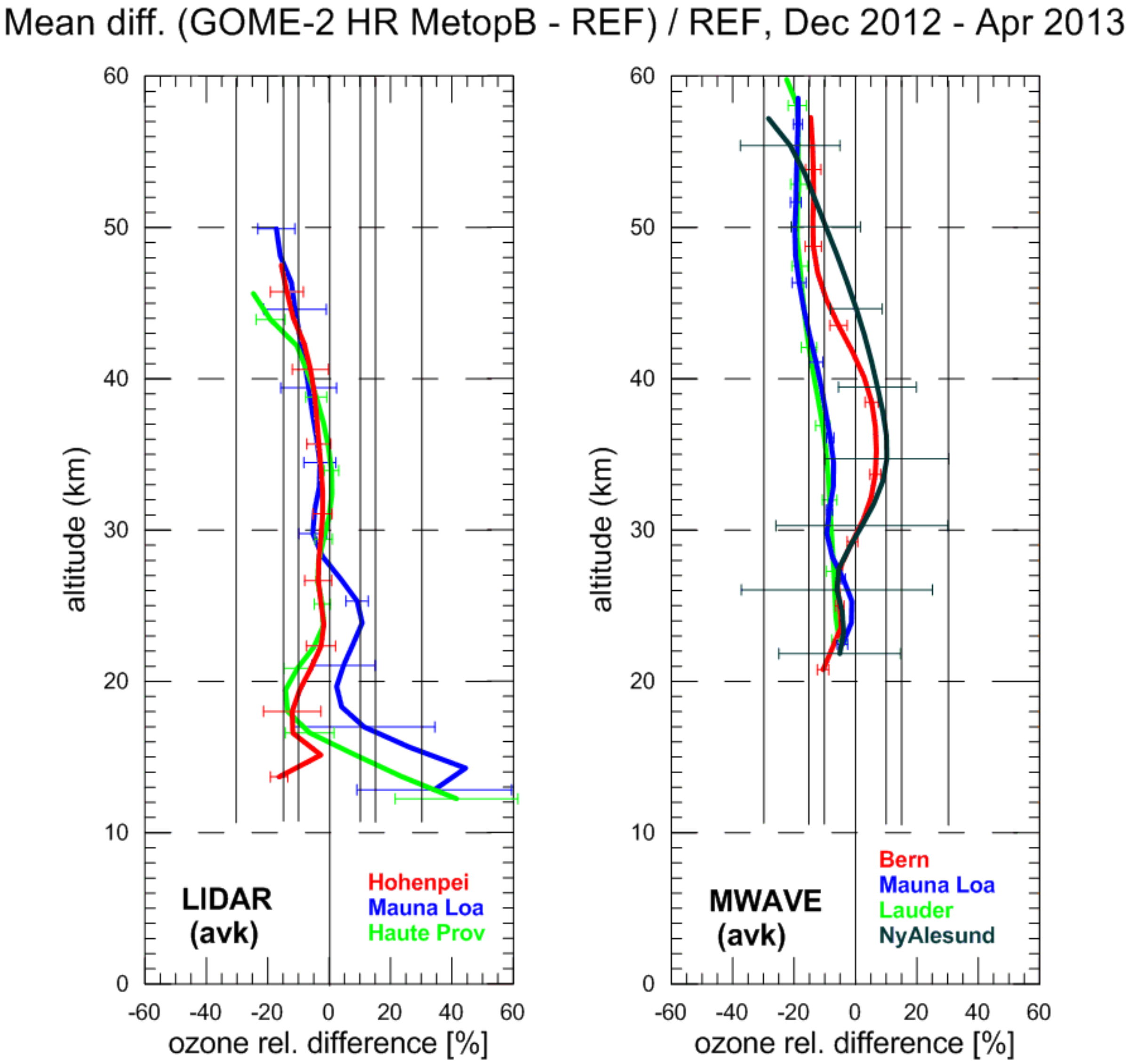


Fig 3.6: Average relative difference between GOME-2 MetopB high resolution profiles and the ground-based reference profiles (X_{avk}). Results are for the December 2012 to April 2013 comparison period. Left panel: Results for lidar stations. Right panel: Microwave stations. The vertical lines at ±30%, ±15%, and ±10% are the threshold, target, and optimum accuracies specified for the GOME-2 product. Error bars are 2σ.

3.8 Ozone variations as seen by GOME-2/MetopB and by lidar and microwave ozone profiles

So far we have focused on average differences between GOME-2 and ground-based ozone profiles. Of course it is desirable to have only small systematic differences between the two. In addition, however, it is necessary that geophysical ozone variations are tracked in the same way by GOME-2 and the ground-based measurements. To check this, it is necessary to have actual ozone variations in the first place. Since we are only considering a short time period here, our validation is restricted here, because we sample only a limited natural variation, and very little variation at some altitudes.

Figures 3.7 to 3.13 present scatter plots that compare the ozone values measured by GOME-2 (along the y-axis) with ozone measured by the ground-based instruments (along the x-axis). The ground-based profiles were smoothed by the GOME-2 averaging kernels (X_{avk}). Using the raw profiles (X_{ref}), or including the a-priori contribution ($X_{avk,apriori}$), gives very similar results (not shown here). Each dot represents one comparison, each panel one altitude level.
Ideally, all dots would line up on the diagonal 1-to-1 line. However, both measurements are subject to instrumental noise and possible systematic biases. Also both instruments nearly always measure different air-masses. These uncertainties and the inevitable spatial and temporal miss-matches scatter the points away from the ideal line.

The red-lines in the plots show the straight lines that best fit the data. Ideally, the red lines would go through the middle of the data dots, and through the origin (0, 0). All data points would line up on the red lines. This alignment is measured by the correlation coefficient (R2 in the plots). A correlation of 1 (or -1) corresponds to perfect alignment, 0 corresponds to a random “cloud” of dots. Slopes (coefficient B) different from 1 indicate systematic biases. Axis intercepts (coefficient A) different from 0 indicate systematic offsets. Note however, that the current least squares regression fit does not account for data uncertainties in the x-direction, and tends to make the regression lines “too flat”. This needs to be fixed in the future.

All scatter plots show a similar picture: Generally, GOME-2 and the reference instruments give a similar range of observed values. Banding occurs when only a small number of profiles is available, usually for the ground-based instruments (e.g. Fig. 3.8 at 18 km). In most cases, the data points align reasonably well with the red fit lines. The red fit lines are not too far away from the ideal 1-to-1 line. Data dots and red fit lines tend to lie below the 1-to-1 line. This indicates that GOME-2 values tend to be lower than the reference – fully consistent with the general low bias discussed in the previous section.

At the highest and lowest altitudes, instrumental noise and/or atmospheric variability become larger. There, data points scatter more widely. Fit lines become flat, or even have negative slope. This can indicate that more data are needed for validation, and that the current validation period is too short. Another example for not enough data is the altitude region around 27 km above Mauna Loa (Figs. 3.11 and 3.12). There the sampled geophysical variation is very small. Then, very little can be said about the tracking of GOME-2 and the ground-based data.

Overall, however, the scatter plots give a strong indication that the GOME-2 ozone profiles do a good job following the atmospheric ozone variations as seen by the ground-based reference profiles.

