

# **GOME-2 on MetOp-A Support for Analysis of GOME-2 In-Orbit Degradation and Impacts on Level 2 Data Products**

## **Final Report**

**Prepared by:**

S. Dikty and A. Richter

With contributions from:

M. Weber, S. Noël, H. Bovensmann, F. Wittrock, and J. P. Burrows



**University of Bremen**  
Institute of Remote Sensing

Institute of Environmental Physics  
Institute of Remote Sensing  
University of Bremen FB1  
P.O. Box 33 04 40  
D-28334 Bremen

Fon: +49-421-218-62079

Fax: +49-421-218-4555

E-Mail: [Sebastian.Dikty@iup.physik.uni-bremen.de](mailto:Sebastian.Dikty@iup.physik.uni-bremen.de)

**Document:** Final Report

**Issue:** Version 2.0 - Final

**Issue Date:** 18.01.2012

## Table of Contents

1	Introduction.....	3
2	Study setup (WP1) .....	3
2.1	Degradation and accuracy of level 2 products.....	3
2.2	Degradation and precision of lv2 products .....	4
2.3	Relevance of throughput loss and systematic spectral structures .....	4
2.4	2nd throughput test.....	5
3	Refinement in input data sets (WP2).....	5
3.1	Data acquired .....	5
3.2	Hardware installed.....	7
3.3	Software tools developed .....	7
4	Data analysis (WP3).....	8
4.1	Time series .....	8
4.1.1	Values shown and selection of geo-locations .....	8
4.1.2	GOME-2 vs. SCIAMACHY time series.....	9
4.2	Intensity dependencies.....	20
4.2.1	Based on daily averages.....	20
4.2.2	Based on all data .....	21
4.2.3	Solar vs. earthshine spectra.....	23
4.2.4	SO <sub>2</sub> – a separate challenge .....	25
4.3	Wavelength dependency of retrieval residuals .....	25
4.4	Line of sight dependency.....	27
4.5	How DOAS retrievals are affected by degradation .....	30
4.5.1	Loss of signal .....	30
4.5.2	Instrument changes not accounted for by calibration.....	30
4.5.3	Differential changes between earth-shine and irradiance .....	31
4.6	Changes in GOME-2 solar spectra.....	32
5	Throughput test (WP 4) .....	36
5.1	Level 0 data.....	36
5.2	Level 2 data.....	39
5.3	Changes in degradation rate after 2 <sup>nd</sup> throughput test.....	42
5.4	Summary throughput test .....	42
6	Summary and Conclusions.....	43

# 1 Introduction

The overall aim of the study was to provide input on the following questions:

1. What is the effect of GOME-2 degradation on the accuracy (absolute values) of level 2 products?
2. What is the effect of GOME-2 degradation on the precision (scatter) of level 2 products?
3. Is the degradation dominated by throughput loss or are there also systematic spectral structures linked to instrument changes or degradation related calibration deficiencies?
4. Are there possibilities to correct for degradation effects on GOME-2 level 2 products?
5. What happened with GOME-2 level 2 products during the 2<sup>nd</sup> throughput test, and what can we learn from these results?

To tackle these questions, different approaches have been selected and have been tested on selected data sets.

The study was divided into several work packages (WP):

During WP1 details of the study setup were defined, input from the individual product suppliers was gathered and preliminary results on selected data sets were derived. The focus of WP2 lay on the acquisition of data sets for analysis done in WP3, the installation of hardware and the preparation of software tools needed as outlined in the statement of work. WP4 dealt with the results from the 2<sup>nd</sup> throughput test in September 2009, while during WP5 all results were updated and prepared for final presentation.

## 2 Study setup (WP1)

Based on the outline given in the project proposal, a number of different approaches have been selected to address the main questions of the study. They will be listed below ordered by questions to answer.

### *2.1 Degradation and accuracy of level 2 products*

Probably the most important issue for data users is the accuracy of the long-term record of GOME-2 data. This could be affected by degradation, and the approach taken here was to investigate the full four and a half year time series (January 2007 to August 2011) of level 2 data products by

1. comparing GOME-2 and SCIAMACHY level 2 time series over selected regions to assess the overall consistency and possible changes with time
2. investigating GOME-2 time series over selected regions thought to have little or no variability in atmospheric concentrations of trace gases (Pacific, Sahara, combined Greenland/Antarctic)
3. repeating the analysis on subsets of data separated by parameters such as viewing conditions, intensity or geographical region

The output of this analysis is an assessment of how strongly the accuracy of the level 2 products is affected by degradation and how it depends on measurement parameters.

## ***2.2 Degradation and precision of lv2 products***

Experience with GOME-2 data shows that the scatter of level 2 products has increased throughout the lifetime of the instrument. This is to be expected from loss of throughput but could also be affected by instrumental changes or issues with calibration resulting from degradation of the instrument. To investigate the loss in precision through degradation, similar approaches as discussed for the accuracy were used, namely

1. analysis of time series over regions where homogeneous trace gas columns are expected and the scatter in values can be attributed to measurement precision
2. comparison to similar time series from SCIAMACHY using both data from the same time period and measurements from the same instrument age
3. evaluation of the effects of limiting the data sets to certain subsets (viewing directions, regions, ...)

This analysis was performed for most level 2 products and the full GOME-2 data set. Resulting figures have been updated on a monthly basis throughout the project.

## ***2.3 Relevance of throughput loss and systematic spectral structures***

While quantifying the effects of degradation on level 2 products is important, it would be preferable to correct data products as much as possible to minimise the impact on data users. A necessary prerequisite for such a correction attempt is the identification of the main reasons for degradation effects. This is not necessarily linked to physical models of degradation in the instrument as empirical corrections are often effectively used, but it is expected that the analysis of the data could also provide hints on instrument behaviour.

From a spectral analysis point of view, two effects of degradation have to be separated; the impact of throughput loss which cannot be reduced and the evolution of spectral artefacts in the level 1 data from instrument changes or imperfect calibration. The latter could in theory be corrected for, either empirically or by applying time dependent instrument calibration.

The main point of this task was to identify and remove the expected effects of throughput loss. This was done by

1. analysing data scatter in “constant” regions selected before as a function of intensity which will highlight the importance of throughput effects
2. separating viewing angle and regional affects normalised by the intensity of the respective measurements
3. systematically evaluating averaged residuals for selected regions, time periods and viewing conditions to investigate the possible development of spectral structures
4. investigating the impact of different calibration settings on the residuals (depending on the outcome of the previous step)

While the first two steps have been performed on most of the data, the third step had to be limited to selected data as a result of processing and data storage limitations. After evaluating the results from the previous steps, step four has been discarded as at this point, calibration problems do not appear to be the main cause of the observed degradation of lv2 product quality.

## **2.4 2nd throughput test**

During the throughput test, the temperature of the GOME-2 instrument was changed in a controlled way while normal measurements were taken. This enabled a test of how the observed degradation depends on instrument temperature and whether or not there is a recovery in performance at any point during the heating.

In order to analyse this and also the effect of additional degradation introduced by the test, the following tests were performed:

1. analysis of throughput data for most level 2 products produced at the University of Bremen
2. comparison of accuracy and precision of the data during 2<sup>nd</sup> throughput test with those taken before and after
3. comparison of trace gas columns with co-located SCIAMACHY data
4. comparison of accuracy and precision of data products taken before and after the throughput test (if necessary corrected for seasonal effects)
5. analysis of residuals obtained during 2<sup>nd</sup> throughput test

The results help to better understand instrument performance during the 2<sup>nd</sup> throughput test.

## **3 Refinement in input data sets (WP2)**

### **3.1 Data acquired**

As a first step done already in WP1, the data sets needed were identified and the status of their availability checked. The results were updated during the whole project and are listed in Table 1. The according spectral ranges of the fitting windows are listed in Table 2.

A consistent set of GOME-2 level 1B data is stored at the University Bremen based on the reprocessed data set until end of June 2008 and the NRT data afterwards. Data for the 2<sup>nd</sup> throughput test (TT2) have been obtained from EUMETSAT and stored on a new crunch-RAID-system (GOMZO).

A level 1A data set covering the period 01.01.2007 – 31.03.2011 has been ordered with UMARF and was delivered to University of Bremen. These data have been also stored on GOMZO. Data after 31.03.2011 will be ordered on a 3-month interval.

Most of the GOME-2 level 2 data sets are available over the complete GOME-2 time period. These are scientific data products produced at the University of Bremen as for these data products full control over input parameters and analysis settings is given.

For comparisons SCIAMACHY level 2 data products are also available from August 2002 to today. An inconsistency in some of the SCIAMACHY data sets has been corrected for by reprocessing.

**Table 1: Overview on data needed in the study and the status of their availability**

Product	Time period	Status
GOME-2 lv1A data	01.01.2007 – 31.03.2011	Available (data thereafter can be ordered via UMARF)
GOME-2 lv1A and lv1B data	2 <sup>nd</sup> throughput test	Available
GOME-2 lv1B data (latest PPF version)	01.01.2007 - today	Available
GOME-2 NO <sub>2</sub> time series	01.01.2007 – today	Available in two versions using different fitting windows
GOME-2 BrO time series	01.01.2007 – today	Available
GOME-2 HCHO time series	01.01.2007 – today	Available
GOME-2 CHOCHO time series	01.01.2007 – today	Available
GOME-2 O <sub>3</sub> time series	01.01.2007 – today	Available
GOME-2 H <sub>2</sub> O time series	01.01.2007 – today	Available
SCIAMACHY NO <sub>2</sub> time series	01.01.2003 - today	Available
SCIAMACHY BrO time series	01.01.2003 - today	Available
SCIAMACHY HCHO time series	01.01.2003 - today	Available
SCIAMACHY CHOCHO time series	01.01.2003 - today	Available
SCIAMACHY O <sub>3</sub> time series	01.01.2003 - today	Available
SCIAMACHY H <sub>2</sub> O time series	01.01.2003 - today	Available

**Table 2: Spectral ranges of fitting windows**

Data product	Spectral region used
SO <sub>2</sub>	312.5 – 317.0 nm
O <sub>3</sub>	326.0 – 335.0 nm
BrO	336.0 – 347.0 nm
HCHO	337.0 – 353.0 nm
OCIO	365.0 – 389.0 nm
CHOCHO	424.0 – 457.0 nm
NO <sub>2</sub>	425.0 – 450.0 nm 425.0 – 497.0 nm
H <sub>2</sub> O	688.0 – 700.0 nm

### *3.2 Hardware installed*

After an analysis of the computational needs, appropriate hardware was selected and installed at IUP for data processing and storage. One combined number crunching - RAID cluster consisting of 24 CPUs and a directly connected (via Bus architecture) RAID with 48 TB of disc space, 32 GB of memory and a Unix operating system serves as platform for the level 1A/1B extractor (provided by EUMETSAT) and the various retrieval algorithms presently available at IUP. The cluster will be referred to as "GOMZO" in the following. A tape recorder to read the LTO-4 tapes (delivered by EUMETSAT) and to backup data from different retrieval/calibration runs is also part of the setup. In addition, a PC was procured which acts both as terminal for GOMZO and as personal computer for the project coordination.

### *3.3 Software tools developed*

Most of the software tools needed for the project had already been developed at the University of Bremen in earlier projects. However, due to the research nature of most of the projects, neither the tools nor the output formats are homogeneous between products and analysis tools are heterogeneous with respect to operating systems, programming languages and interfaces used. IDL based software has been developed during WP1 part to

1. read all level 2 data products produced needed,
2. perform basic statistical analysis on the data for different regions separating the data sets according to their SZA, VZA and other geophysical parameters,
3. create graphical output of the results.

The tools have continuously been updated and improved during the project. They have been validated against existing tools and were applied to existing level 2 data sets. Throughout the project, this software was extended as needed. Auxiliary information on the usage of the software will be made available if requested by Eumetsat as decided during progress meeting 1 (PM1) in May 2010.

In addition, the level 1A/1B extractor was installed and runs without unanticipated difficulties. The EPS GOME-2 level 1 Product Generation Specification served as reference for individual calibration steps that could be turned on/off according to the needs of the investigation. The speed of level 1A/1B calibration was tested several times. When parallelized, calibrating one year of GOME-2 level 1A data takes no more than 48 hours. Extracting the level 1B files takes app. another 48 hours per year of data, but unlike the calibration has only to be done once. A shell script has been written to simplify the handling of the level 1A/1B extractor and data management. This script can be extended as needed, possibly to quasi-parallelize the calibration process by distributing and queuing jobs on an orbit-to-orbit basis, minimizing calibration times.

In addition to the level 1A/1B calibration step, retrieval algorithms for various level 2 products can be run on GOMZO to complete the level 1A to level 2 data processing on a single cluster of CPUs. The individual retrieval algorithms are faster than calibration. Depending on the trace gas, i.e. fit window size, one year of calibrated and extracted data can be processed within 14-20 hours. Altogether, from level 1A to level 2 would take app. 4-5 days per one complete year of GOME-2 data.

## 4 Data analysis (WP3)

The focus of the activities during WP3 was substantial data analysis on selected data sets. In addition, GOME-2 data were compared to SCIAMACHY data, covering the same period on the one hand and the first mission years on the other. Results will be presented in the following.

### 4.1 Time series

#### 4.1.1 Values shown and selection of geo-locations

Operational IUP level 2 products (BrO, NO<sub>2</sub>, HCHO, O<sub>3</sub>, and H<sub>2</sub>O) were checked for degradation signals in the time series. In the following, an emphasis is put on BrO since it is an important GOME-2 product and shows clear degradation effects. The quantity of interest is the vertical column (VC) which is deduced from the retrieved slant column (SC) through division by appropriate air mass factors (AMF). The VC has also been cosine corrected for line of sight effects. The other main values going along with the VC are the root mean square (RMS) giving basically the spread of the VCs within a given box of geo-coordinates, the earthshine fit window intensity (INT; only for SO<sub>2</sub>, BrO, NO<sub>2</sub>, and HCHO), and the spread of the retrieval residuals, Chi-square ( $\chi^2$  or ChiSq).

**Table 3:** Overview of  $\chi^2$  as computed for GOME-2 level 2 products

Trace gas	Shown as
SO <sub>2</sub> , OCIO, CHOCHO, BrO, NO <sub>2</sub> , and HCHO	$\chi^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu_i)^2$ #
O <sub>3</sub>	$\chi^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_i)^2$
H <sub>2</sub> O	$\begin{aligned} \text{Error} &= \frac{\sigma_K \chi}{\sqrt{n_d - n_p}} \\ &= \frac{\sqrt{\text{cov}(p_K, p_K)}}{\sqrt{n_d - n_p}} \sqrt{\sum_{i=1}^n (X_i - \mu_i)^2} \end{aligned}$

Table 3 shows an overview of how the residuals are being computed within the specific trace gas retrieval.  $X_j$  is the fitted optical depth and  $\mu_i$  is the measured optical depth.  $\sigma_i$  is the standard deviation of the fitted parameters, and  $n_d$  and  $n_p$  are the degrees of freedom. For all following results the *Error* of H<sub>2</sub>O has been also squared to comply with the other trace gases.

GOME-2 and SCIAMACHY time series from January 2007 to July 2011 have been plotted for geo-locations where natural or anthropogenic variations are minimum. In the following, two regions were selected, one over the Southern Pacific Ocean (25°S-15°S and 150°W-110°W) and one over the Sahara Desert (20°N-30°N and 0°E-30°E). In addition, an additional box consisting of two geo-locations has been added. This “Ice-and-Snow-Box” is a combination of Greenland (70°N-75°N, 50°W-30°W) and Antarctic (70°S-75°S, 130°E-150°E) values. The locations of all the boxes are shown in Fig. 1.