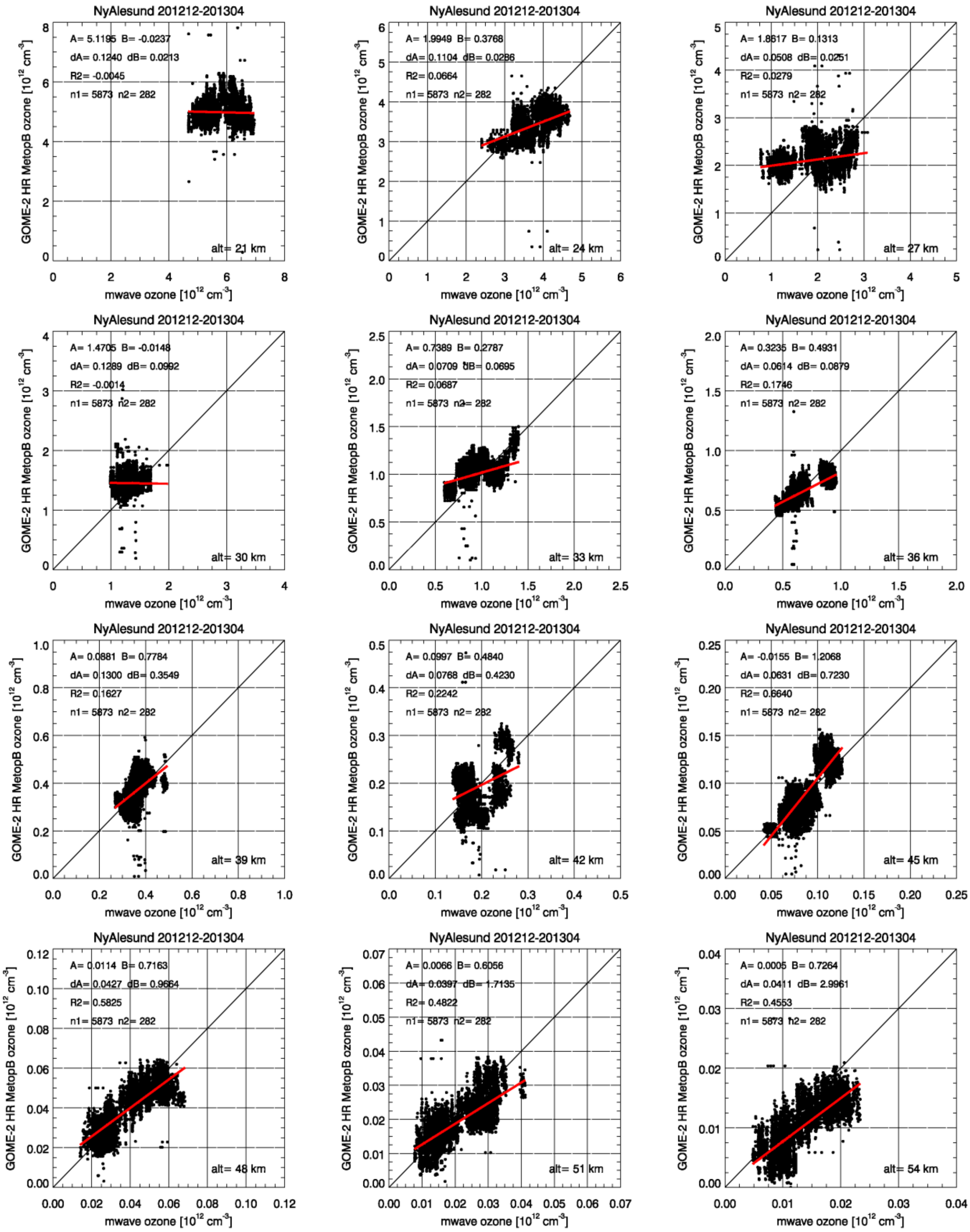


Fig 3.7: Scatter plot of GOME-2 MetopB high resolution ozone versus ozone from the NDACC microwave radiometer at NyAlesund (X_{avk}), i.e. smoothed by the GOME-2 averaging kernels. The different panels are for different altitudes. Dots are individual profile comparisons. The diagonal is the ideal 1-to-1 line. Red lines are the least squares fitted straight line. A is the resulting offset, with 1σ uncertainty dA, B is the slope, with 1σ uncertainty dB. R2 is the linear correlation between GOME-2 and the reference. The number of GOME-2 profiles is given by n1, the number of ground-based profiles by n2.

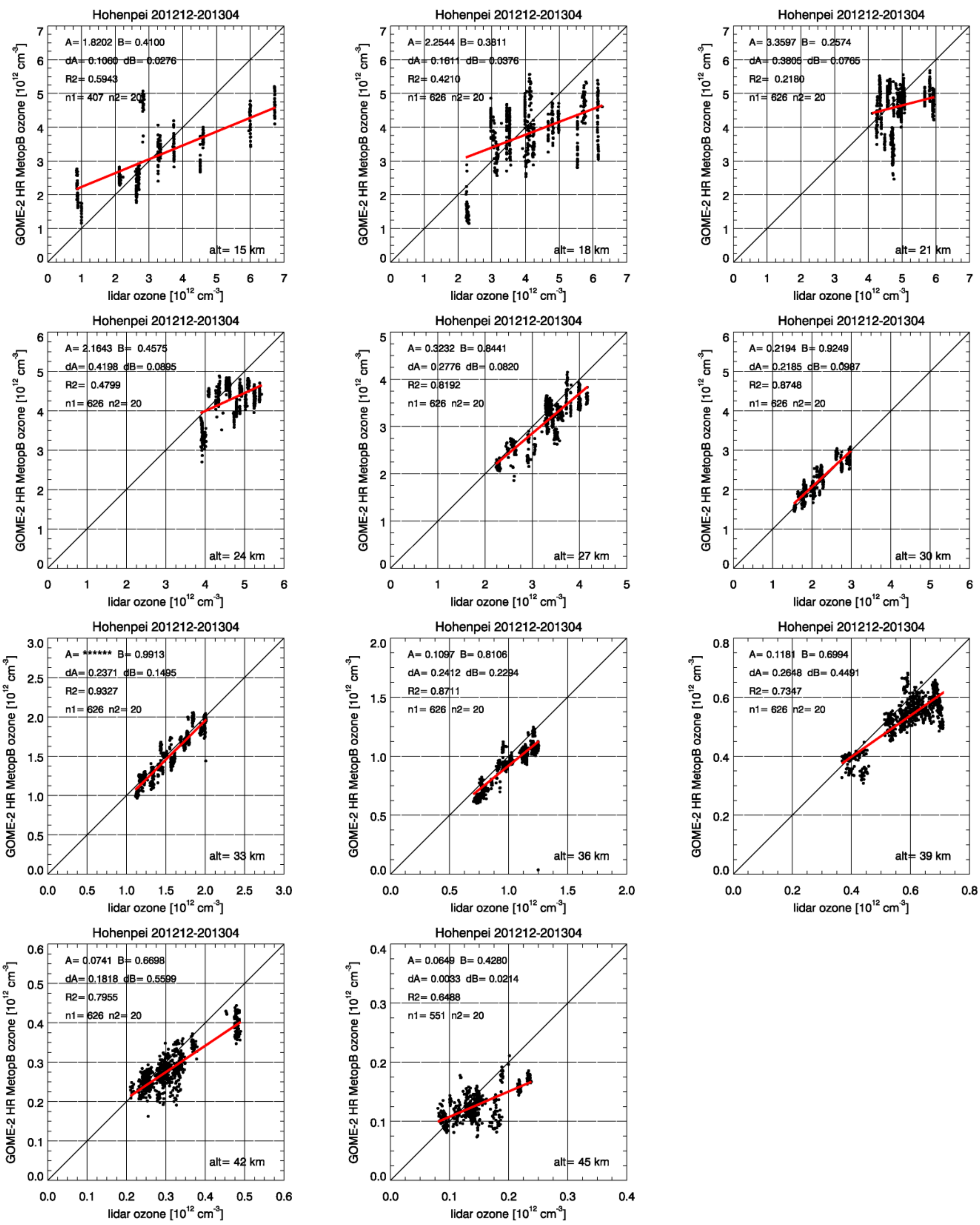


Fig 3.8: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC lidar at Hohenpeissenberg