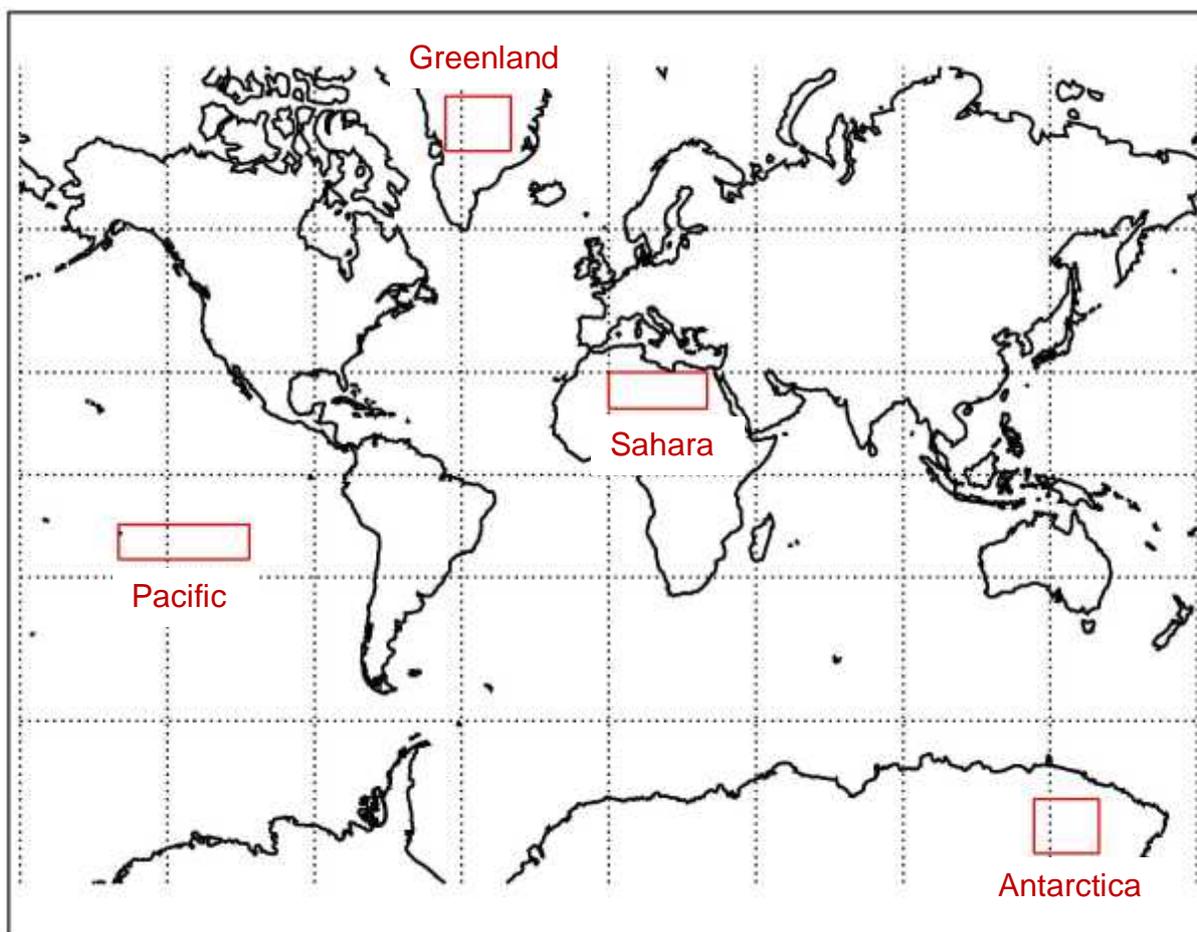
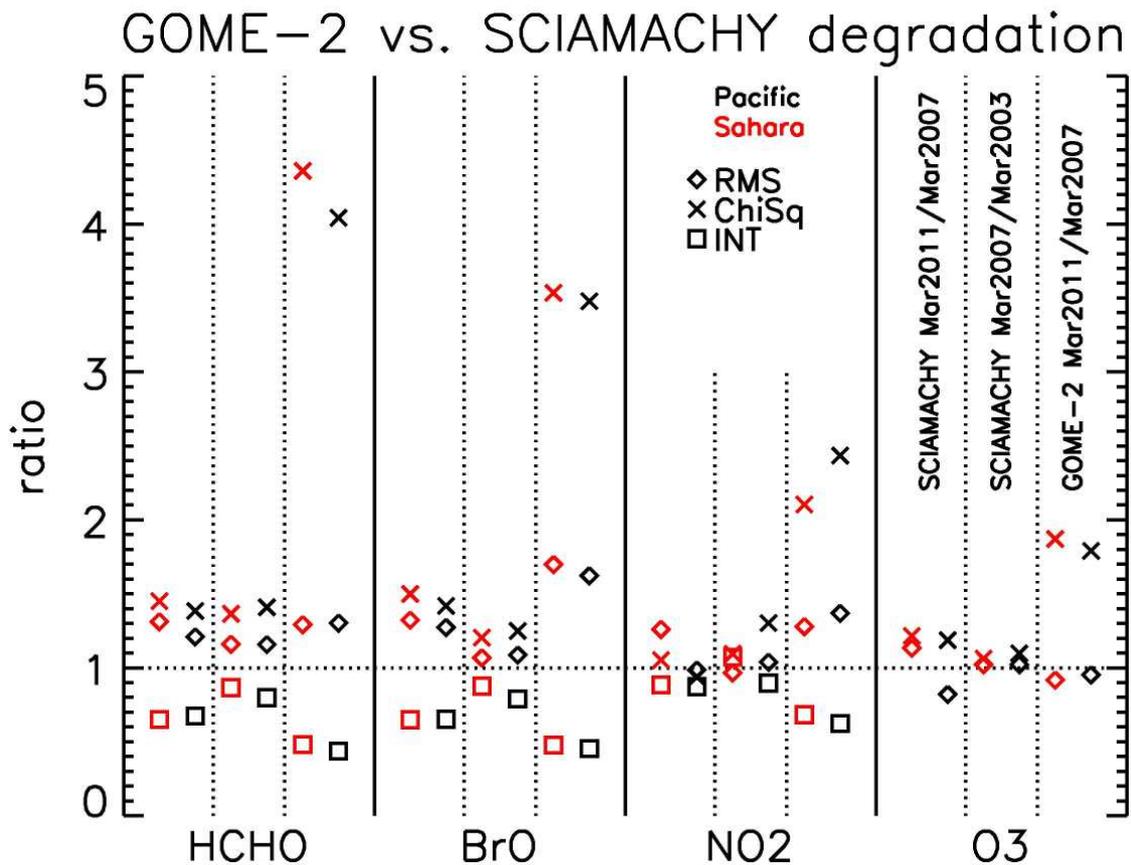


Fig. 1: Boxes of geo-locations chosen for the investigations.

#### 4.1.2 GOME-2 vs. SCIAMACHY time series

A summary of GOME-2 and SCIAMACHY degradation in level 2 data sets is shown in Fig. 2. The following scientific level 2 products from IUP Bremen are depicted: BrO, HCHO, NO<sub>2</sub>, and O<sub>3</sub>, over the boxes of geo-locations Sahara and Pacific. Apart from BrO (cf. Figs. x.7 and x.8) there are no signs of changes in the vertical columns. Therefore the presentation of results was restricted to the scatter of vertical columns, RMS, the retrieval residuals,  $\chi^2$ , and the fit window intensity, INT.



**Fig. 2:** Summary of GOME-2 and SCIAMACHY degradation separated into the species BrO, HCHO, NO<sub>2</sub>, and O<sub>3</sub>. Shown are the ratios of monthly mean values of March 2011 w.r.t. March 2007 and March 2007 w.r.t. March 2003 (only SCIAMACHY) for the scatter of data within a box of geo-locations (RMS), the retrieval residuals  $\chi^2$  (ChiSq), and fit window intensity (INT). The boxes of geo-locations are highlighted in red (Sahara), and black (Pacific).

For technical reasons, there is no information on the fit window intensity available for O<sub>3</sub>. For BrO, and HCHO, having been retrieved with the help of information in the UV region of the spectrum, the loss of intensity (throughput) is in the order of 50 to 70 %. As for NO<sub>2</sub> with its fitting window already in the visible part of the spectrum between 425 and 450 nm (channel 3), the loss of intensity ranges around 20 to 30 %.

When it comes to the scatter of values for the vertical columns within a given box of geo-locations, O<sub>3</sub> shows no or little change, whereas the other species show signs of loss in the precision of measurements. Again, NO<sub>2</sub> being a little better off than BrO, and HCHO. On average the precision worsened by 30 to 80 %. As to why O<sub>3</sub> actually shows less scatter in the most recent data than at the beginning of GOME-2 operations there is currently no good explanation.

The most important values to look at are the ones for the quality of the fits, the retrieval residuals  $\chi^2$ . H<sub>2</sub>O (not shown here) with its fitting window at the far end of the visible near the infrared (channel 4) and because of its character as a strong absorber (strong signal in the spectra) shows no signs of degradation in  $\chi^2$ . Also a strong absorber but being retrieved in the UV part of the spectrum (channel 2), where GOME-2 has almost its largest loss of throughput, O<sub>3</sub> shows stronger signs of degradation. The residuals are in the order of 80 to 100 % larger than they used to be in the beginning of the mission in 2007. An increase in the same order of magnitude is visible for NO<sub>2</sub>, although being a weaker absorber, but having the

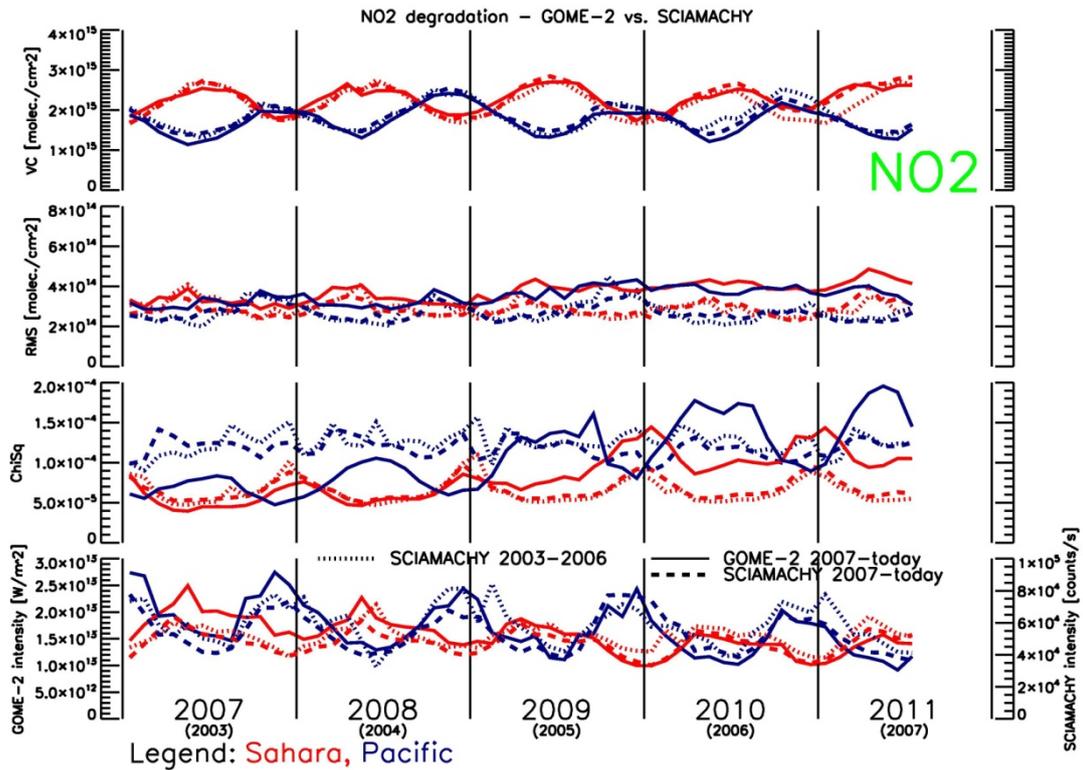
advantage of being retrieved in the visible with more light reaching the sensor. BrO and HCHO, being retrieved in the UV and being even weaker absorbers than NO<sub>2</sub> show the largest increase in retrieval residuals and thus the largest loss of accuracy. The increase ranges between 350 and 450 %. This cannot be explained alone with the loss of throughput (although it's the largest contribution to the degradation), but additional systematic components must contribute to this high level of  $\chi^2$  increase.

For comparisons with SCIAMACHY performance, data have been added to Fig. 2, which shows the degradation summary for the period 2007-2011 and 2003-2007, respectively.

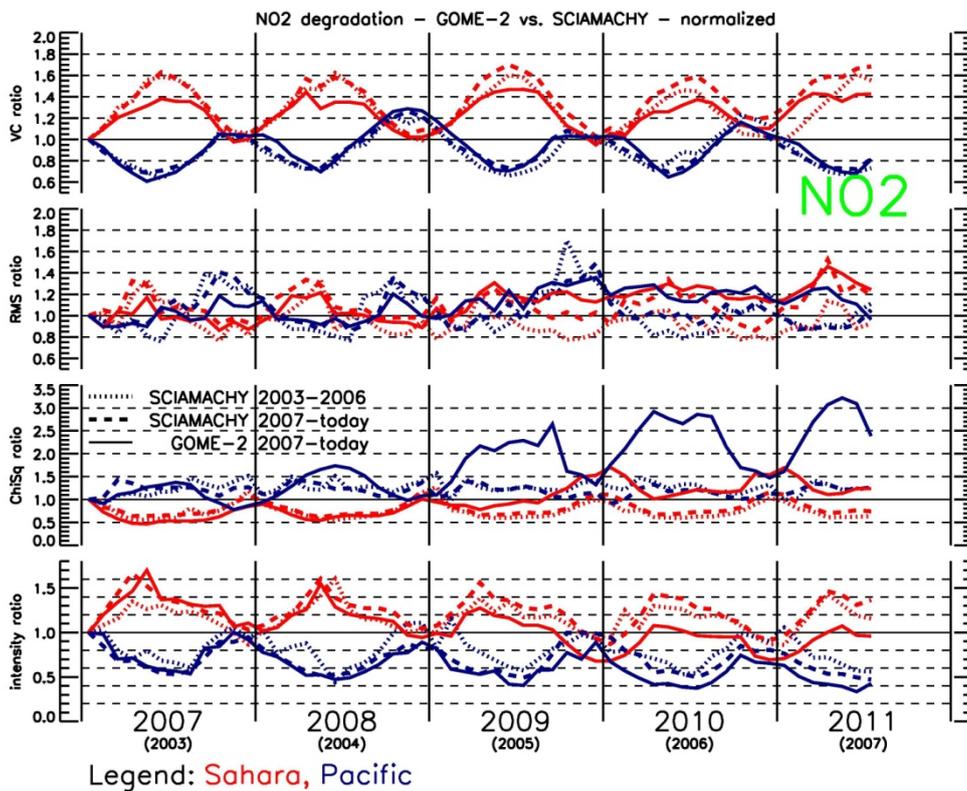
All individual GOME-2 time series (January 2007 to July 2011) are shown in Figs. 3 to 17. The upper plots show the vertical column (VC) time series, in the middle the scatter of VCs within a box of geo-locations (RMS) is depicted, and below are the time series for retrieval residuals  $\chi^2$  and fit window intensity (INT).

As WP3 puts an emphasis on the comparisons of GOME-2 performance with its sister instrument SCIAMACHY, Figs. 3 to 17 also include time series of SCIAMACHY. With dotted lines the period from January 2003 to July 2007 is depicted and with dashed lines the period from January 2007 to July 2011 is covered. The latter was chosen to compare instrument performance in the same time frame and the earlier was chosen to compare the instruments performance within their respective first years of operation.

There are 3 sets of plots for each trace gas investigated. One shows the time series in absolute values over the Sahara desert (red) and the Pacific Ocean (dark blue), the next set of plots shows the same time series (same colour code) normalized to January 2007 (March 2007 for H<sub>2</sub>O, no level 2 data available for January and February 2007, reprocessing planned with version 5.0 level 1B data after final report) for a more direct comparison between GOME-2 and SCIAMACHY. The third set of time series shows the absolute values over Greenland (green) and Antarctica (blue).



**Fig. 3:** GOME-2 NO<sub>2</sub> time series of VC, RMS,  $\chi^2$ , and INT (top to bottom) from 2007-2010 (solid lines). SCIAMACHY NO<sub>2</sub> time series are also plotted from 2007-2010 (dashed lines) and from 2003-2006 (dotted line). Time series are plotted for the Sahara (red) and Pacific (blue) box.



**Fig. 4:** Same as Fig. 3 but time series are normalized to January 2007 and 2003, respectively.

The last set of plots is dedicated to the polar boxes of geo-locations. In green time series for Greenland are shown and in blue the time series for Antarctica are drawn.

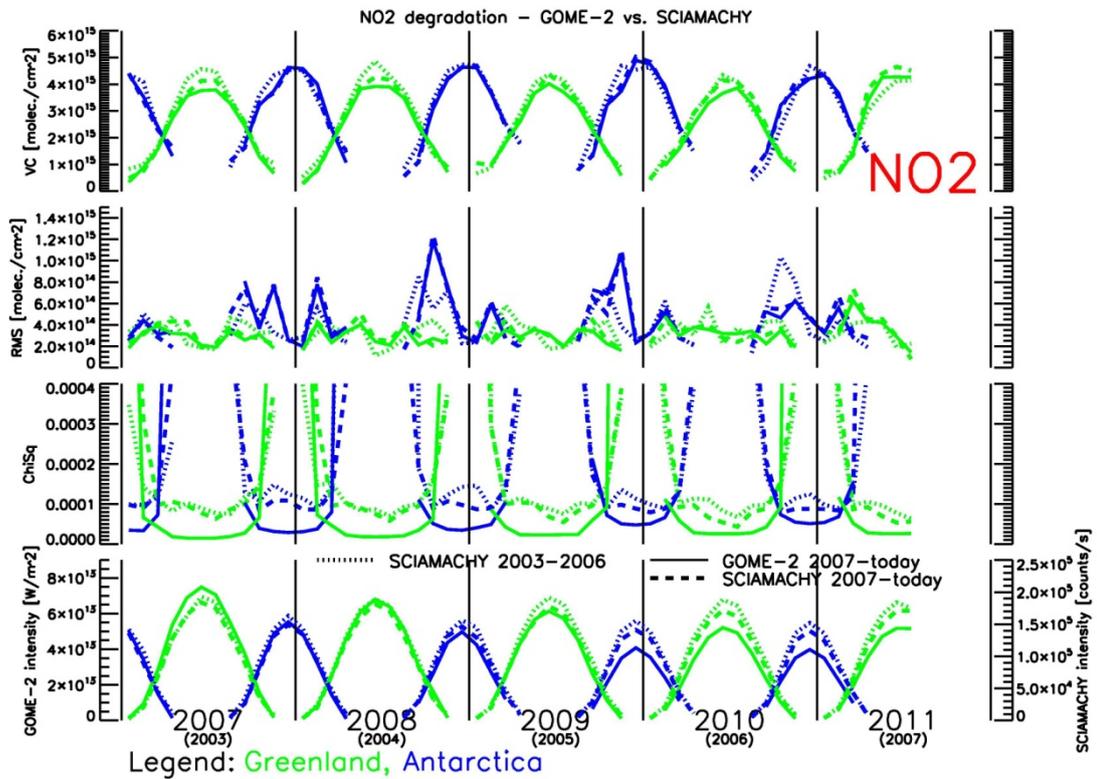


Fig. 5: Same as Fig. 3 but for the combined Antarctica (blue) and Greenland (green) box.