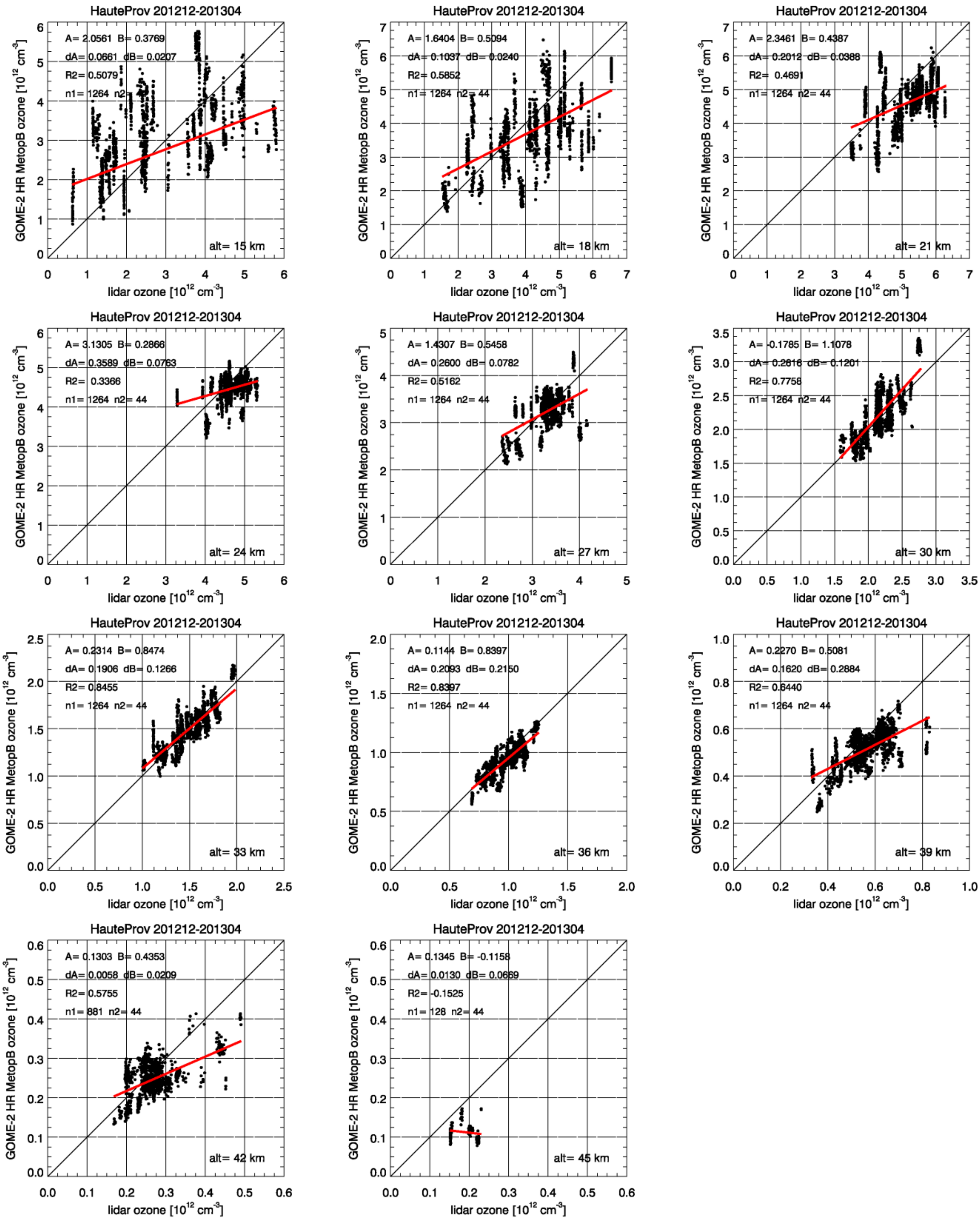


Fig 3.9: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC lidar at Haute Provence

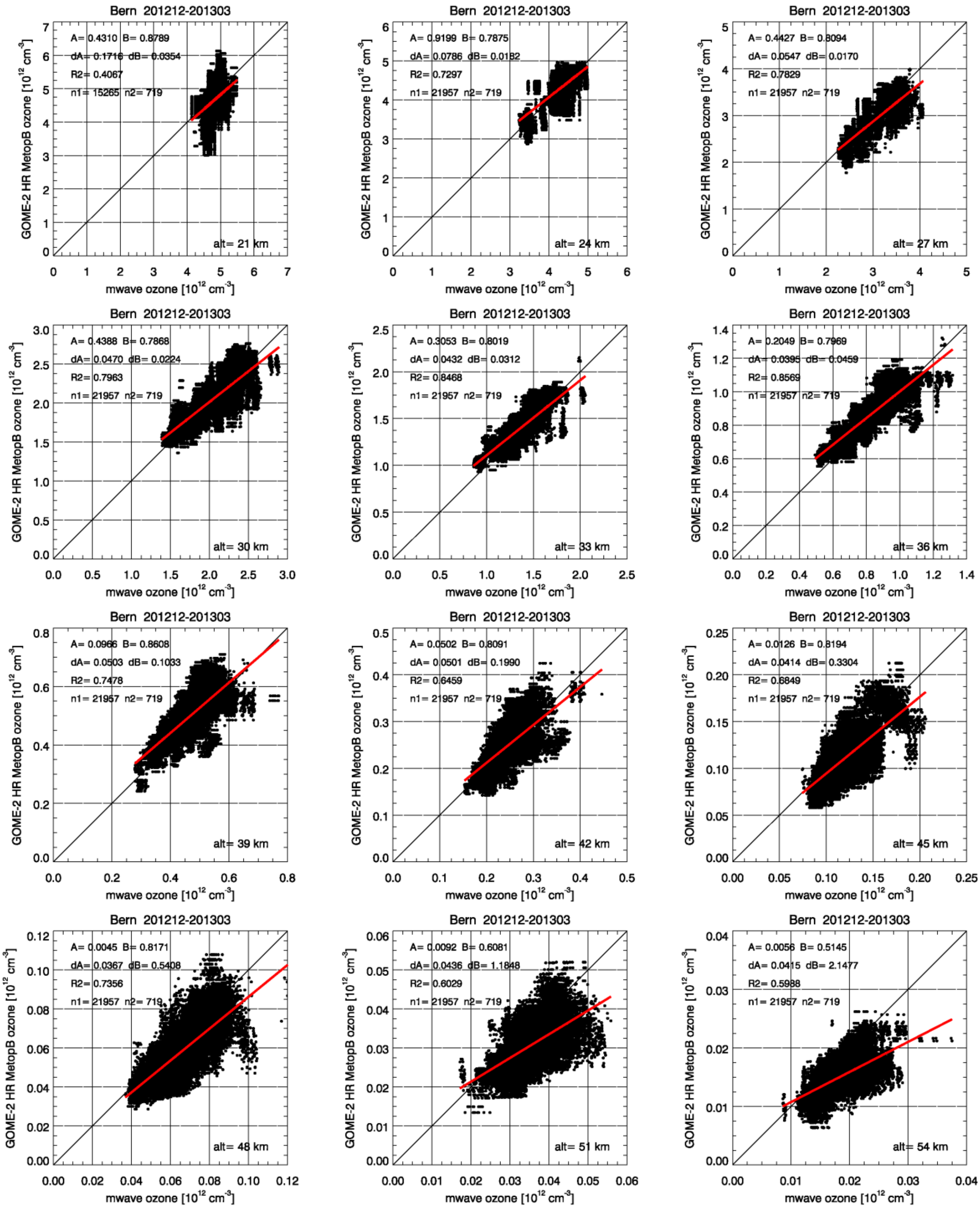


Fig 3.10: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC microwave at Bern