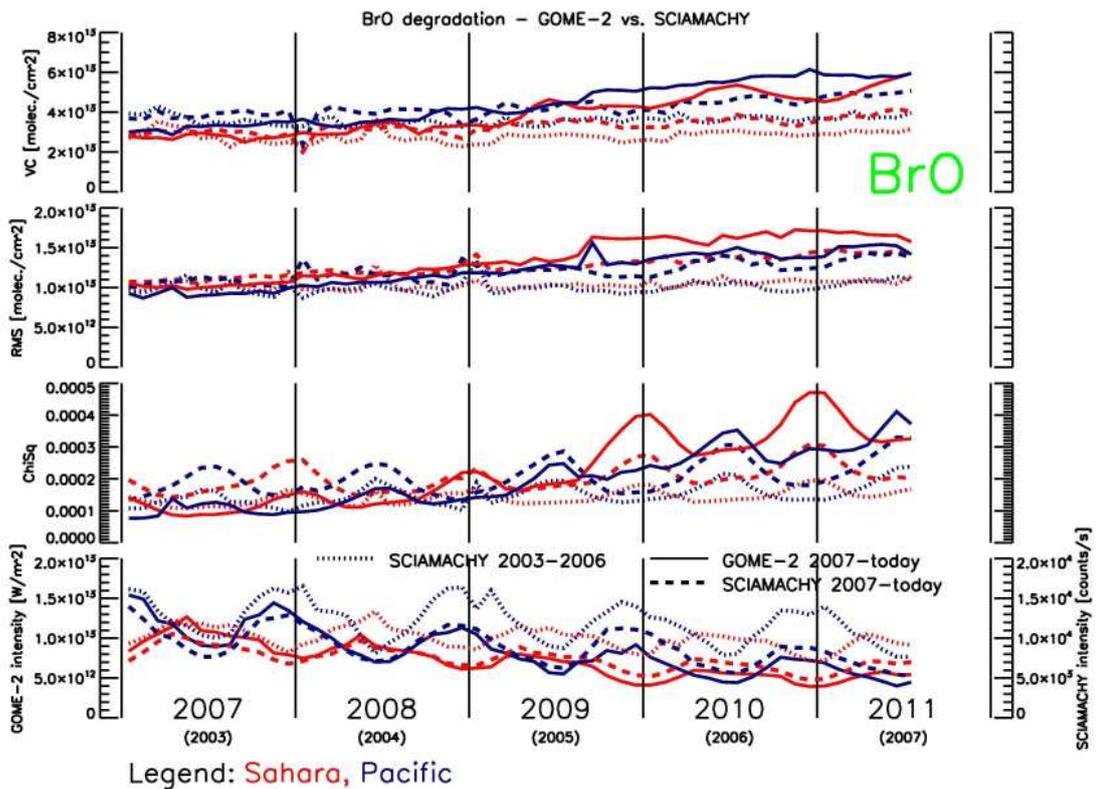


Fig. 6: Same as Fig. 3 but for BrO.

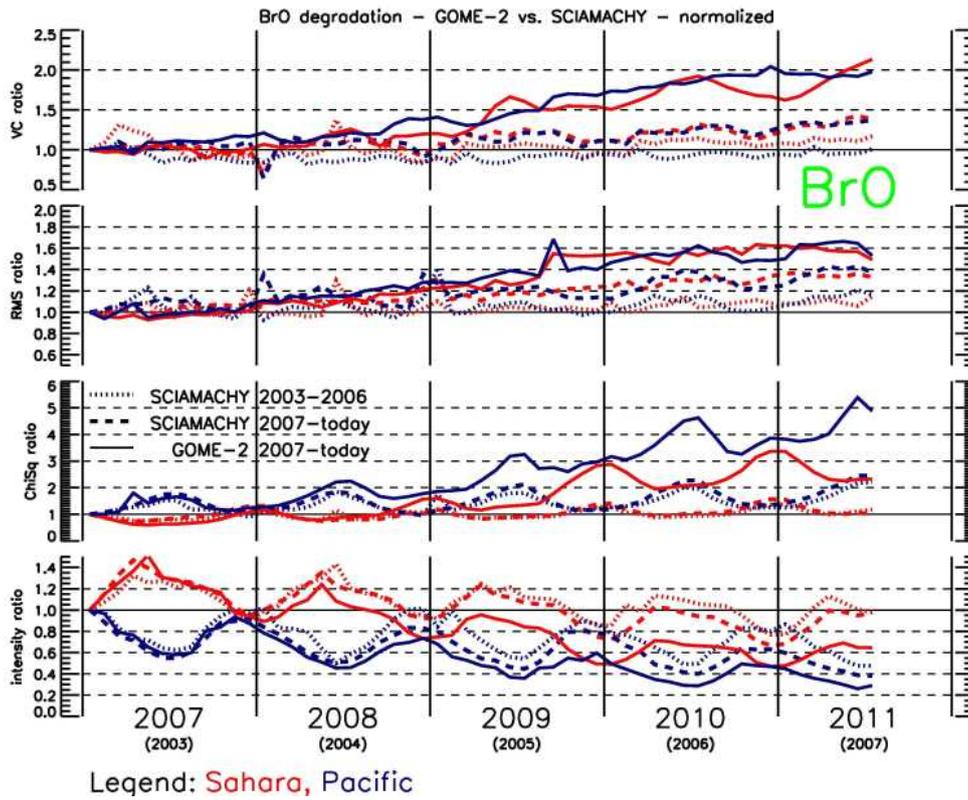


Fig. 7: Same as Fig. 6 but time series are normalized to January 2007 and 2003, respectively.

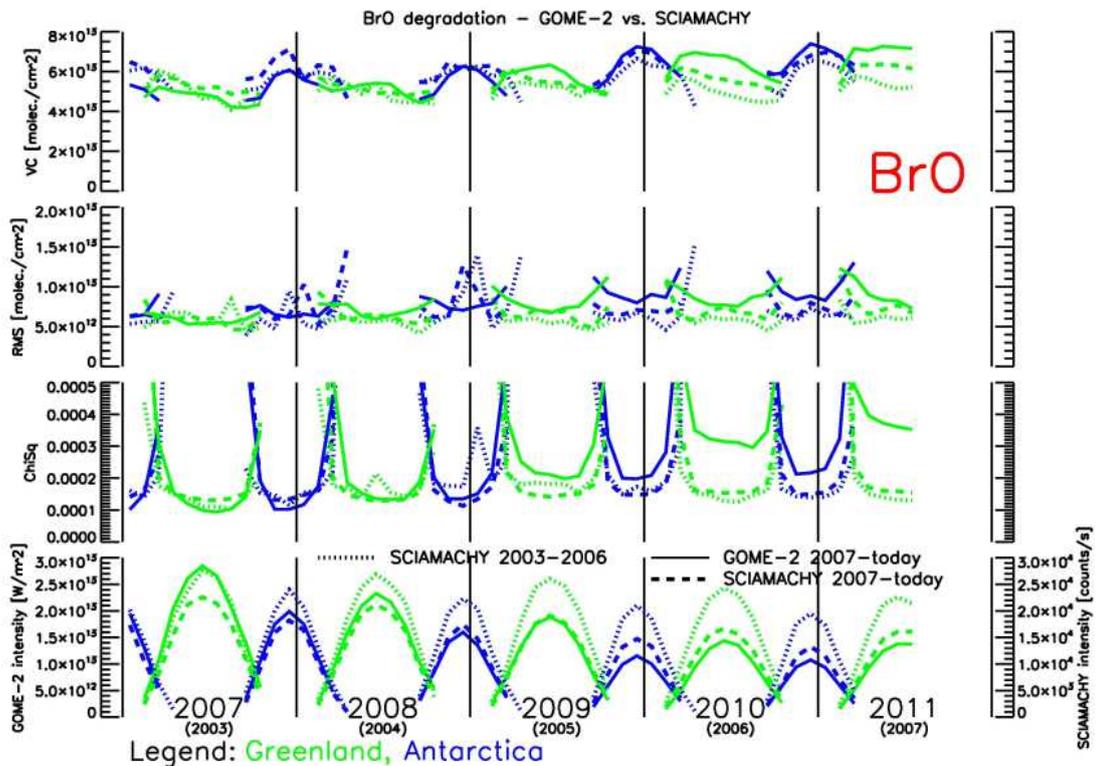


Fig. 8: Same as Fig. 6 but for the combined Antarctica (blue) and Greenland (green) box.

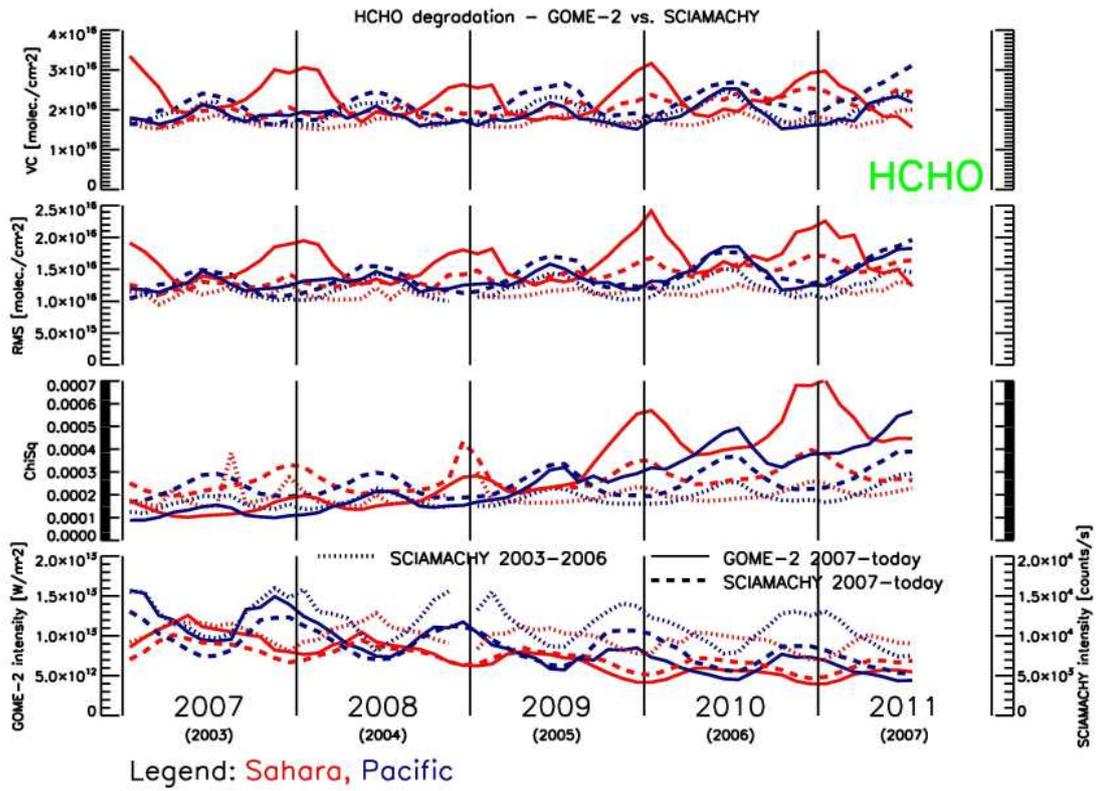


Fig. 9: Same as Fig. 3 but for HCHO.

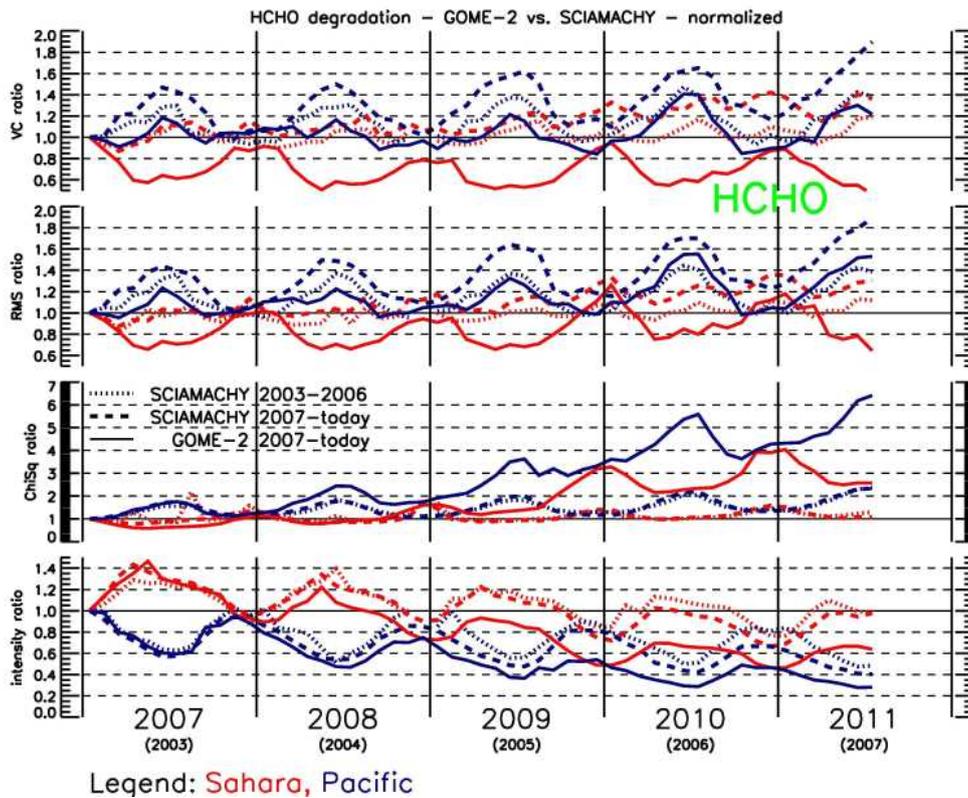


Fig. 10: Same as Fig. 9 but time series are normalized to January 2007 and 2003, respectively.

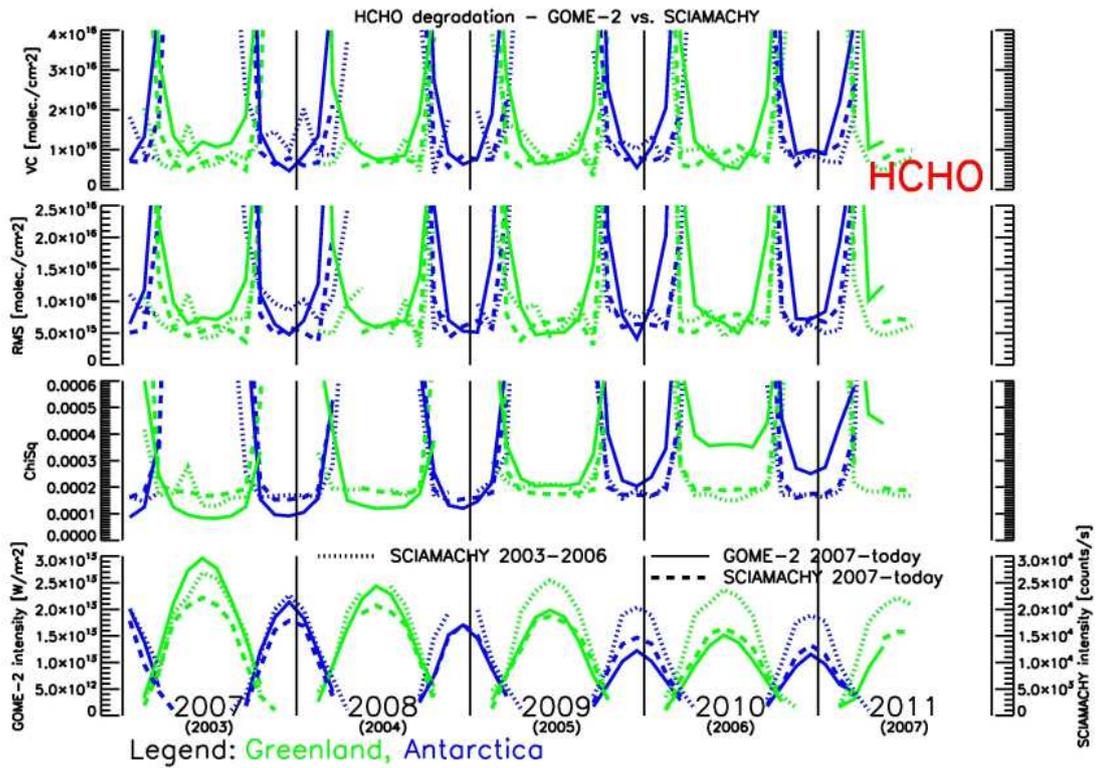


Fig. 11: Same as Fig. 9 but for the combined Antarctica (blue) and Greenland (green) box.

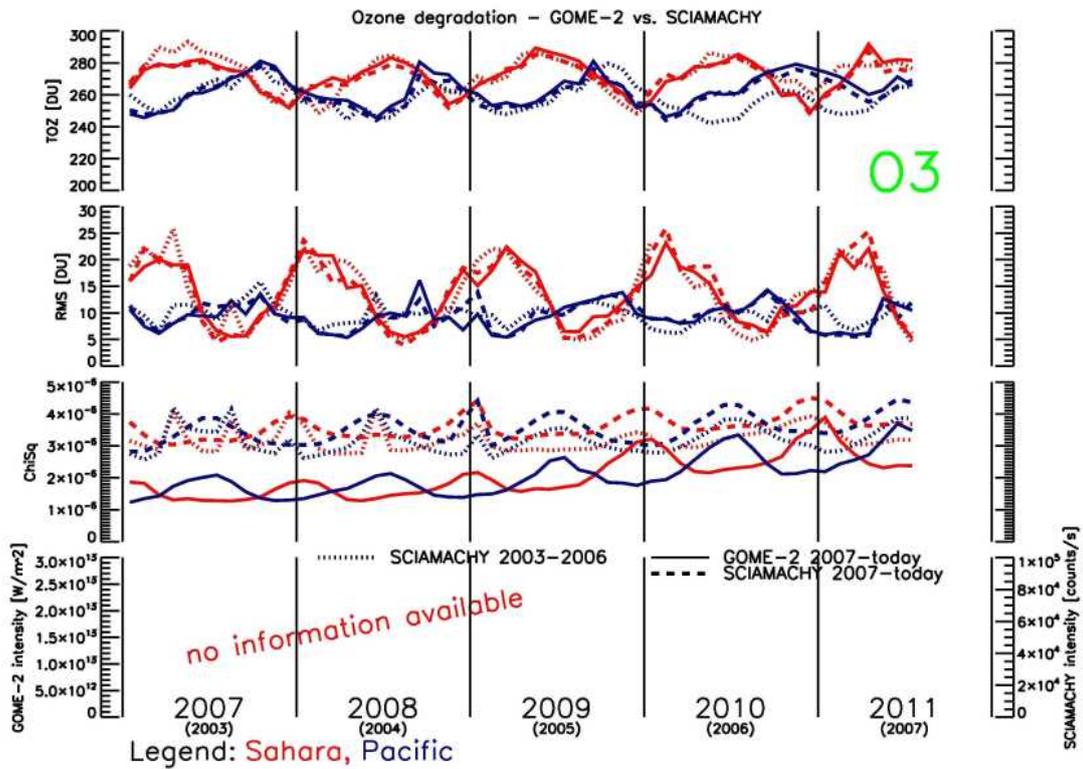


Fig. 12: Same as Fig. 3 but for O<sub>3</sub>.