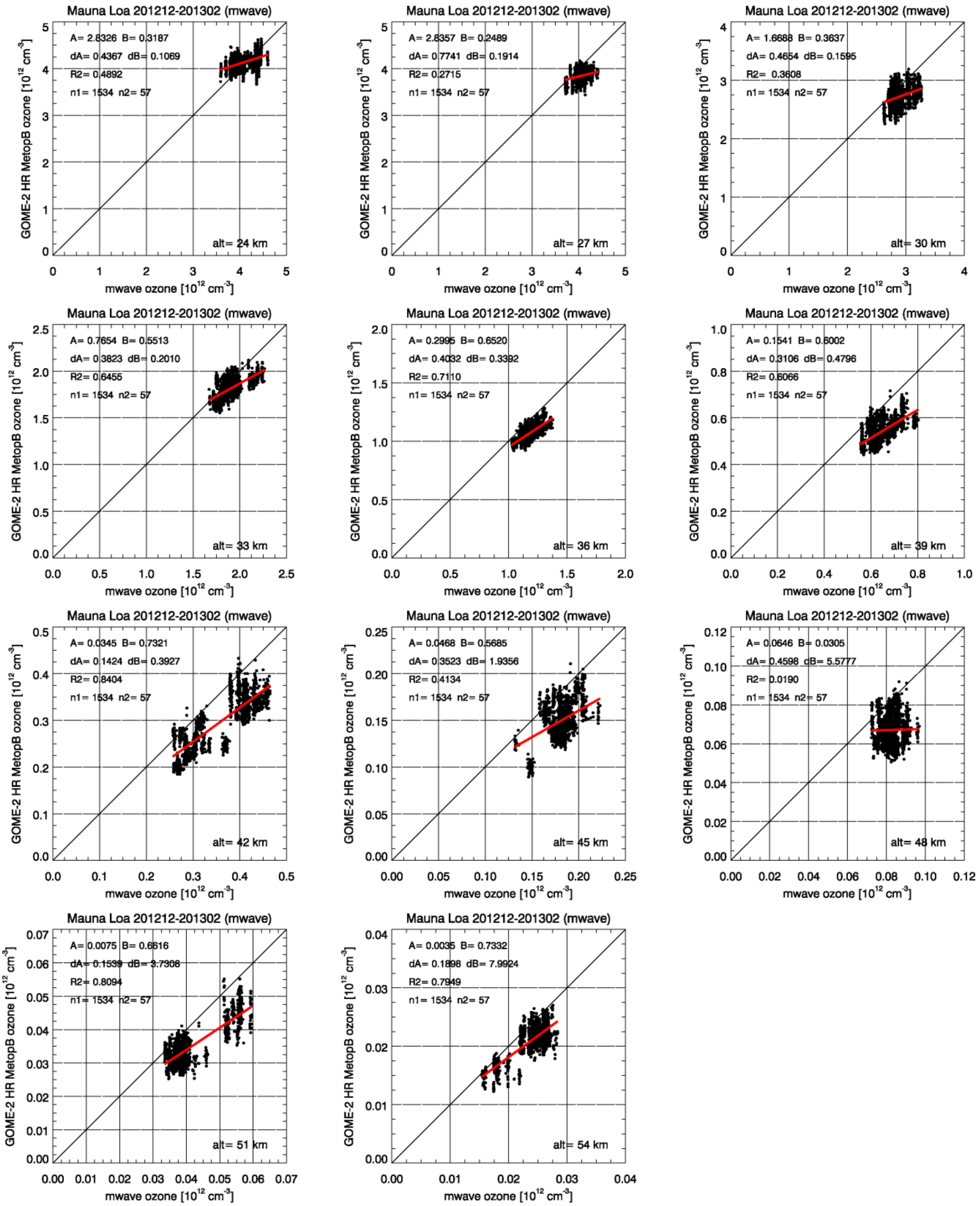


Fig 3.11: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC microwave at Mauna Loa.

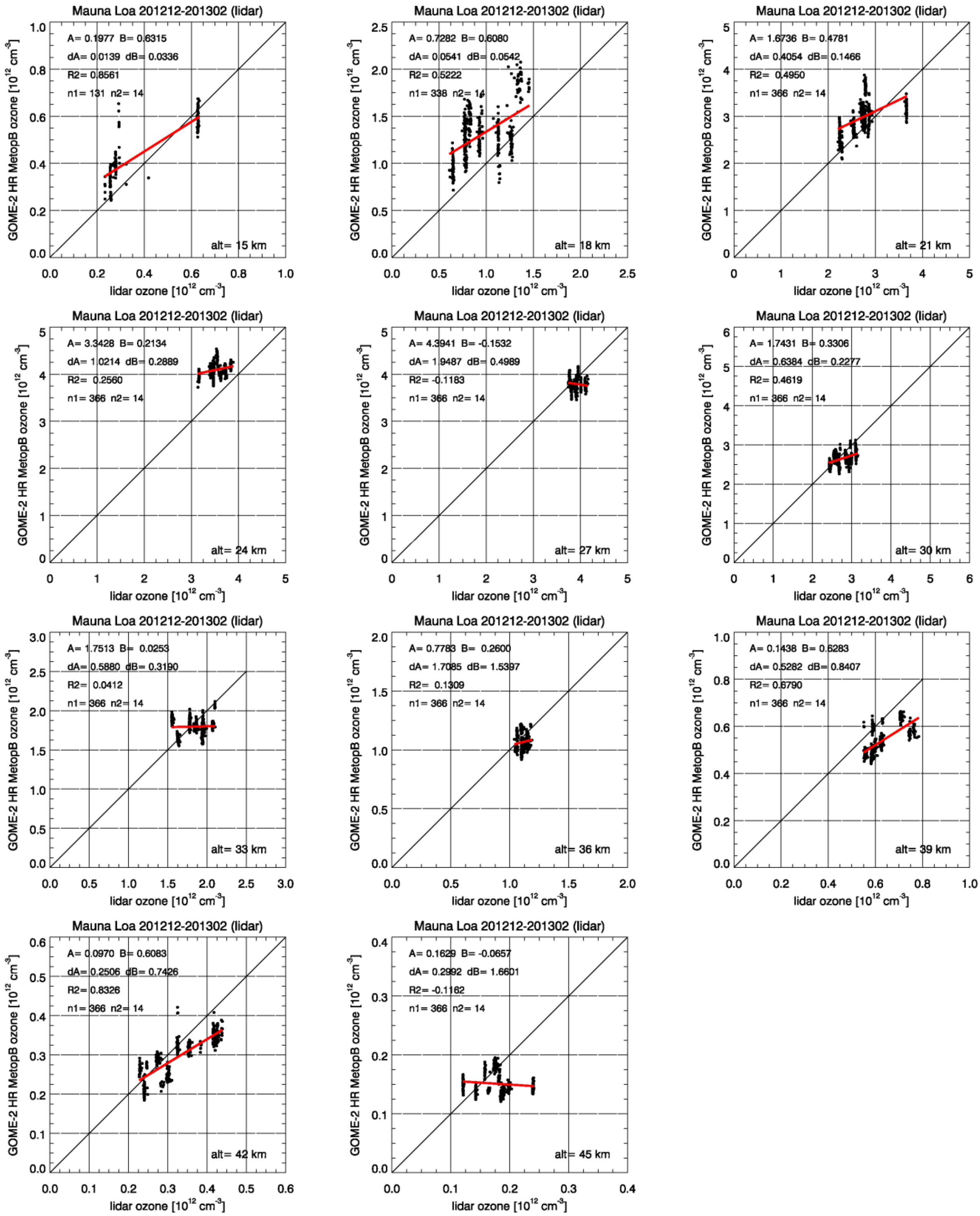


Fig 3.12: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC lidar at Mauna Loa.