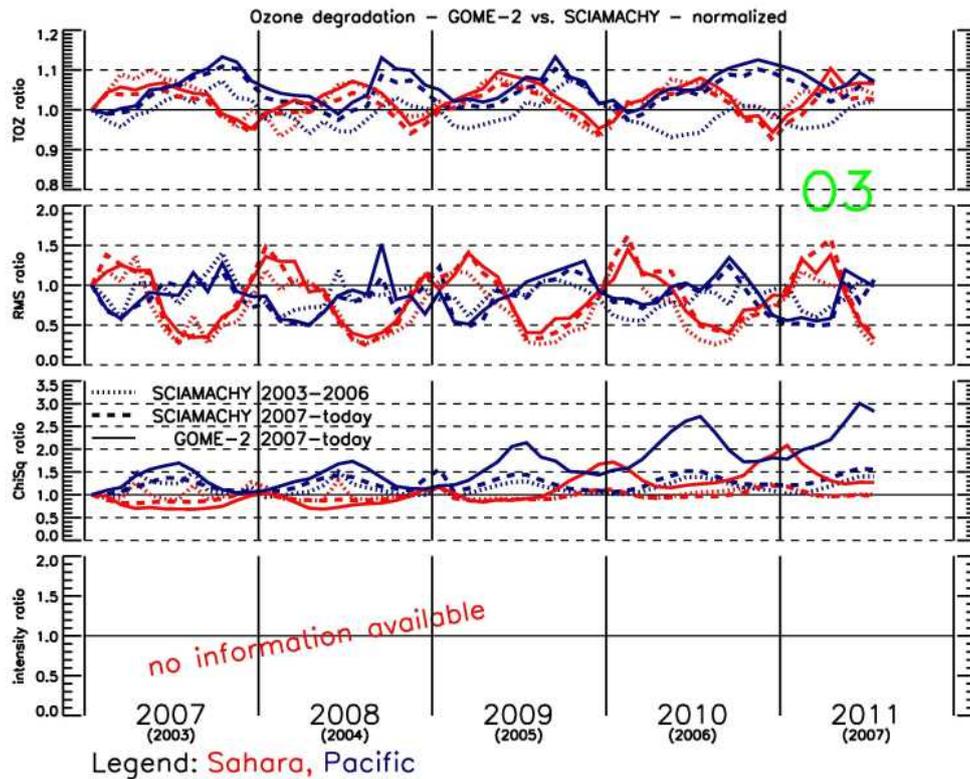


Fig. 13: Same as Fig. 12 but time series are normalized to January 2007 and 2003, respectively.

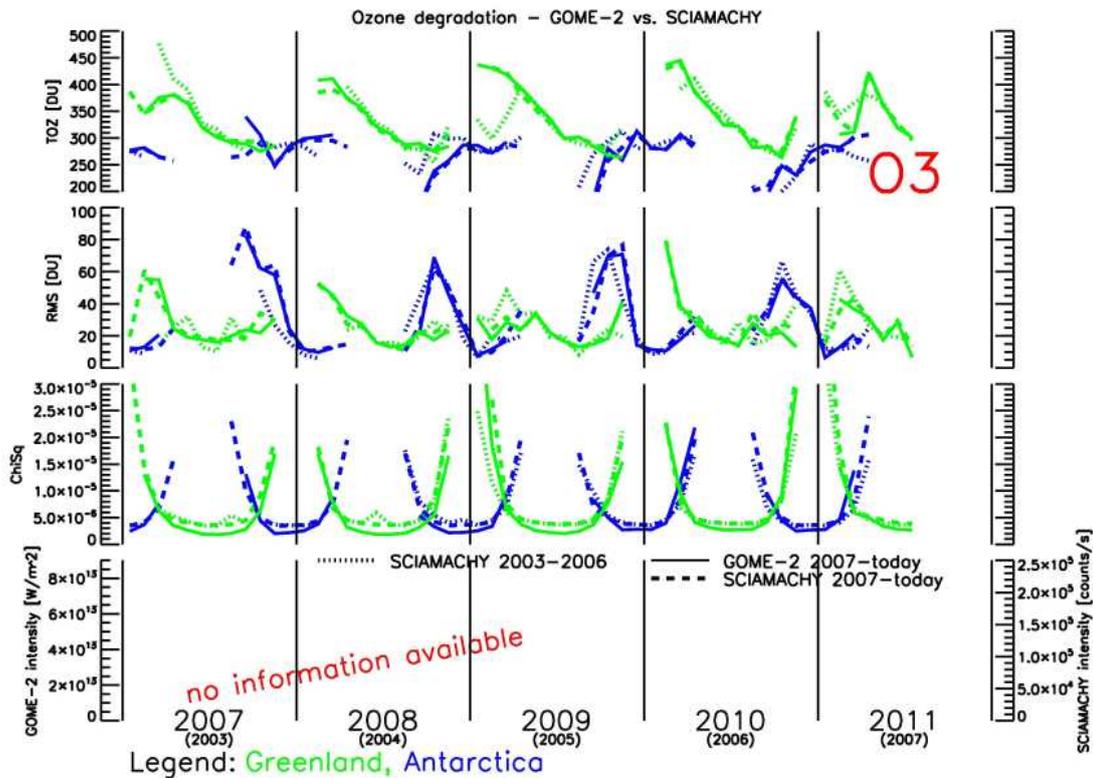


Fig. 14: Same as Fig. 12 but for the combined Antarctica (blue) and Greenland (green) box.

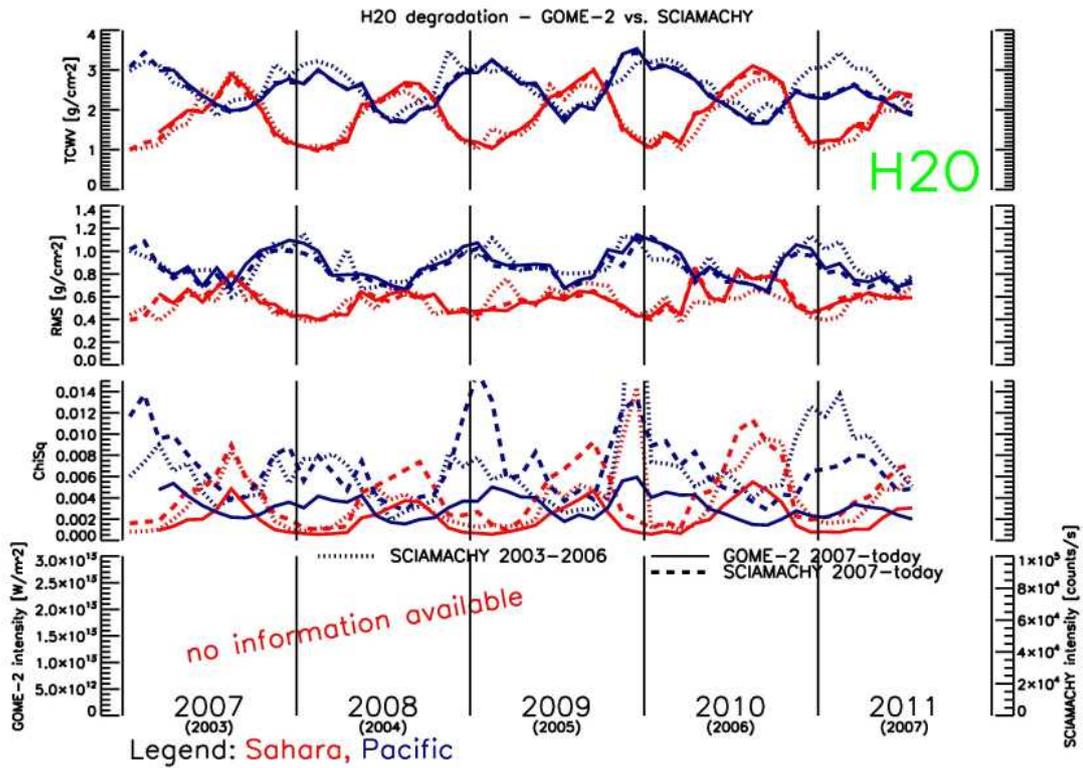


Fig. 15: Same as Fig. 3 but for H<sub>2</sub>O.

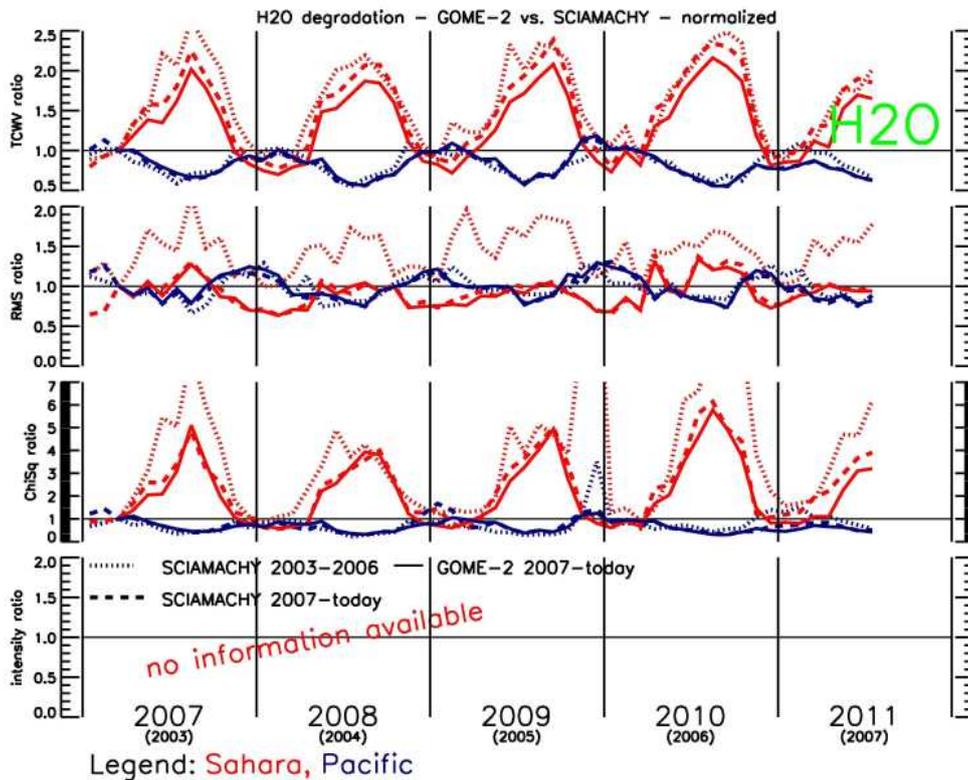


Fig. 16: Same as Fig. 15 but time series are normalized to January 2007 and 2003, respectively.

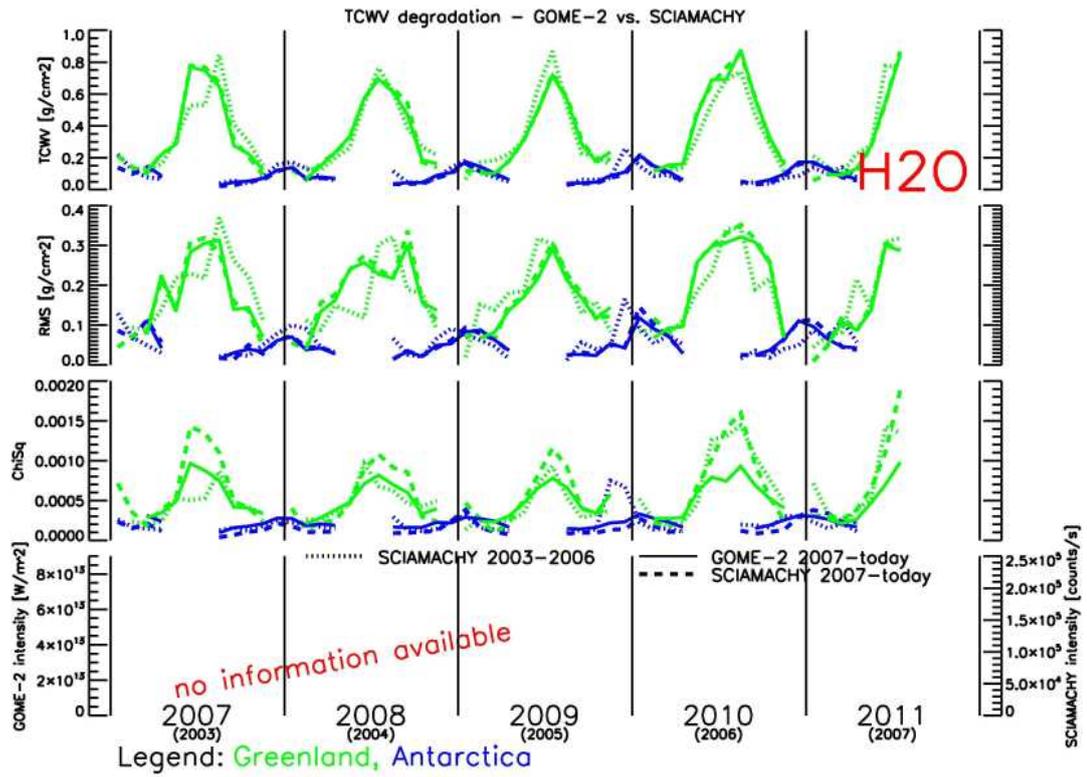


Fig. 17: Same as Fig. 15 but for the combined Antarctica (blue) and Greenland (green) box.

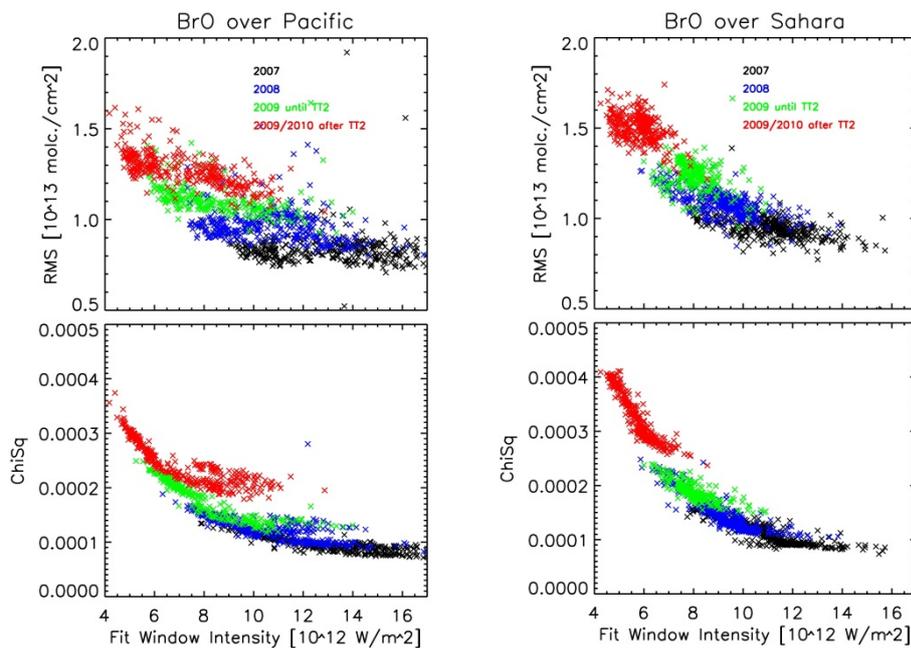
## 4.2 Intensity dependencies

### 4.2.1 Based on daily averages

Based on daily means, the dependence of  $\chi^2$  and RMS on the earthshine fit window intensity has been investigated for the whole GOME-2 time series within the example boxes over the Pacific and the Sahara. Fig. 18 shows results based on daily averages. From photon shot noise alone, it is expected to see a  $1 / (\text{INT})^2$  behaviour.