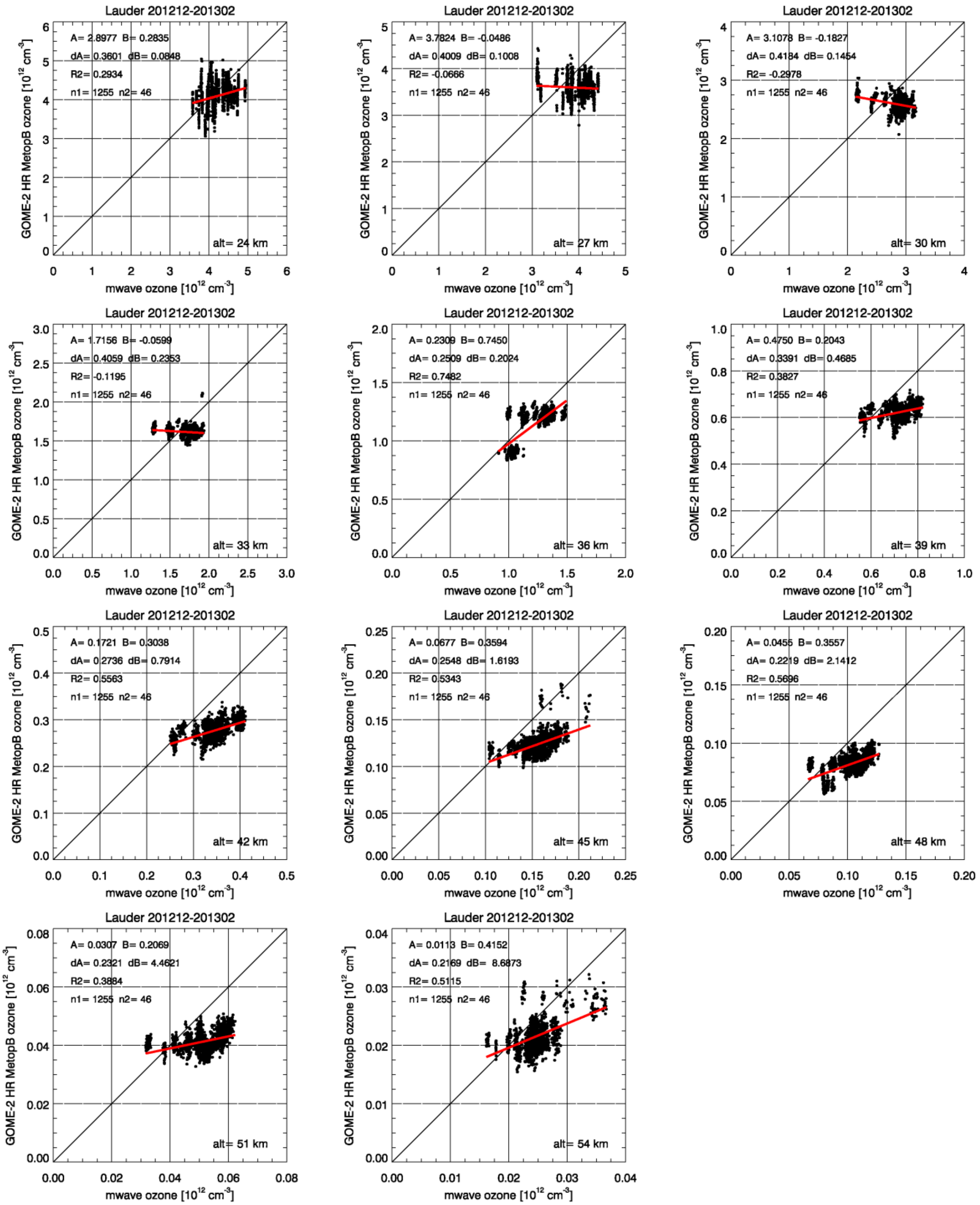


Fig 3.13: Same as previous Figure, but comparing GOME-2 MetopB high resolution with the NDACC microwave at Lauder.

3.9 GOME-2 MetopB coarse resolution ozone profiles.

Nearly all the results described so far are for the high resolution profile, where a ground pixel corresponds to about 80 x 40 km². For the coarse resolution profiles, ground pixels correspond to 640 x 40 km². Processing, meteorological and other input data are virtually identical for coarse and high resolution ozone profiles. The only major difference lies in cloud-recognition and clearing, which tends to be more precise for the high resolution profiles.

Not surprisingly, our validation found very similar results for coarse resolution and high resolution profiles. Bias and scatter are very similar. One difference is the much lower number of coarse resolution profiles. Since validation results for high resolution and coarse resolution profiles are so similar in the mid- and upper stratosphere, we will restrict ourselves to a few examples.

Fig. 3.14 shows individual and average comparison results for GOME-2 MetopB coarse resolution profiles and the lidar at Mauna Loa. Fig. 3.14 should be compared directly with Fig. 3.5, which shows the same thing, but for high resolution profiles. Apart from the smaller number of GOME-2 comparison profiles (dots), both Figures give very similar results, and almost the same difference profiles. This is also found for the other NDACC stations/ instruments.

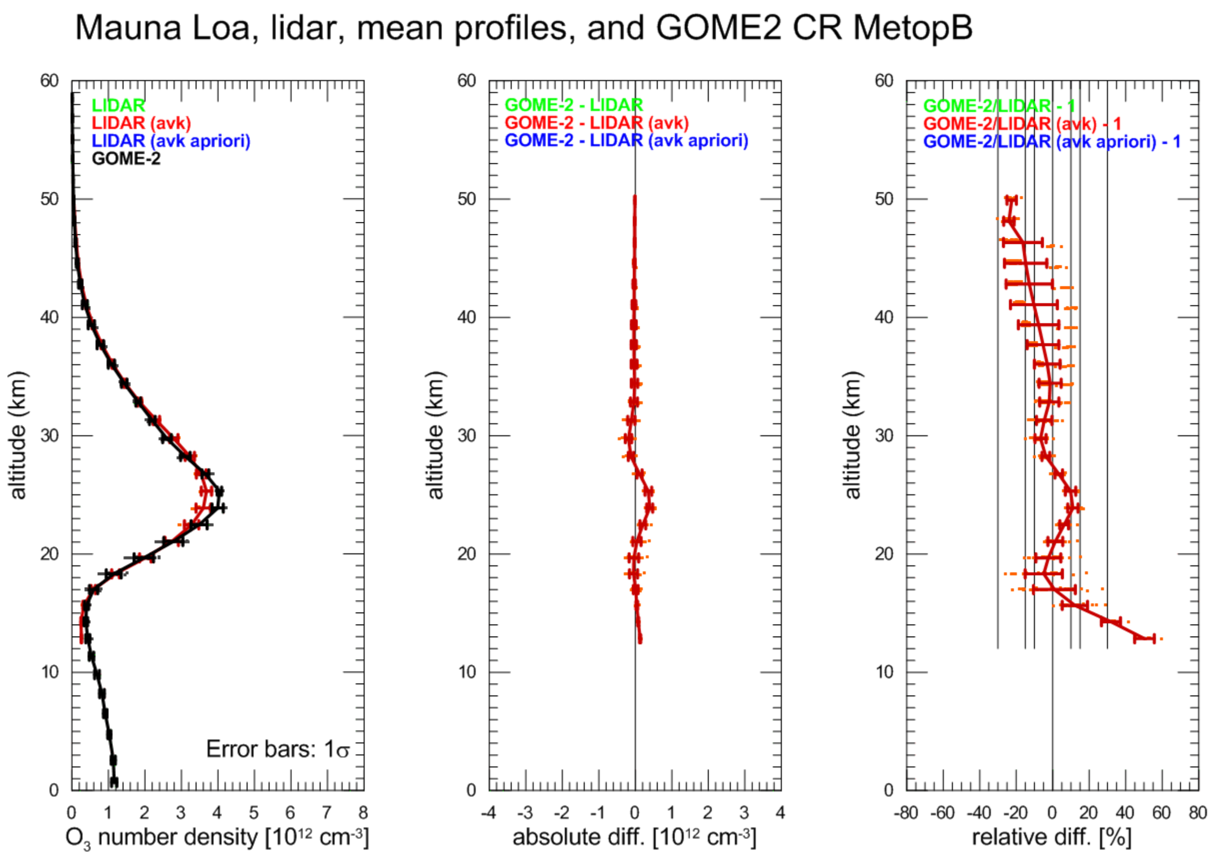


Fig 3.14: Same as Fig. 3.5, but comparing GOME-2 coarse resolution profiles and lidar profiles at Mauna Loa.

Exemplary scatter plots for the GOME-2 MetopB coarse resolution profiles are shown in Figs. 3.15 and 3.16. These can be compared directly with the corresponding Figs. 3.9 and 3.13 for the high resolution profiles. Apart from the smaller number of data points, overall impression and more importantly the red fit lines are very similar for coarse resolution and high resolution profiles.

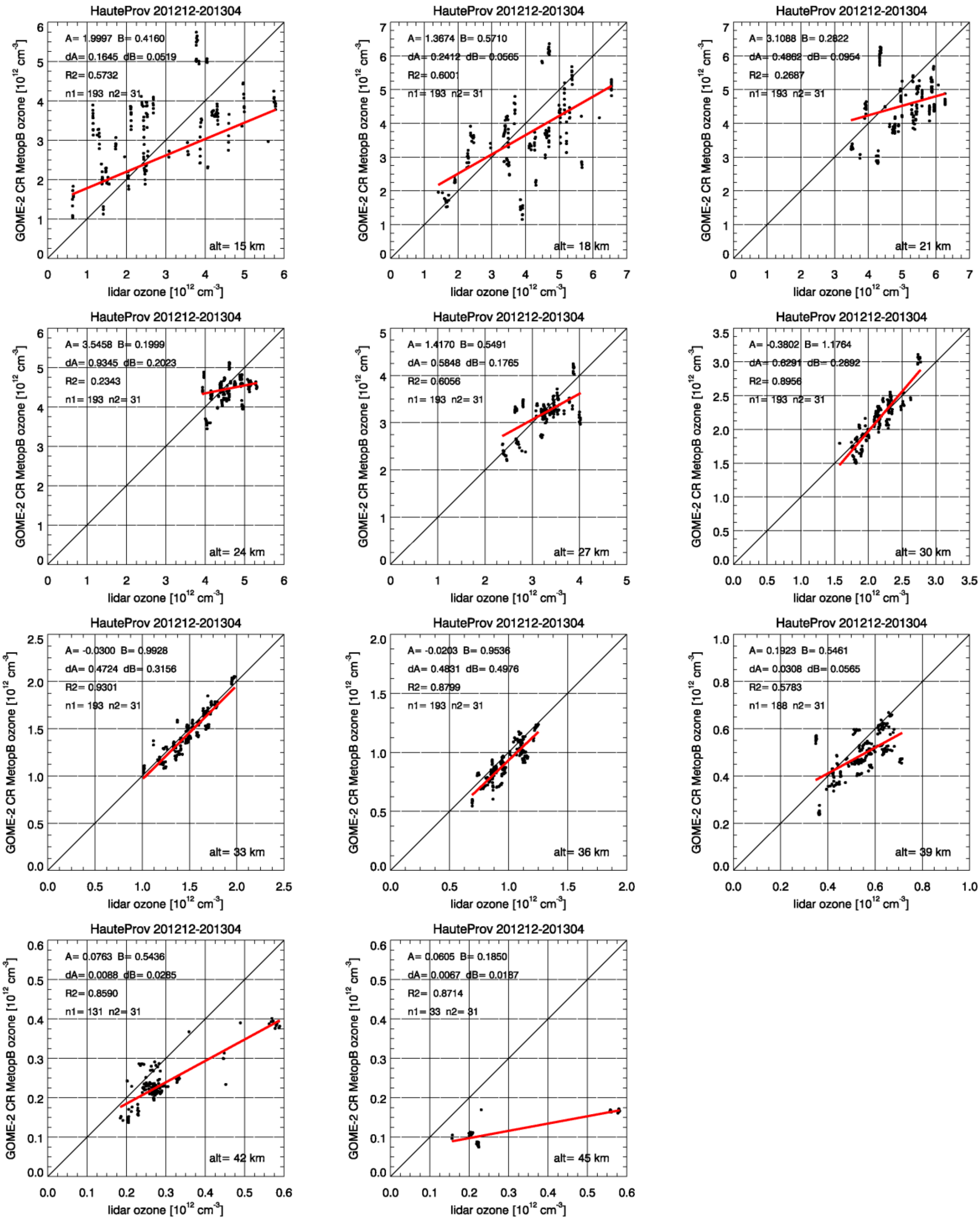


Fig 3.15: Same as Fig. 3.9, but comparing GOME-2 coarse resolution profiles versus lidar profiles at Haute Provence.

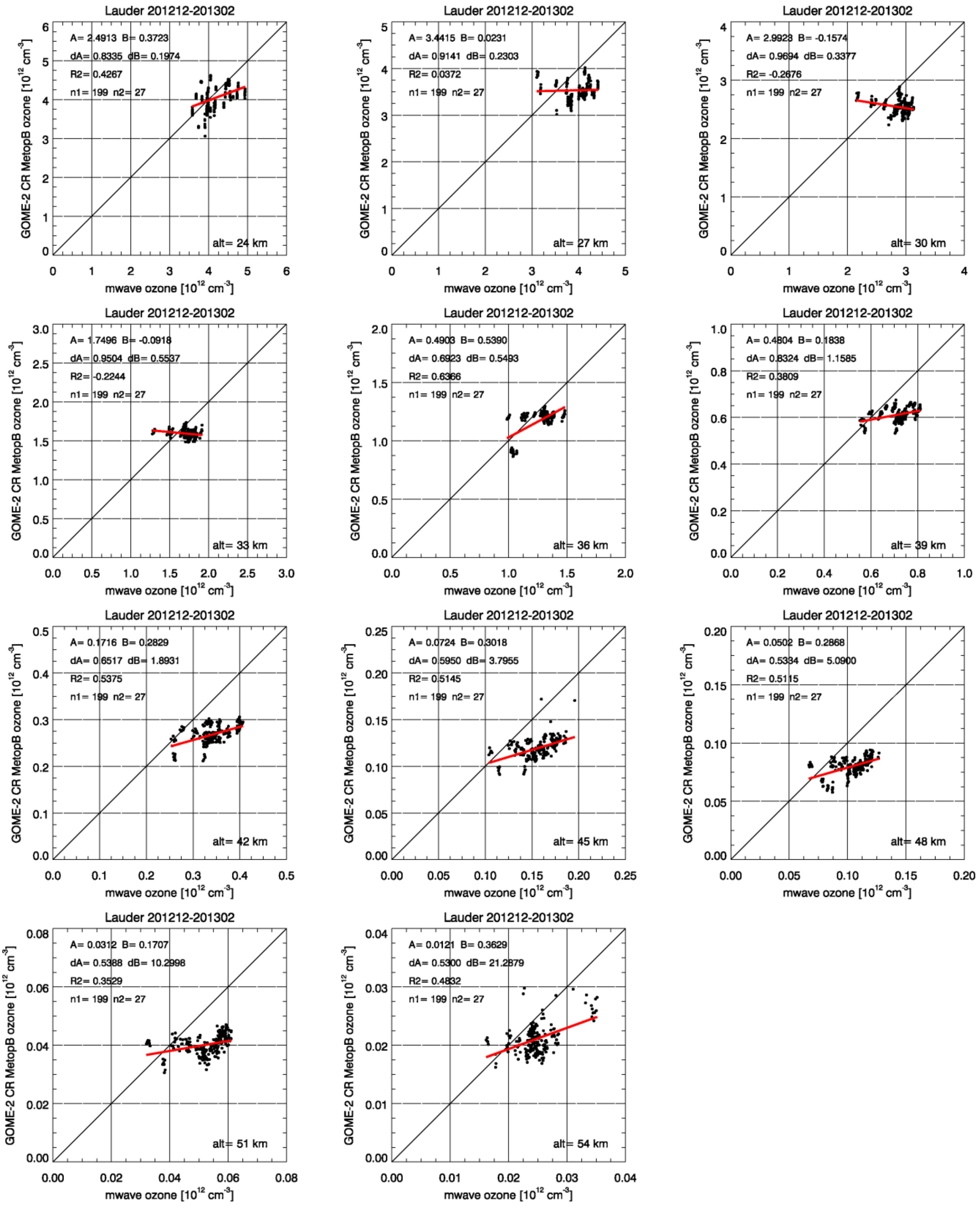


Fig 3.16: Same as Fig. 3.13, but comparing GOME-2 coarse resolution profiles versus microwave profiles at Lauder.

3.10 Comparing GOME-2 on MetopB with GOME-2 on MetopA

An important question for this validation report is whether GOME-2 on MetopB is ready to take over from the aging GOME-2 instrument on MetopA, in orbit since 2007. Overall, our results indicate that GOME-2 MetopB ozone profiles are as good as GOME-2 MetopA ozone profiles at the beginning of the lifetime of that instrument. Processing is much more mature now (and better than for GOME-2 in 2008 or 2009), since many improvements have been made over the life of MetopA. This now provides a very good basis for MetopB processing.
Over the years, GOME-2 MetopA ozone profiles have developed a significant low bias compared to the ground-based data. Also, GOME-2 MetopA data have become noisier over the years due to instrument degradation. Both the larger bias and the larger noise of GOME-2 MetopA can be seen in Fig. 3.17, which can be compared directly with Fig. 3.8, giving the same results for GOME-2 MetopB.