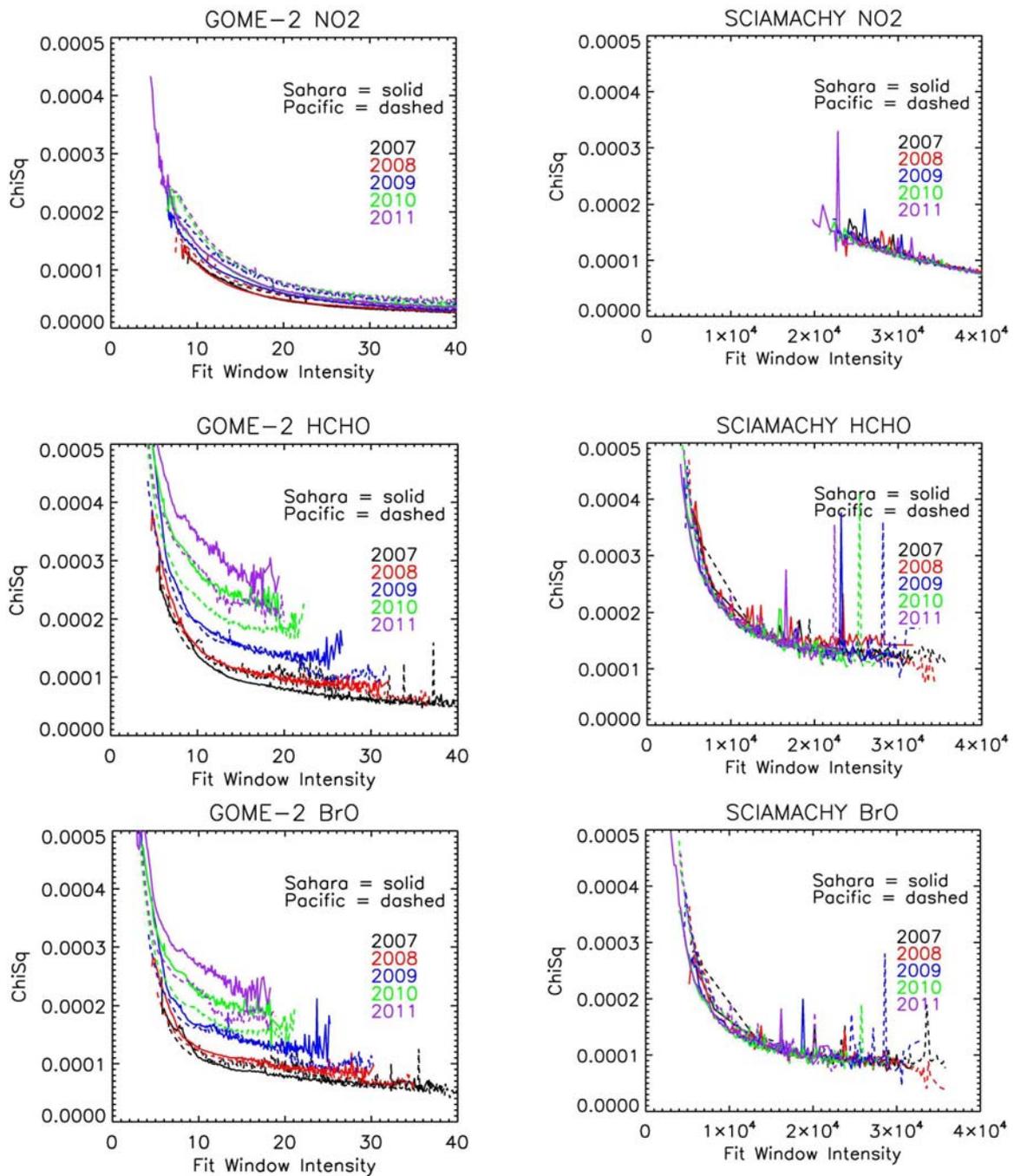
The figures show several interesting features:

1. The pattern for  $\chi^2$  over the Sahara follows more or less the expected behaviour, indicating that throughput loss is responsible for most of the loss in fitting precision.
2. The relationship over the Pacific is much less compact and residuals after the last throughput test are higher than for similar intensities earlier. This points at additional spectral degradation or calibration issues.
3. The distributions for RMS are less compact but follow a similar pattern, indicating that throughput loss is the main reason for the observed scatter in the VC results, again with some differences between the regions. However, it is unexpected to see such a clear stratification with years in the RMS over the Pacific – here, intensity does not appear to be the only variable as for the same intensity, RMS is increasing over the years.



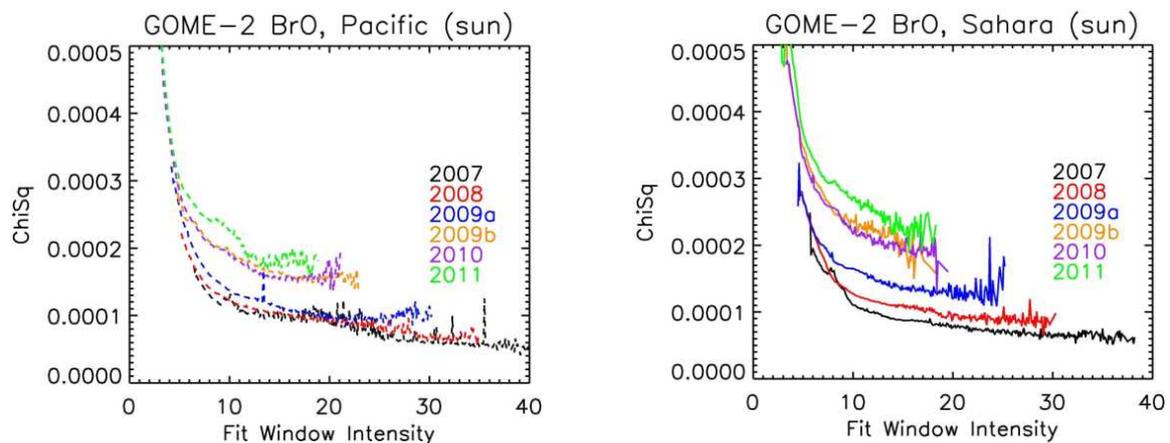
**Fig. 18:** Scatter plots of BrO RMS (top) and  $\chi^2$  (bottom) over earthshine fit window intensity for BrO over the Pacific (left) and the Sahara (right). Each cross represents one day of data.

#### 4.2.2 Based on all data



**Fig. 19:** Comparison between GOME-2 (left) and SCIAMACHY (right). Intensity dependent behaviour of  $\chi^2$  spitted into a year-by-year view. Plots are shown for NO<sub>2</sub> (top), HCHO (middle), and BrO (bottom). Two boxes of geo-locations are shown: Sahara (solid lines) and Pacific (dashed lines). GOME-2 NO<sub>2</sub> shows little changes over the years, whereas GOME-2 BrO and HCHO show higher fit residuals for the same intensity, year after year. Also visible is the different behaviour for  $\chi^2$  in 2010 and 2011 when compared to the previous years, fit residuals are lower over the Pacific than they are over the Sahara. SCIAMACHY data do not deviate from the  $1/(\text{INT})^2$  behaviour. All scan angles have been considered.

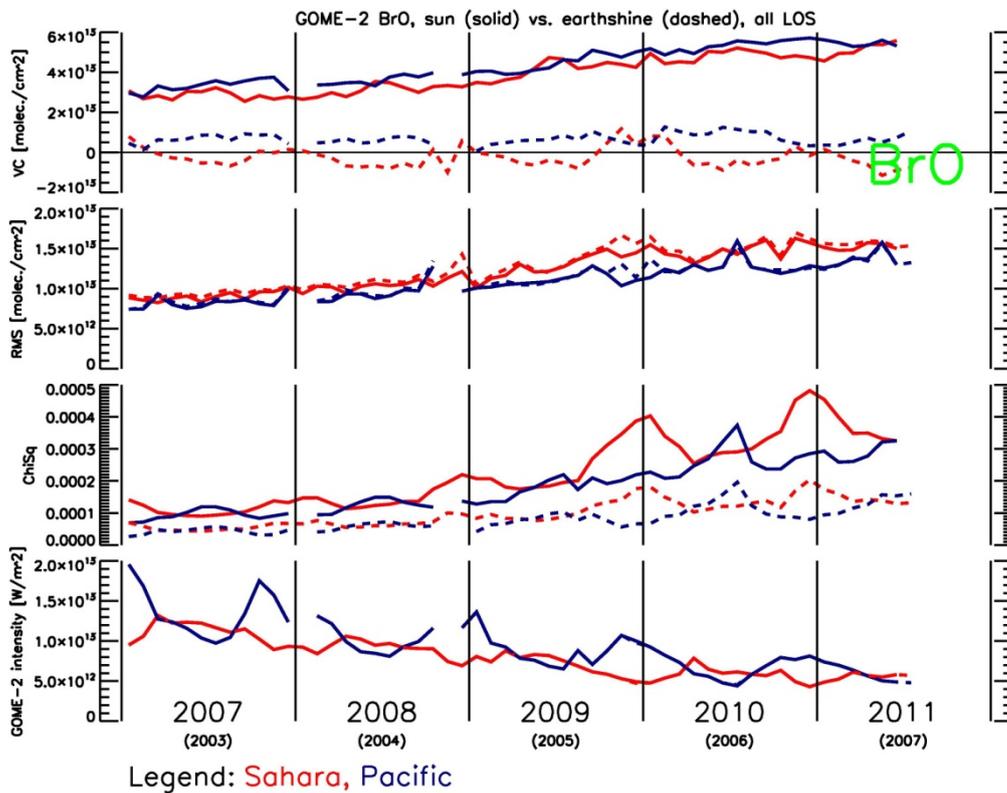
The comparison between GOME-2 and SCIAMACHY shows overall good agreement in the intensity dependence of  $\chi^2$ . This indicates that the basic relationship is consistent. A quantitative comparison cannot be done as intensity values from the SCIAMACHY data used are uncalibrated and therefore have been arbitrarily scaled in Fig. 19. This figure shows the intensity dependent behaviour of  $\chi^2$  split into a year-by-year view. Plots are shown for NO<sub>2</sub> (top), HCHO (middle), and BrO (bottom). Two boxes of geo-locations have been investigated: Sahara (solid lines) and Pacific (dashed lines). GOME-2 NO<sub>2</sub> shows little changes over the years, whereas GOME-2 BrO and HCHO show higher fit residuals for the same intensity, year after year. Also visible is the different behaviour for  $\chi^2$  in 2010 and 2011 when compared to the previous years, fit residuals are lower over the Pacific then they are over the Sahara. SCIAMACHY data do not deviate from the  $1/(\text{INT})^2$  behaviour and are quite consistent, year after year. This result clearly illustrates that while most of the degradation in fit quality for BrO and HCHO in GOME-2 data is from throughput loss, there is another, systematic component that increases over time. This systematic effect is also linked to the 2<sup>nd</sup> throughput test as Fig. 20 suggests. Over the Pacific (and the Sahara) the increase in  $\chi^2$  for constant fit window intensity leaps up when data before and after the 2<sup>nd</sup> throughput test are being compared. No similar effect is observed for SCIAMACHY data. The same is true for NO<sub>2</sub> but at a lower level (less systematic increase).



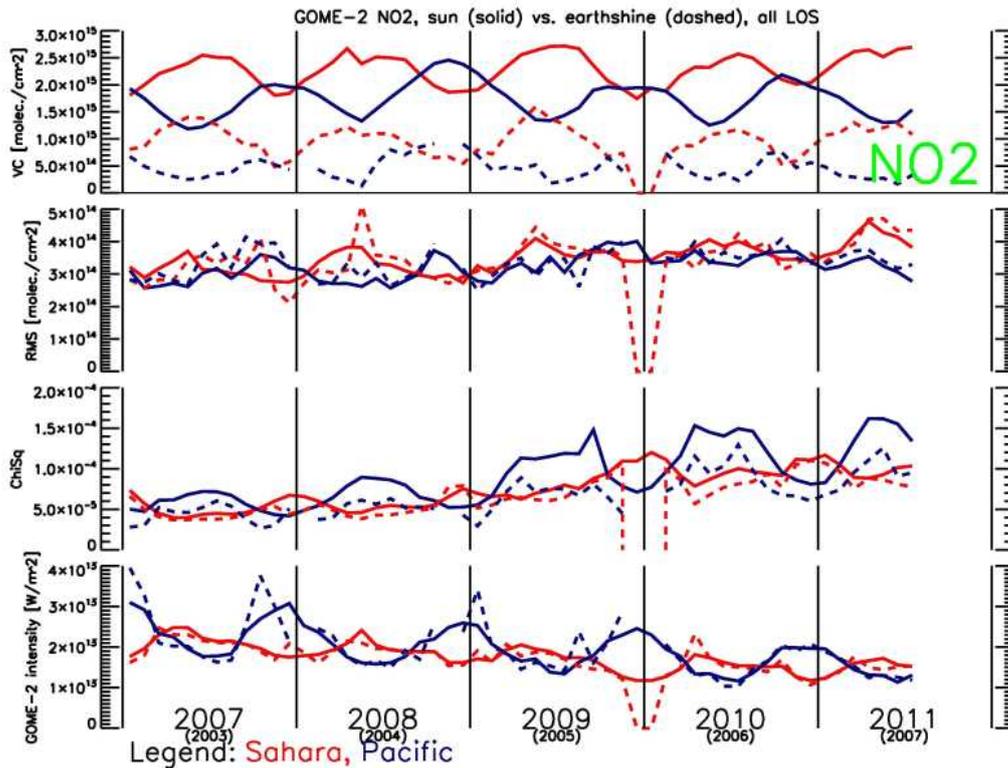
**Fig. 20:** Intensity dependent behaviour of  $\chi^2$  for GOME-2 BrO over the Pacific (left) and the Sahara (right). Similar to Fig. 19 data from January 2007 to July 2011 have been plotted. Data from the year 2009 have been divided into before (2009a, blue) and after (2009b, orange) the 2<sup>nd</sup> throughput test. A gap between both data sets is visible.

### 4.2.3 Solar vs. earthshine spectra

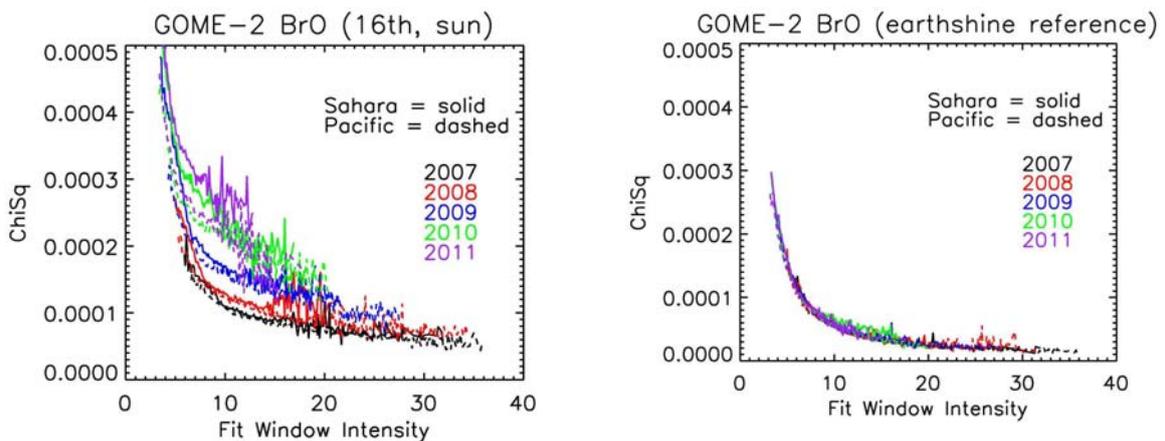
To investigate a possible impact of relative changes between solar irradiance and earthshine spectrum, daily background spectra over the Pacific were created by averaging all measurements in a box between 10°S and 10°N and 180°W to 140°W. A new BrO and NO<sub>2</sub> level 2 data set covering one day per month (the 16<sup>th</sup>) was created by using these earthshine spectra as background for comparison to data generated with solar spectra. As expected, the fitting residuals are clearly smaller when using the earthshine background (roughly a factor of 2.5 over the complete time series, cf. Fig. 21, not so much for NO<sub>2</sub>, cf. Fig. 22), mainly because of less Ring-effect which is not fully corrected for in the retrieval. As expected the RMS is as large as with the solar background, indicating that growing differences between solar and earth-shine spectra are either not happening or not affecting the scatter in retrievals. Absolute columns cannot be directly compared between the two data sets as using a background spectrum derived over the Pacific forces BrO and NO<sub>2</sub> in equatorial regions to very small values. The match in intensity must be expected because the same earth-shine spectra were used for both the sun and the earthshine background.



**Fig. 21:** GOME-2 BrO time series as shown in Fig. 6 using the solar background (solid lines) and earthshine background (dashed lines). Data have been generated for only one day per month (the 16<sup>th</sup>) and all scan angles have been considered.



**Fig. 22:** GOME-2 NO<sub>2</sub> time series as shown in Fig. 3 using the solar background (solid lines) and earthshine background (dashed lines). Data have been generated for only one day per month (the 16<sup>th</sup>) and all scan angles have been considered. Monthly means for December 2009 and January 2010 could not be generated for the earthshine background, explaining the dip in the time series.



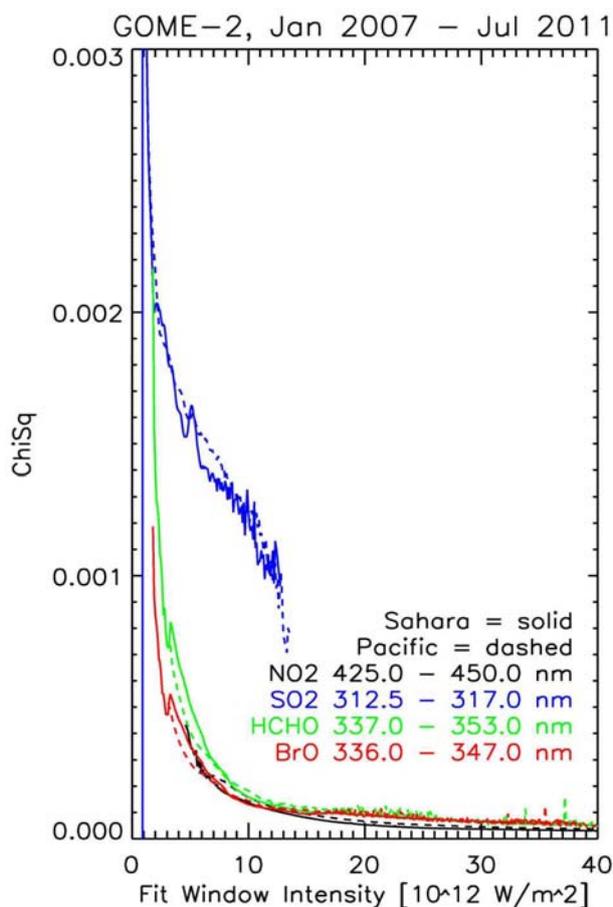
**Fig. 23:** Intensity dependent behaviour of  $\chi^2$  for GOME-2 BrO over the Sahara and the Pacific. Similar to Fig. 19 data from January 2007 to July 2011 have been plotted on the left side using only measurements from the 16<sup>th</sup> of each month. On the right side also data from only the 16<sup>th</sup> of each month have been plotted, this time the solar reference spectra have been substituted by earthshine spectra form over the Pacific (10°S and 10°N and 180°W to 140°W).