Signal degradation, and relatively larger noise of the GOME-2 MetopA data over the years also make it harder to derive good profile information from the spectra. This can be seen in Fig. 3.18, which compares (for a typical day, in the upper panels) the number of iterations in the OPERA inversion processor between MetopA and MetopB. More than 8 iterations usually mean that convergence was not achieved, and a meaningful ozone profile could not be retrieved. Clearly, MetopA has many more pixels, where no ozone profile could be retrieved (and N iter was greater than 8). The lower panels of Fig. 3.18 compare the DFS indicators (degrees of freedom for signal) between GOME-2 on MetopA and MetopB. Higher values of DFS mean that more detailed profile information could be retrieved (see Algorithm Theoretical Basis Document, Eq. 12, van der Oss et al., 2013). The DFS plots also show clearly that MetopB currently provides better profile information than MetopA.

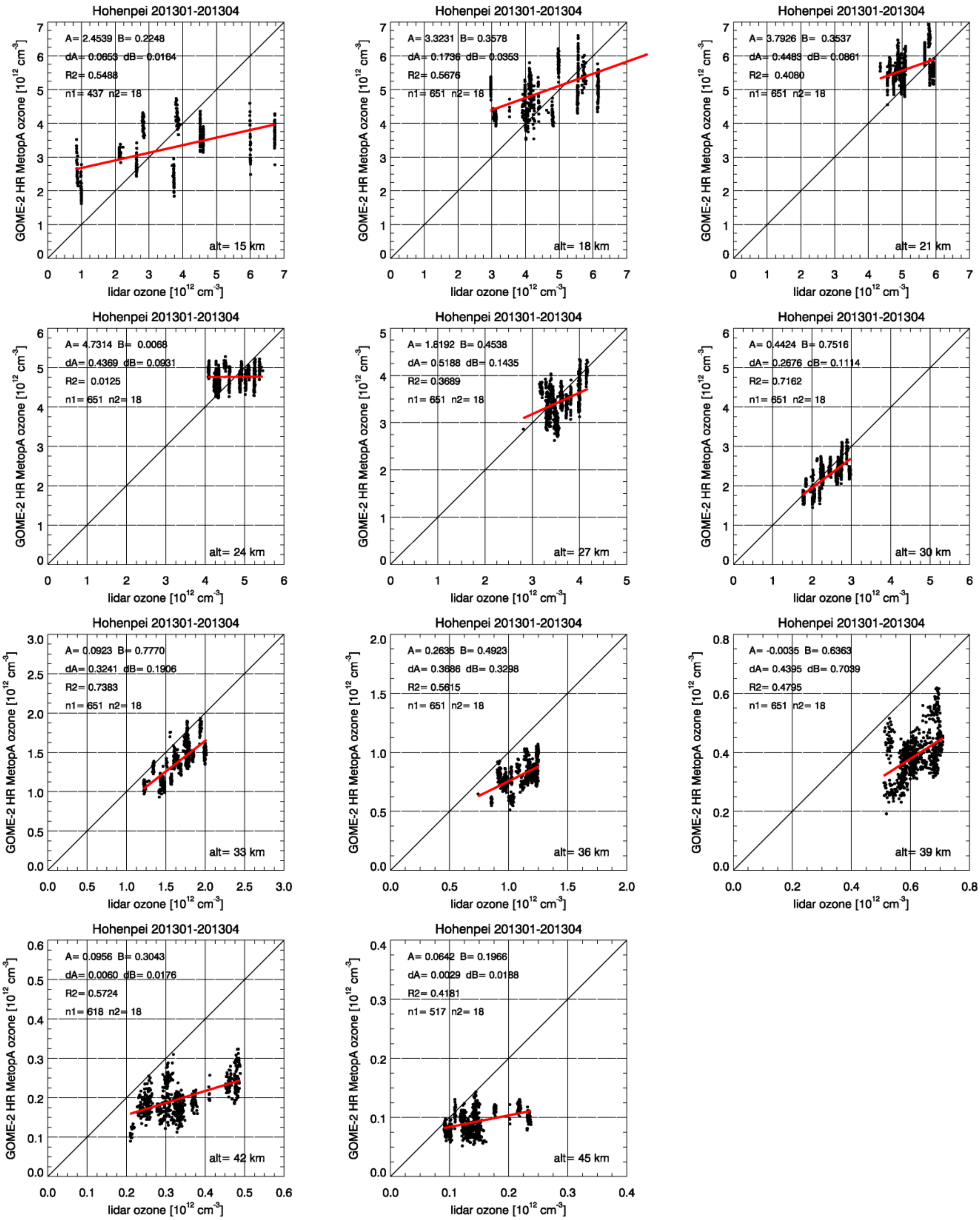


Fig 3.17: Same as Figure 3.8, but comparing GOME-2 MetopA high resolution with the NDACC lidar at Hohenpeissenberg

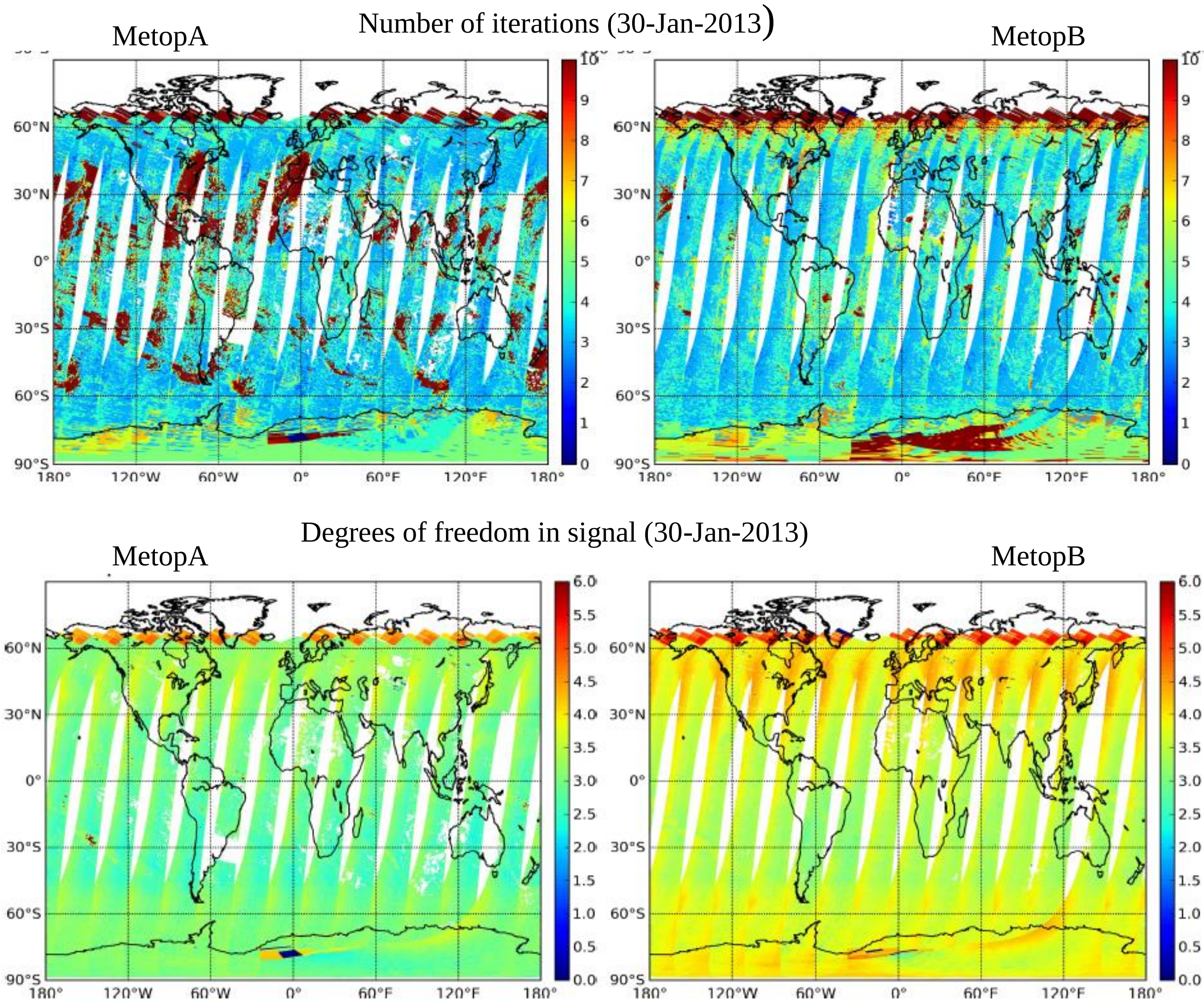


Fig 3.18: Maps comparing processor performance and retrieved profile information content between GOME-2 on MetopA (left panel) and MetopB (right panels). Top panels: Number of iterations before convergence. More than 7 iterations (red patches) usually mean that a successful retrieval was not possible. Bottom panels: DFS indicator (degrees of freedom for signal) giving the number of independent profile information pieces. Higher DFS usually means that a better profile could be retrieved.