Very interesting to observe is the difference in the intensity dependence of  $\chi^2$  when BrO was retrieved using an earthshine reference spectrum over the Pacific. Fig. 23 shows the intensity dependency of  $\chi^2$  using only data from the 16<sup>th</sup> of each month within the observation period of GOME-2. For one (left figure) it has been plotted the

same fashion as illustrated in Fig. 19. The year-to-year increase in  $\chi^2$  is visible as previously shown. When the sun spectra are substituted by earthshine spectra over the Pacific the overall values for  $\chi^2$  decrease and – what is more important – the systematic increase in  $\chi^2$  from year-to-year vanishes. The conclusion to be drawn from this result is very interesting as it hints to suggesting that the solar measurements of GOME-2 introduce uncertainties for the retrieval of BrO. Instrument components affected by this conclusion must be in the light path for the sun measurements that are not in the light path of the earthshine measurements (i.e. mainly the sun mirror and the sun diffusor).

#### 4.2.4 SO<sub>2</sub> – a separate challenge

GOME-2 SO<sub>2</sub> has also been investigated in this context. Fig. 24 shows the intensity dependent behavior of SO<sub>2</sub> w.r.t. HCHO, BrO, and NO<sub>2</sub>. For the same intensity, SO<sub>2</sub> has by one order of magnitude higher  $\chi^2$ . SO<sub>2</sub> is being retrieved close to the edge of channel 2 with even higher losses of throughput and an etalon correction is only partially performed. Therefore, SO<sub>2</sub> fits have not been considered further.



**Fig. 24:** Intensity dependent behaviour of  $\chi^2$  for NO<sub>2</sub> (black), SO<sub>2</sub> (blue) HCHO (green), and BrO (red), separated into boxes of geo-locations over the Sahara (solid lines) and the Pacific (dashed lines). Altogether, NO<sub>2</sub>, BrO, and HCHO show the  $1/(\text{INT})^2$  behaviour when compared to SO<sub>2</sub>. The SO<sub>2</sub> fitting window is located close to the edge of Channel 1 where the degradation of the throughput is the strongest. In addition, the fitting window is fairly small in contrast to the other species.

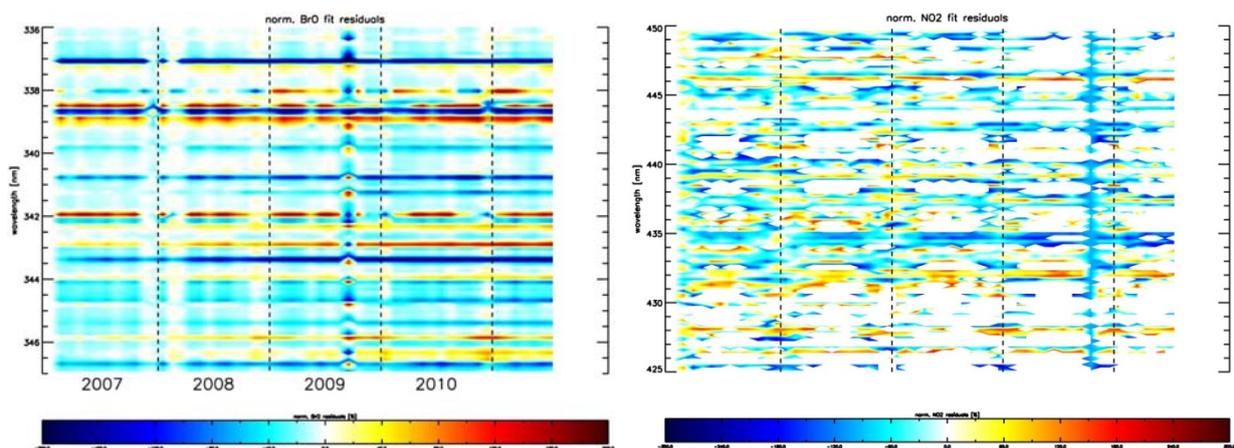
### 4.3 Wavelength dependency of retrieval residuals

For the investigation of the wavelength dependency of the retrieval residuals, residuals from all BrO and NO<sub>2</sub> fits over the Sahara region have been averaged for one day per month (the 16<sup>th</sup>). As a large number of spectra is averaged (about 4000

per day), the results are the non-random component of the fitting residual. Any systematic changes in calibration quality resulting from degradation should be visible in the temporal evolution of these average residuals.

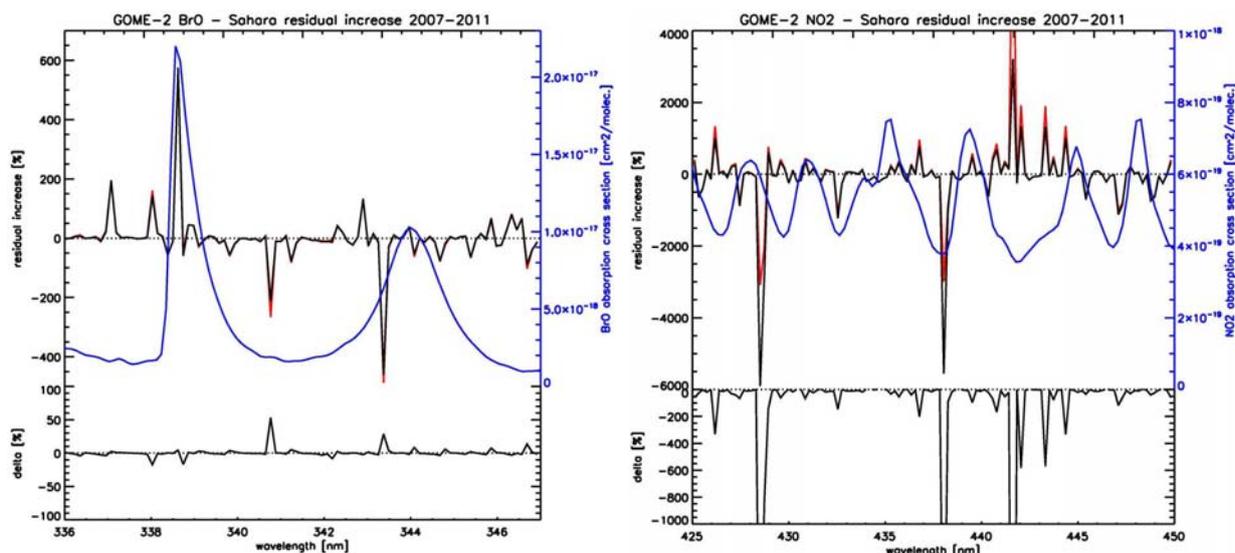
Results (cf. Fig. 25) indicate that no particular degradation effect is visible when the fitting window is split into its contributing wavelengths. For BrO there are some time periods with larger changes, in particular during the 2<sup>nd</sup> throughput test, but overall there is only a moderate increase of the mean residual without significant change in pattern. This indicates that at least in this wavelength region, there is no systematic calibration problem developing over time that is not taken care of by the polynomial in the DOAS retrieval. For NO<sub>2</sub> the evolution of fitting residuals does not show any pattern. The residuals seem to vary randomly. The vertical structure in October 2010 is due to a sampling issue as there was no data of the 16<sup>th</sup>.

These results were not expected as the initial assumption was, that the systematic increase in residuals should be related to an instrument change that should show up in change of the averaged residuals over time. In fact, the results seem to indicate that the residual remains remarkably constant but increases in amplitude. This point warrants further investigations, for example with respect to a possible impact of changes in instrument slit function.



**Fig. 25:** Wavelength dependent in-/decrease of fit residuals for GOME-2 BrO (left) and NO<sub>2</sub> (right). Data have been normalized to January 2007.

The overall in- and decrease in GOME-2 BrO and NO<sub>2</sub> fit residuals over the Sahara has been plotted in Fig. 26 in units of % change between January 2007 and July 2011. The signature does not match the absorption cross sections of BrO and NO<sub>2</sub>, respectively. The increase of fitting residuals for BrO is continuously whereas for NO<sub>2</sub> it varies randomly in time (cf. Fig. 25).

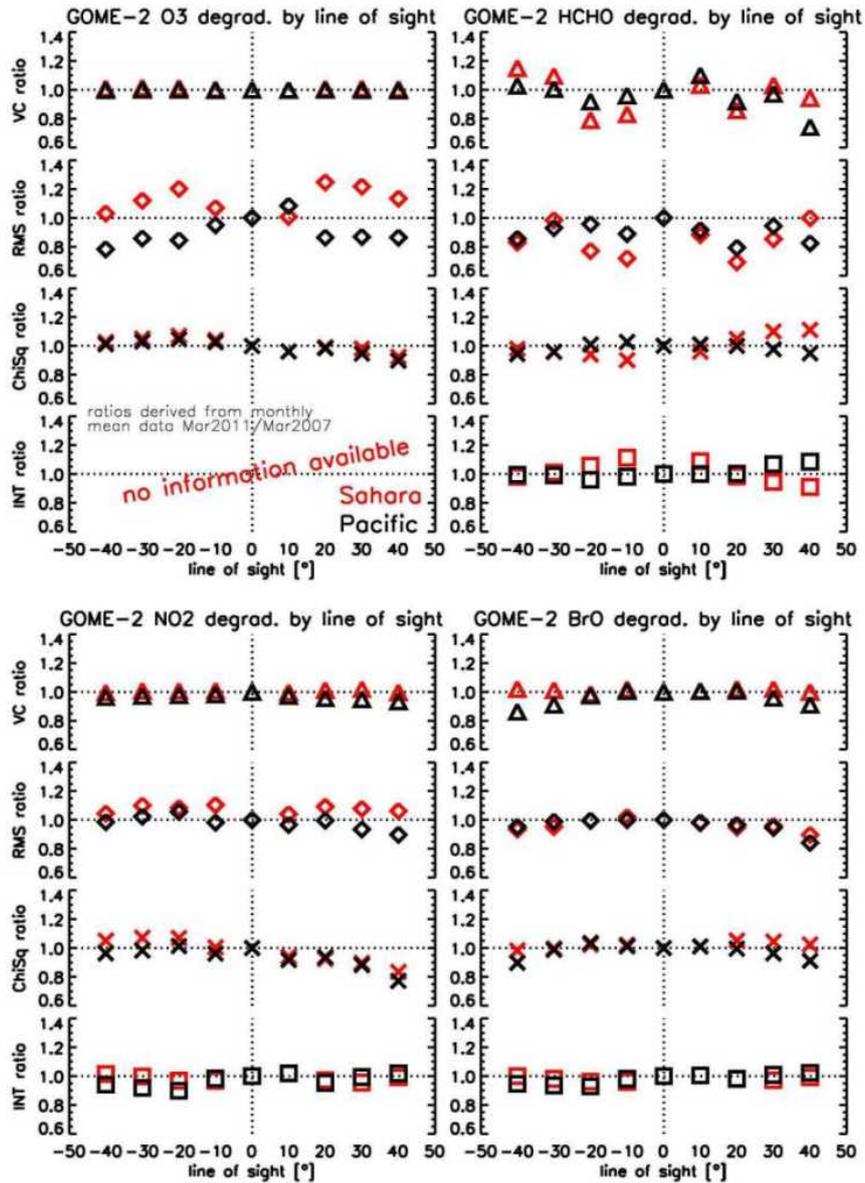


**Fig. 26:** GOME-2 BrO (left) and NO<sub>2</sub> (right) fit residual increase over the Sahara between January 2007 and July 2011. As September 2009 (BrO) and October 2010 (NO<sub>2</sub>) show large deviations from the normal behaviour, the in-/decrease of fit residuals has been determined without the particular month (red line). The difference is plotted below. Absorption cross sections are plotted as blue line.

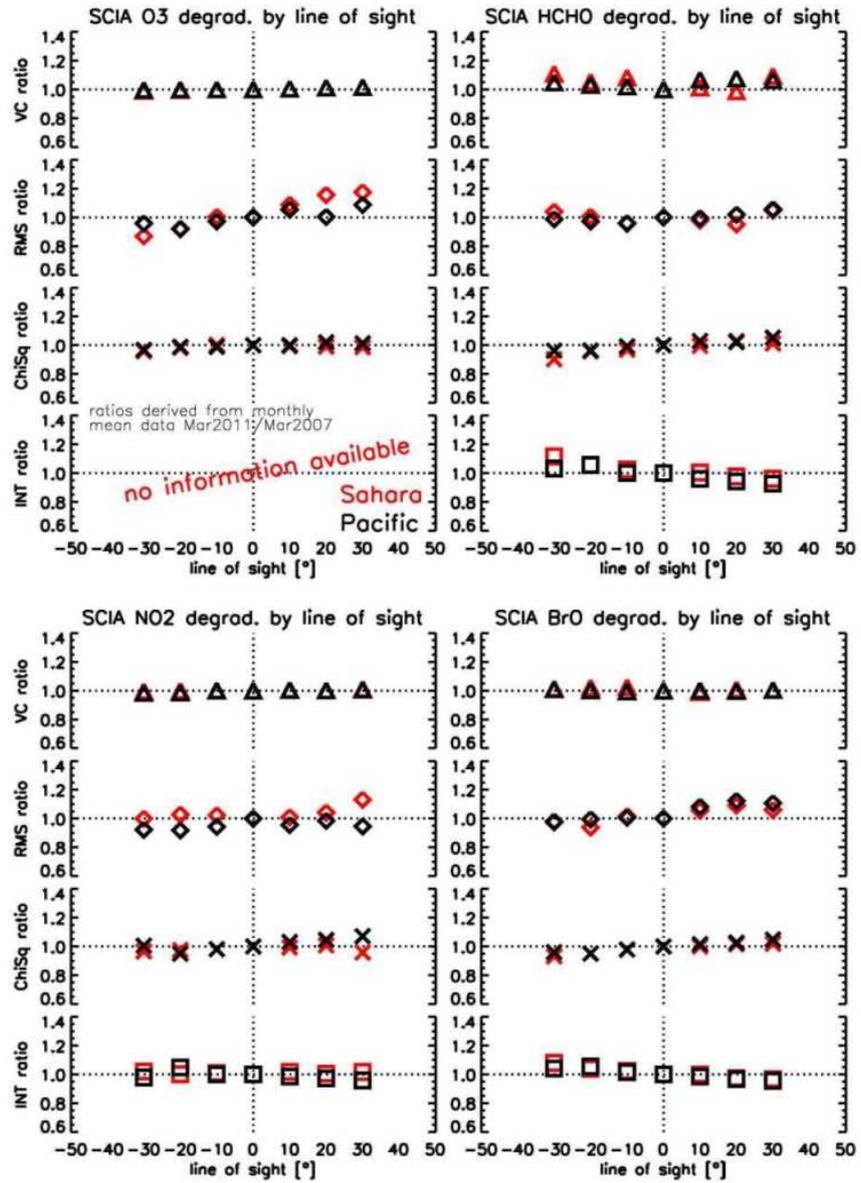
#### 4.4 Line of sight dependency

No changes w.r.t. the time series, yet a summary similar to Fig. 2 is given below for GOME-2 (cf. Fig. 27) and SCIAMACHY (cf. Fig. 28) time series. For the presentation of these results, a separation into negative (East) and positive (West) lines of sight (LOS) has to be performed. This can be achieved with the help of the relative azimuth angle included in the level 2 data files. For SCIAMACHY the clear East-West LOS difference in VC, RMS, and  $\chi^2$  correlates with respective fit window intensities. For GOME-2 degradation due to contamination is believed to be the main driver. Note the wider range of scan angles for GOME-2. All LOS have been converted to top of atmosphere.

The differences between the patterns of GOME-2 and SCIAMACHY LOS dependence are probably due to scan angle position at which the solar measurements are being done. Differences between the Sahara and the Pacific are due to differences in the surface properties of these regions. The light coming from the Sahara desert is more polarized when compared to light reaching the instrument from the Pacific.

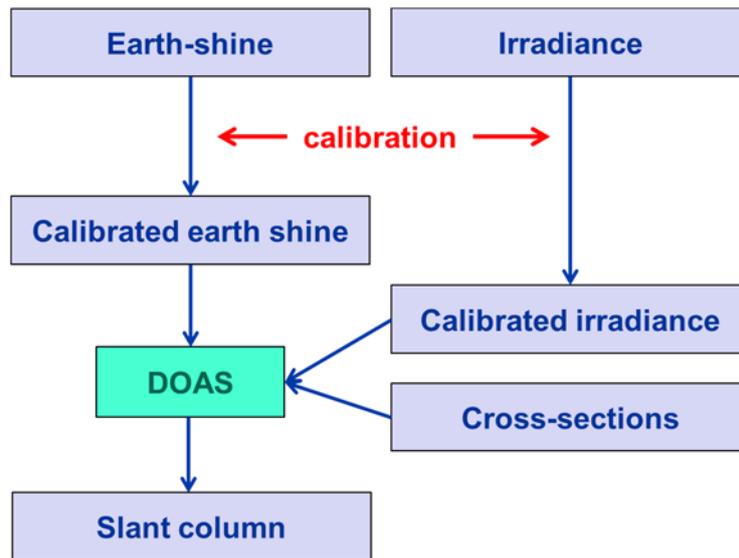


**Fig. 27:** Summary of line of sight dependencies for GOME-2 O<sub>3</sub> (upper left), HCHO (upper right), NO<sub>2</sub> (bottom left), and BrO (bottom right). Values are shown for the Sahara and the Pacific box of geolocations.



**Fig. 28:** Summary of line of sight dependencies for SCIAMACHY O<sub>3</sub> (upper left), HCHO (upper right), NO<sub>2</sub> (bottom left), and BrO (bottom right). Values are shown for the Sahara and the Pacific box of geolocations.

## 4.5 How DOAS retrievals are affected by degradation



**Fig. 29:** Schematic of the DOAS retrieval and the individual steps where degradation can affect the results. See text for discussion

One of the advantages of the DOAS retrieval technique is the inherent cancelling of many instrumental features which do affect other retrieval methods. The reason for the robustness of the approach is the use of a background measurement taken on the same day in the normalisation  $\ln I/I_0$  where  $I$  is the earth-shine and  $I_0$  usually is the solar irradiance measurement.

However, instrument degradation can still affect DOAS retrievals in different ways. These will briefly be discussed in the following

### 4.5.1 Loss of signal

The most obvious effect of degradation is a reduction of signal. This ultimately limits the precision of the measurements as shot noise in the light signal is proportional to the square root of the signal, and thus the signal to noise ratio decreases as intensity reduces. This will increase the residuals of the fit and the scatter in the retrieved slant (and vertical) column densities but should not result in a change of the average values.

As intensity decreases, there also is the effect of increasing importance of correlations between different absorption cross-sections. In many cases, this limits the accuracy of DOAS retrievals. For example, in the BrO or SO<sub>2</sub> retrievals, the strong absorption by O<sub>3</sub> has some effect on the results for the smaller absorbers. At decreasing signal to noise ratio, this effect gets more and more important.

Loss of signal problems are best investigated using scatter plots of fit residual as a function of intensity where they should appear as a clear  $1 / \text{SQRT}$  dependency.

### 4.5.2 Instrument changes not accounted for by calibration

In addition to general throughput loss, the spectral dependency of throughput can change by degradation, for example if a substance depositing on optical components has an absorbance that depends on wavelength or if a prism or disperser changes its spectral properties.

In principle, these changes should be corrected by calibration. However, if this calibration is not perfect or if the input to the calibration (e.g. polarisation measurements) is affected by degradation, the earth-shine measurements will have spectral features that potentially can correlate with absorption cross-sections used in the retrieval. If this is the case, the derived slant and vertical column densities change with degradation.

Spectral changes occurring through degradation often are broadband with wavelength and are to a large degree compensated by the polynomial used in the DOAS retrieval. However, the slant columns retrieved for absorbers, in particular if they have weak signatures are not entirely independent of the polynomial and thus even broadband changes can have an effect on the results for minor absorber.

Calibration related problems can usually be identified by their systematic dependency on parameters such as line-of-sight, solar zenith angle or azimuth angle. They do not increase the random noise but usually lead to increases in fitting residual.

In order to identify calibration related problems, several tests can be performed:

- Plots of  $\chi^2$  (changes) as function of LOS. This is useful to identify problems with the first mirror.
- Global maps of trace gases and  $\chi^2$  sometimes show clear patterns that can be linked to special calibration situations as e.g. in the case of OCIO.
- Evaluation of spectral change of residual over time without absorber. This can be used to track instrumental changes but can also be affected by changes in the atmosphere.
- Evaluation of  $\chi^2$  changes with number of spectra averaged. Ideally, this should follow the Poisson statistics. Any deviation from this behavior indicates systematic residuals which are either linked to problems in the DOAS retrieval (for example from non-perfect cross-sections) or to calibration issues.

#### 4.5.3 Differential changes between earth-shine and irradiance

If degradation affects earth-shine and solar irradiance measurements in a different way, e.g. because of an angle dependence on the mirrors used or because the degradation affects the diffuser plate, the cancelation of instrumental features in the DOAS retrieval is compromised. Any spectral feature in this differential degradation can then potentially correlate with the absorption cross-sections of the absorbers leading to biases in the retrieved results. Again, weak absorbers are affected most.

Differential changes between earth-shine and irradiance result in daily and seasonal offsets in the slant columns which can easily be identified as they affect all measurements independent of latitude and longitude in the same manner. The best known example is the effect of the diffuser on the GOME-1 instrument which introduced large seasonal changes in all small absorption signals.

In order to test for this type of degradation effect, two approaches can be taken

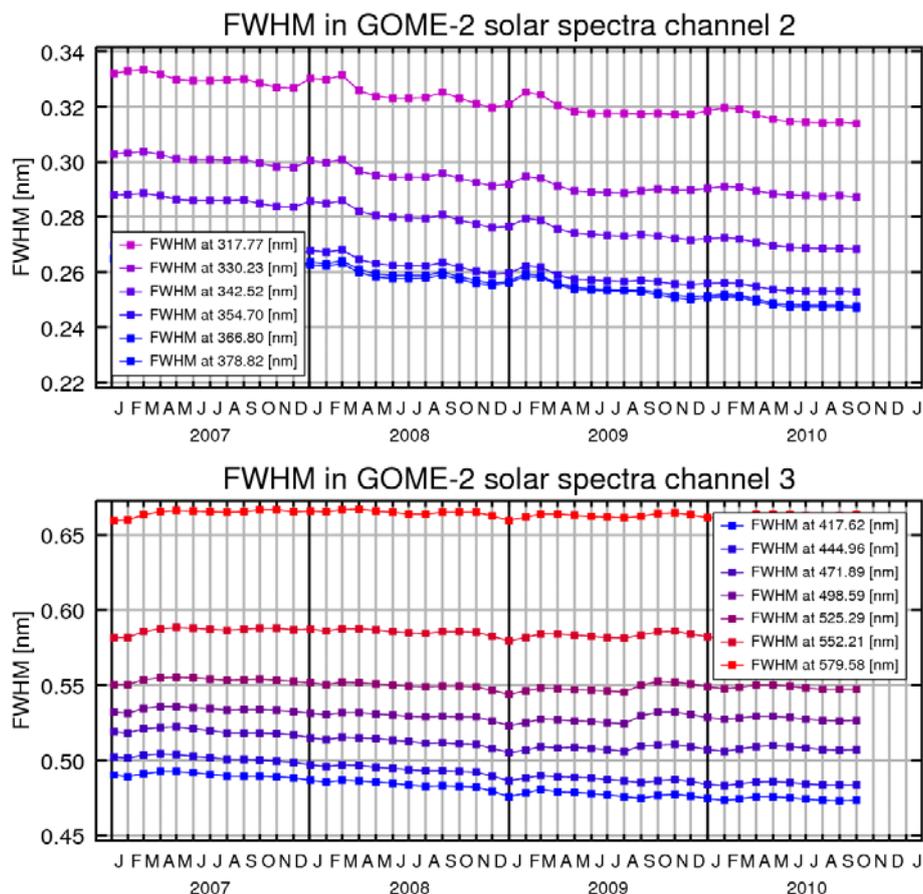
- Trace gas retrievals acting on solar spectra taken at different times. They should result in negligible slant columns for all absorbers and any deviation indicates a change in solar spectra
- Trace gas retrieval using earth-shine data as background. If there is a differential degradation between solar and earth-shine measurements, these fits should not be affected.

#### 4.6 Changes in GOME-2 solar spectra

Some aspects of instrument degradation can best be studied using GOME-2 solar irradiance measurements. This includes the evolution of instrument slit function as no complicating effects such as filling-in of Fraunhofer lines has to be considered. In Fig. 28, the results of a non-linear fit are shown in which the Kurucz solar spectrum is convoluted with a Gaussian and shifted on the GOME-2 irradiance measurements with the shift and the width of the Gaussian as fitting parameters. The fit is performed in small wavelength windows facilitating determination of the instrument slit width as a function of wavelength.

As can be seen from the figure, the FWHM of the GOME-2 slit function slightly decreased over time in both channel 2 and 3. There appears to be slight wavelength dependence in this effect with larger changes in the UV than in the visible. There also is a small seasonality in the retrieved FWHM at least in channel 2. The throughput-test had no discernible effect on the results in channel 2 but in channel 3, a slowdown of slit function changes is observed.

The origin of the changes in slit function is not clear. Possible explanations are mechanical deformation of the instrument as result of thermal changes which could bring the detectors closer to the focal plane. Accumulation or evaporation of etalons on optical parts cannot be fully excluded either but appear less probable.



**Fig. 30:** Evolution of FWHM of solar irradiance measurements at selected wavelength intervals for channel 2 (top) and channel 3 (bottom).

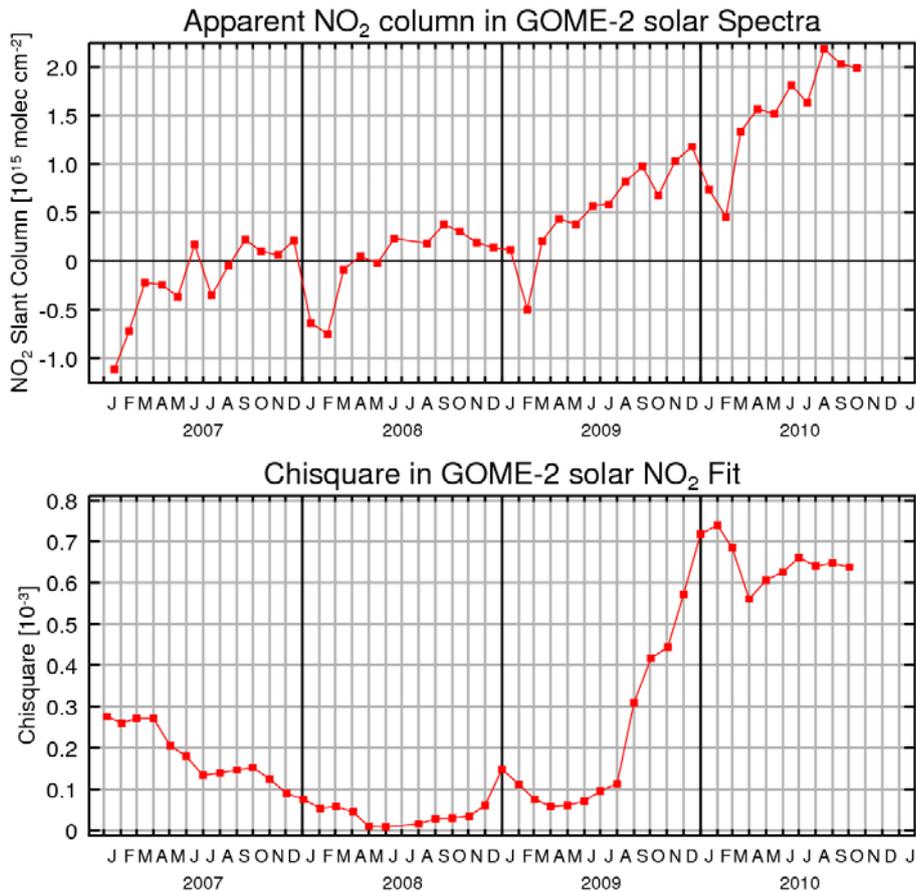
Another test that can shed some light on possible instrument changes are retrievals performed on solar irradiance measurements. In this test, a standard DOAS retrieval

is performed on the daily solar measurements using a single solar measurement as background. In theory, this should yield columns which are within the uncertainties as the solar irradiance is expected to be constant and is not affected by changing absorber signals. However, in practice, changes resulting from degradation or calibration affect the retrievals and led to the retrieval of non-zero slant column amounts.

In Fig. 31, the results of this analysis are shown for NO<sub>2</sub>. A solar measurement from July 2008 was selected as background. Instead of the expected constant and very small NO<sub>2</sub> column, NO<sub>2</sub> values of up to several 10<sup>15</sup> molec./cm<sup>2</sup> are retrieved, indicating a change in the solar spectra in 2009 and 2010 which correlates spectrally with NO<sub>2</sub> absorption structures. These columns are large enough to have a significant impact on retrieved atmospheric NO<sub>2</sub> fields in the tropics if a fixed solar background spectrum is used in the analysis.

Also shown in the figure is the  $\chi^2$  of the analysis which gives an indication of the change in solar spectrum. As can be seen,  $\chi^2$  increases with increasing time difference from the background spectrum (July 2008) which is behaviour also observed in other instruments. After the throughput test, a large increase in  $\chi^2$  is observed, indicating that solar spectra have changed strongly. No clear effect of the throughput test is visible in the apparent NO<sub>2</sub> columns.

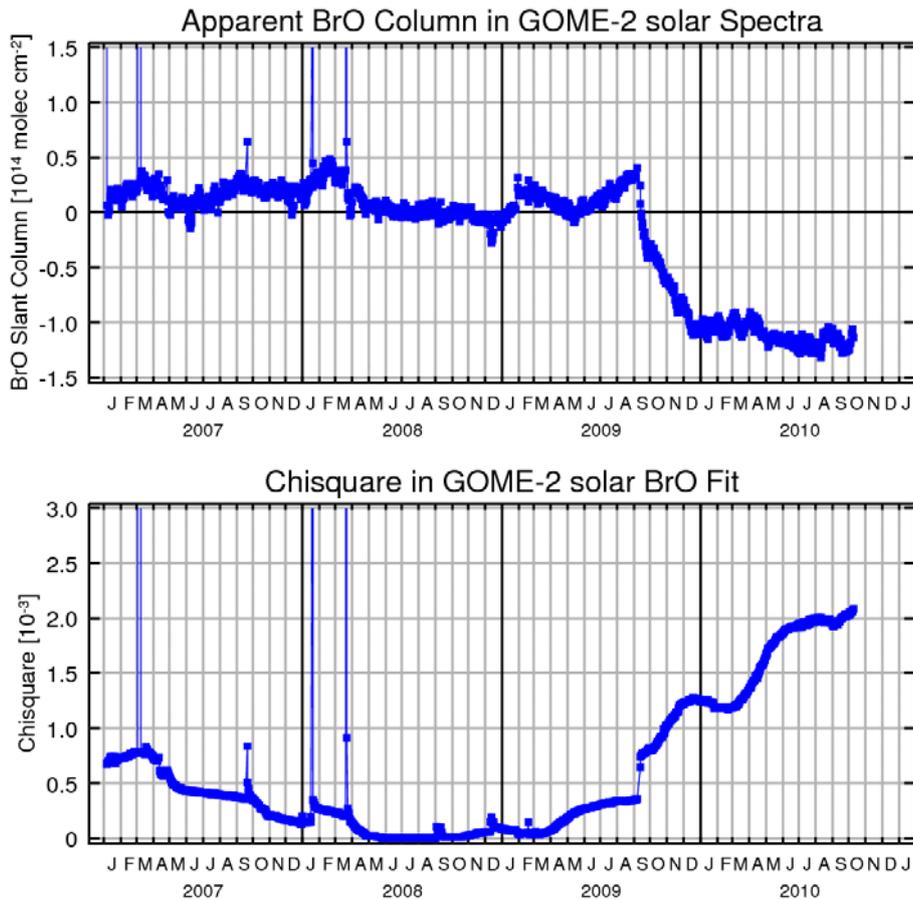
A specific feature of the time series are the systematically low apparent NO<sub>2</sub> column values in January and February which are linked to a clear and reproducible spectral pattern in the ratio of solar measurements from e.g. February and March. The origin of this specific pattern is still under investigation.



**Fig. 31:** Evolution of apparent NO<sub>2</sub> column (top) and chisquare (bottom) of a NO<sub>2</sub> DOAS analysis on daily solar irradiance measurements. An arbitrarily selected solar measurement from July 2008 is used as background spectrum in the analysis.

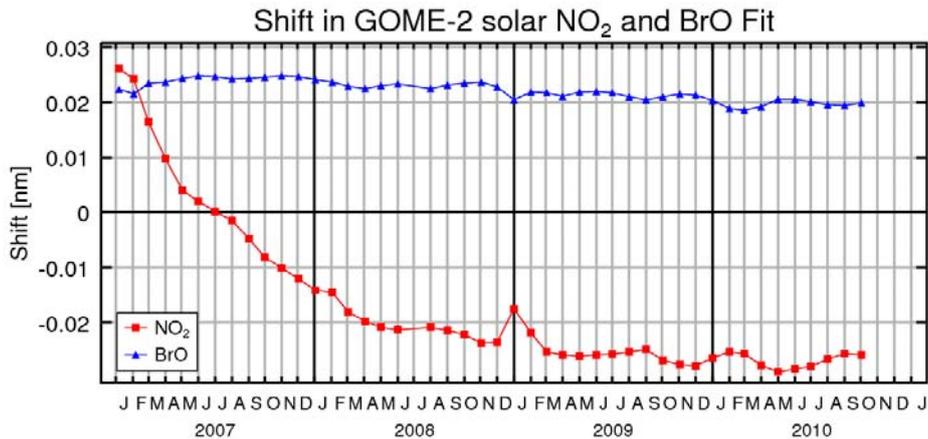
The same analysis as for NO<sub>2</sub> was also performed for BrO, and results are shown in Fig. 32. Similar results are obtained as for NO<sub>2</sub>, with non-zero columns retrieved, increasing  $\chi^2$  with increasing time-difference and a clear impact of the throughput test on the  $\chi^2$ . However, in contrast to NO<sub>2</sub>, the throughput test also lead to large changes in the apparent BrO columns, indicating a change in solar measurements after the test. Interestingly, this change is not abrupt as one would expect but rather extends over several months until the instrument appears to stabilise again. This is unexpected and should be investigated further.