4. Summary and conclusions

This report presents initial near global validation results for GOME-2 MetopB high resolution and low resolution ozone profiles for the relatively short time period from December 12th 2012 to April/ May 2013.

The *threshold* accuracies for the GOME-2 ozone profiles are 30% for the stratosphere and 70% for the troposphere. These targets are met for all the stations used in the analysis.

The *target* accuracies for the ozone profiles are respectively 30% in the troposphere and 15% in the stratosphere. This accuracy is reached in the stratosphere for most of the ozone profiles for the high- and coarse resolution pixels. For the troposphere, this accuracy is reached for the high resolution pixels and for most of the coarse resolution pixels.

Note that for these comparisons, the ground-based reference profiles have undergone substantial vertical smoothing. This is necessary to match the vertical resolution of the ground-based profiles to the vertical resolution of the GOME-2 profiles. The accuracies mentioned here refer to the smoothed reference profiles. In reality, ozone values in narrow layers may lie far away from the smoothed profiles, so true ozone values can be further away than statedby the accuracies here. These “smoothing errors”, both in the vertical and horizontal direction (due to pixel size and viewing geometry) can exceed 5 % in some cases, and should be kept in mind (see also Section 8 of the Algorithm Theoretical Basis Document, van Oss et al., 2013).

Comparisons with lidar and microwave radiometers reveal best results from the subtropics to the Mid-Latitude stations. The mean relative difference between GOME-2 and $X_{AVK\text{-}sonde}$ is in general smaller than ±15 % in the lower troposphere (for heights below 7 km) and in the stratosphere (for heights between 15 and 30 km, also compared to avk_lidar and avk_microwave). Between 30 km and 50 km an increasing negative bias is observed when GOME-2 is compared to lidar and microwave instruments. It has its maximum around 45 to 50 km with –15% ± 5% compared to vertically smoothed reference profiles (not shown) and average kernel smoothed reference profiles. There is little difference between high resolution and coarse resolution GOME-2 profiles. Both show similar bias and noise.

GOME-2 high- and coarse resolution ozone profiles give sensibly better results at Mid-Latitude stations than at the other latitude belts when compared to ozonesondes. Since it was not possible yet to validate for a full year period and because of well-known seasonal behaviour of the GOME-2 ozone profile product on METOP-A (Delcloo and Kins, 2009/2012), preliminary results shown in this report confirm that the GOME-2 METOP-B ozone profile products shows similar behaviour to the METOP-A high- and coarse resolution ozone profile products.

To summarize the statistics for the three different latitude belts included in this report (applying the averaging kernels) when compared with ozonesondes;

For the Northern Polar stations, the ozone concentrations in the troposphere are overestimated for the HR and CR ozone profile product respectively by about 17%, 18%, for the UTLS this is around 29%, 40% and for the stratosphere the underestimation equals -10% on average.
For the Northern Mid-Latititude stations, the ozone concentrations in the troposphere are overestimated for the HR and CR ozone profile product respectively by about 29%, 33%, for the UTLS this is around 49%, 64% and for the stratosphere the underestimation equals -5% on average.

For the Neumayer station, the ozone concentrations in the troposphere are overestimated for the HR and CR ozone profile product respectively by about 8%, 6%, for the UTLS this is around 2%, 0% and for the stratosphere the underestimation equals -10% on average.

The information content is lower in the troposphere, UTLS-zone and around the ozone maximum.

The algorithm has problems with the UTLS-zone and tends to overestimate the ozone in the lower stratosphere during winter- and early spring season and underestimates the ozone concentrations in the upper stratosphere. By using the high resolution pixels, the retrieved profiles improved significantly and all the target values are met for all stations when using the high resolution pixels.

From this initial validation we can conclude that the quality of the GOME-2 Metop-B ozone profile products is as good as GOME-2 MetopA ozone profile products for 2008 or 2009. The quality is better than the current GOME-2 Metop-A ozone profile products. In our opinion, Metop-B data are ready for release, and can take over from Metop-A for the monitoring of ozone profile changes.

5. Acknowledgement

The ozonesonde data was made available by WOUDC (http://www.woudc.org), the SHADOZ network (http://croc.gsfc.nasa.gov/shadoz/) and the NILU's Atmospheric Database for Interactive Retrieval (NADIR) at Norsk Institutt for Luftforskning (NILU) (http://www.nilu.no/nadir/). Lidar and microwave radiometer ozone profiles were retrieved from the Network for the Detection of Atmospheric Composition Change (NDACC) (http://www.ndsc.ncep.noaa.gov/) and the NILU Atmospheric Database for Interactive Retrieval (NADIR) at Norsk Institutt for Luftforskning (NILU).

6. References

- 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): Ozone SAF Validation reports
- Godin S., G. Megie, J. Pelon, Systematic Lidar Measurements of the Stratospheric Ozone vertical Distribution, Geophys. Res. Lett., 16, 547-550, 1989.

- Hocke, K.; N. Kämpfer, D. Ruffieux, L. Froidevaux, A. Parrish, I. Boyd, T. von Clarmann, T. Steck, Y.M. Timofeyev, A.V. Polyakov, E. Kyrölä, Comparison and synergy of stratospheric ozone measurements by satellite limb sounders and the ground-based microwave radiometer SOMORA, Atmos. Chem. Phys., 7, 4117-4131, 2007
- Keckhut, P.; I. S. McDermid, D. Swart, T. J. McGee, S. Pal, S. Godin-Beekmann, A. Adriani, J. Barnes, H. Bencherif, H. Claude, G. Fiocco, G. Hansen, A. Hauchecorne, T. Leblanc, C. H. 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 retreived 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.
- 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., J.C. Witte, R.D. McPeters, S.J. Oltmans, F.J. Schmidlin, J.A. Logan, M.Fujiwara, V.W.J.H. Kirchhoff, F. Posny, G.J.R. Coetzee, B. Hoegger, S. Kawakami, T. Ogawa, B.J. Johnson, H. Vömel and G. Labow, Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2000 Tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements, *J. Geophys. Res.*, Vol. 108 No. D2, 8238, doi: 10.1029/2001JD000967, 30 January 2003.
- Thompson, A.M., J.C. Witte, S.J. Oltmans, F.J. Schmidlin, J.A. Logan, M. Fujiwara, V.W.J.H. Kirchhoff, F. Posny, G.J.R. Coetzee, B. Hoegger, S. Kawakami, T. Ogawa, J.P.F. Fortuin, and H.M. Kelder, Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2000 Tropical ozone climatology 2. Tropospheric variability and the zonal wave-one, *J. Geophys. Res.*, Vol. 108 No. D2,8241, doi: 10.1029/2002JD002241, 31 January 2003.
- van Oss, R., J. De, O. Tuinder, Algorithm Theoretical Basis Document for the GOME-2 Near-Real-Time and Offline Ozone Profile Products, EUMETSAT Ozone SAF, 2013 available from http://o3msaf.fmi.fi/documents.html