In summary, there is clear evidence that the GOME-2 solar spectra change over time in spite of calibration, that the throughput test lead to additional changes in these spectra and that the changes have some spectral correlations with absorbers such as NO<sub>2</sub> and BrO.



**Fig. 32:** Evolution of apparent BrO column (top) and  $\chi^2$  (bottom) of a BrO DOAS analysis on daily solar irradiance measurements. An arbitrarily selected solar measurement from July 2008 is used as background spectrum in the analysis.

In the DOAS retrievals, an additional parameter is the wavelength alignment between Kurucz and GOME-2 spectra. In this test, the wavelength calibration coefficients provided in the operational product are not used so the observed changes reflect real instrument changes without calibration. The results of this analysis are shown in Fig. 33 for both the BrO and NO<sub>2</sub> fitting windows. As can be seen, there are very small changes in the BrO spectral region but a slow change of more than 0.05 nm in the NO<sub>2</sub> wavelength region which looks like a slow relaxation process.



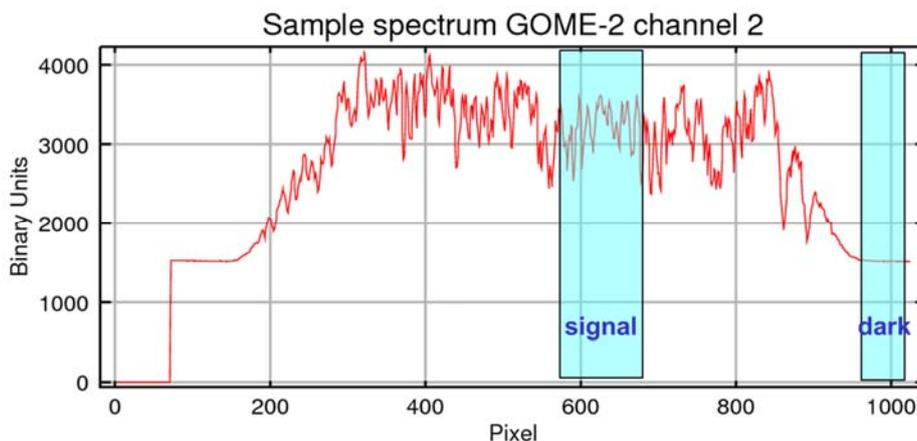
**Fig. 33:** Evolution of the spectral shift retrieved from the BrO and NO<sub>2</sub> retrievals on solar spectra. The Kurucz solar measurement is used as standard for the wavelength axis.

## 5 Throughput test (WP 4)

During the 2<sup>nd</sup> throughput test, detector temperatures were changed systematically in order to investigate how throughput and SNR depend on temperature. Several unexpected effects were observed including the absence of the strong hysteresis observed during temperature changes resulting from instrument switch-offs and a systematic and persistent decrease in overall instrument throughput during the period with the highest temperatures. Not all aspects of the throughput test are fully understood, but some interesting conclusions can be drawn from the data taken in this period.

### 5.1 Level 0 data

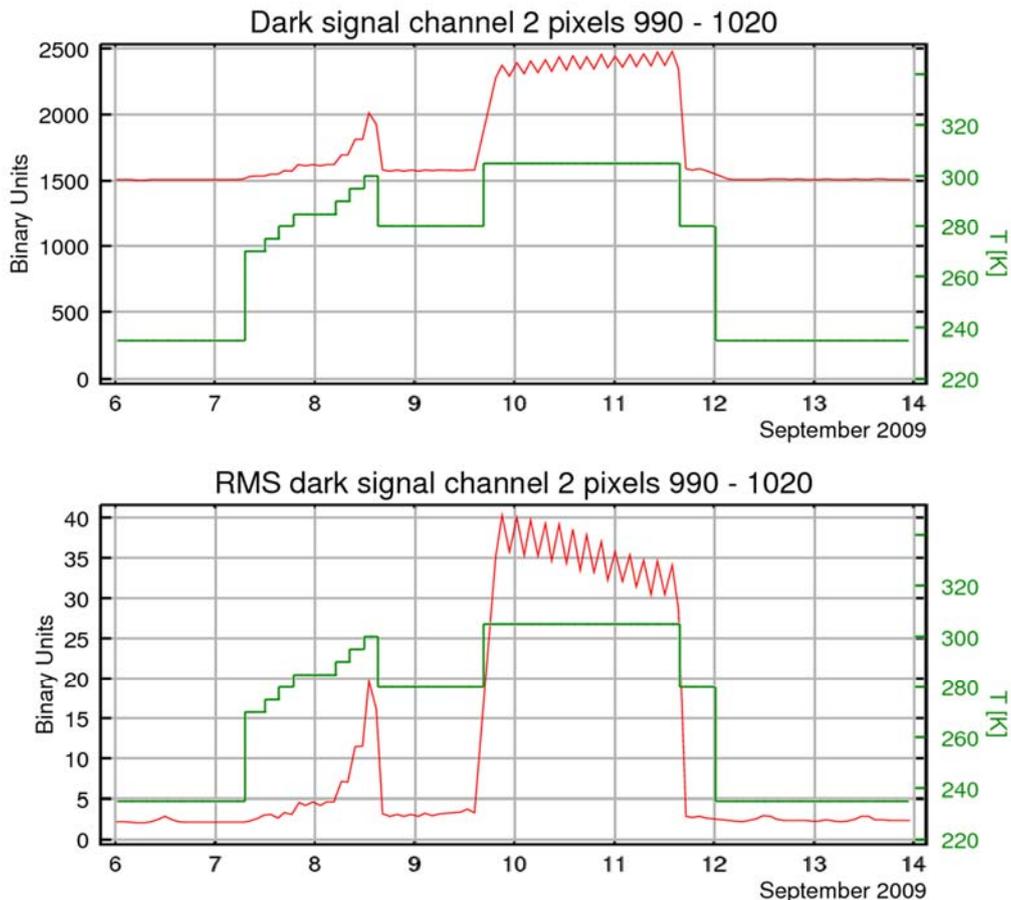
In the first step, raw (lv0) GOME-2 data are used to investigate the development of throughput and RMS of the measurements. While similar studies were performed by EUMETSAT on calibrated data using special observations, here raw data from the nadir observations are analysed. IN all the following figures, data from a complete orbit were averaged so that each point represents an average over a large number of observations taken at varying conditions. While some effects of different ground scenes and cloud conditions are still noticeable in the earth-shine data, they are largely averaged out.



**Fig. 34:** Typical spectrum in channel 2 taken during the 2<sup>nd</sup> throughput test. The shaded areas indicate the pixel ranges used to track the earth-shine signal and the dark signal.

First, two regions have been identified in the spectra to represent not illuminated and well illuminated parts of the detectors. In Fig. 34, an example is shown for channel 2. The dark region is used to investigate the dark signal behaviour, the illuminated parts to check for possible changes in throughput.

The region selected for dark signal analysis is unfortunately not completely dark. Comparison of the temporal evolution of the signal in this region over one orbit with that over the full detector indicates strong correlation in the variability of the two quantities. However, the absolute size of the additional signal in the dark region is very small and therefore can be neglected here.



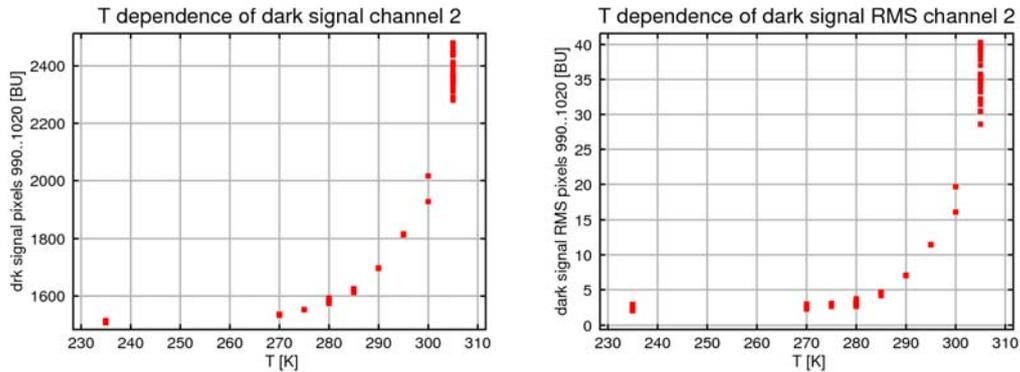
**Fig. 35:** Temporal evolution of the dark signal (top) and the RMS of the dark signal (bottom) over the 2<sup>nd</sup> throughput test. Also shown is the evolution of the detector temperature (in green).

As shown in Fig. 35, the dark signal closely follows the evolution of detector temperature as expected. There are two unexpected observations: i) during the phase with a temperature of 305 K, the dark signal appears to increase and ii) there is a saw-tooth signal on the temporal evolution. Both effects are currently unexplained.

Also shown in Fig. 35 is the evolution of the RMS of the dark signal. If the dark signal is the result of leakage current of the detector, it should be a random process and the RMS should be linked to the square root of the absolute signal following Poisson

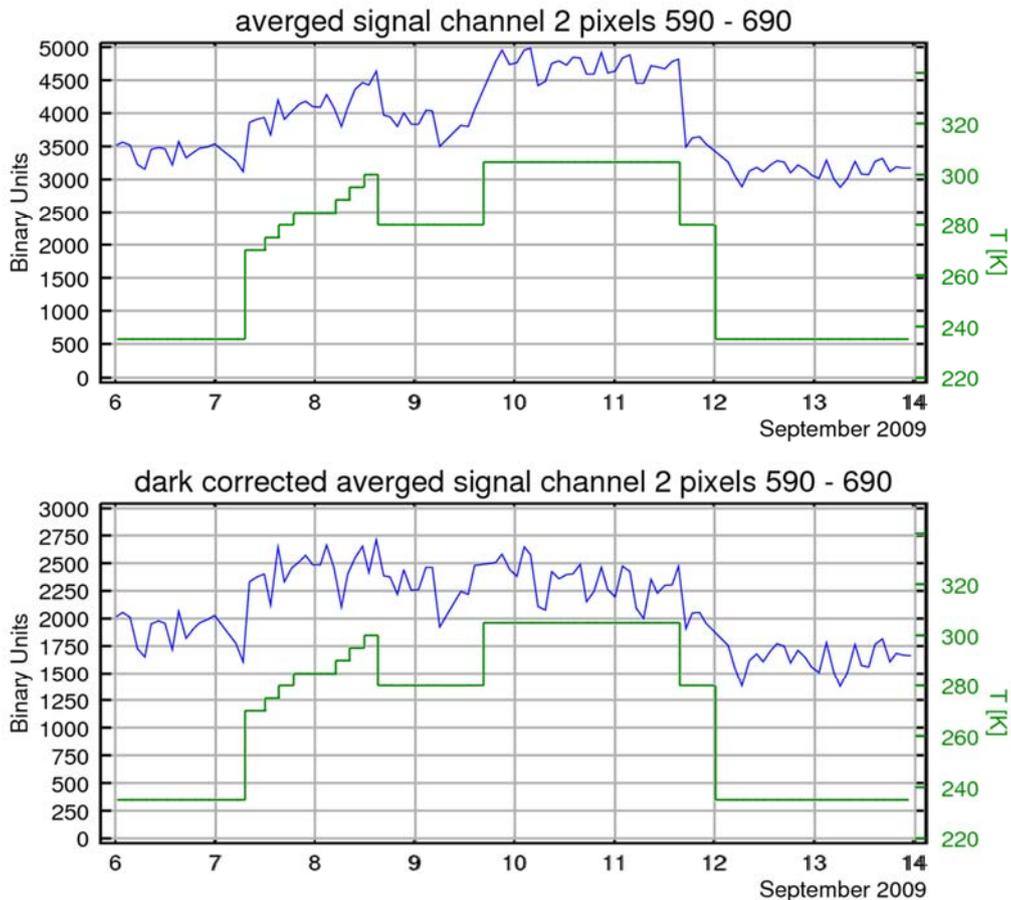
statistics. The electronic offset in contrast should not contribute to the RMS. With an electronic offset of about 1500 BU, this relation is more or less observed in the data.

During the period with measurements at 305 K, the behaviour of the dark signal RMS does not follow this simple pattern. Instead, the RMS decreases while the signal increases. This can only be understood if the non-random component of the signal (the electronic offset) increased at the expense of leakage current.



**Fig. 36:** temperature dependence of the dark signal (left) and the RMS of the dark signal (right)

The leakage current of a photodiode detector is expected to show an exponential behaviour as a function of temperature. Some measurements during the throughput tests appeared to indicate a deviation of the GOME-2 detectors from this behaviour. However, as shown in Fig. 36, analysis of the dark signal and its RMS as function of temperature show a clear exponential T-dependence as expected. The only exceptions are the measurements at 305 K and afterwards (one point at 300 K) as the dark signal increased and the RMS decreased over time independently of temperature in this period.

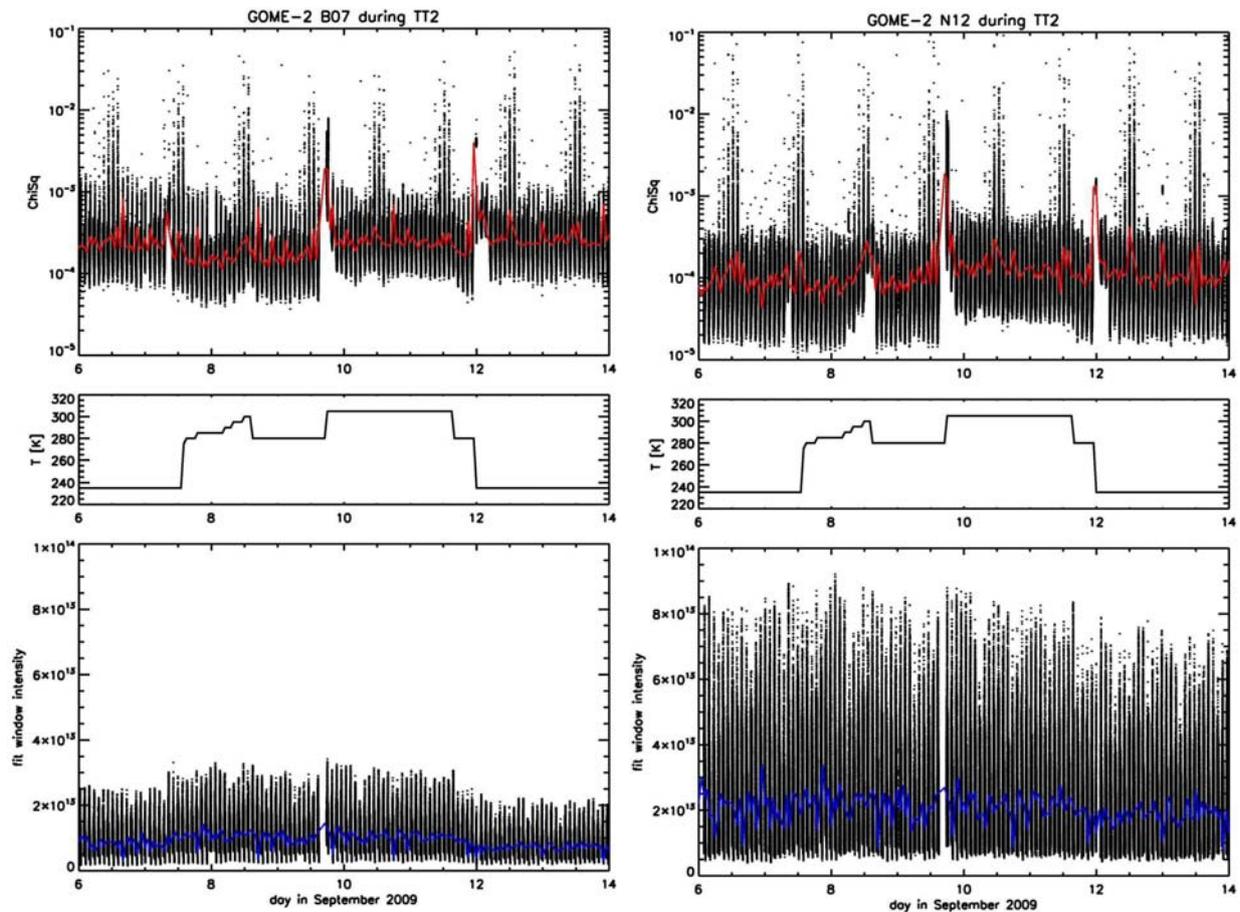


**Fig. 35:** Evolution of earth-shine signal during the 2<sup>nd</sup> throughput-test. In the upper panel, raw signals are shown while in the lower panel, the dark signal from the not illuminated part of the detector was subtracted to correct for dark signal.

As shown in Fig. 37, the temporal evolution of the raw signal from the earth-shine observations also follows the temperature evolution with larger values at higher temperatures. However, after correction for dark signal by simply subtracting the signal from the non-illuminated pixels in the same observation the picture looks different. Now, a throughput gain of about 30% at the beginning of the throughput test becomes apparent, followed by a slow reduction in throughput during the time period with 305 K and resulting in an overall loss of 15 – 20% of signal over the complete test.

## 5.2 Level 2 data

Level 1A and 1B from September 7-12, 2009 (2<sup>nd</sup> throughput test) was exclusively provided by Eumetsat to investigate the behaviour during this time period. The respective IUP retrieval algorithms for NO<sub>2</sub> and BrO were used to fill in this data gap in the level 2 time series. Fig. 38 shows the temporal evolution of the fit residuals (log-scale!) and the fit window intensity for BrO and NO<sub>2</sub>, respectively. Black dots indicate the information for a single ground pixel between 60°S and 60°N and the solid lines denote an hourly running mean. The plot in the middle shows the sensor temperature during this time period.

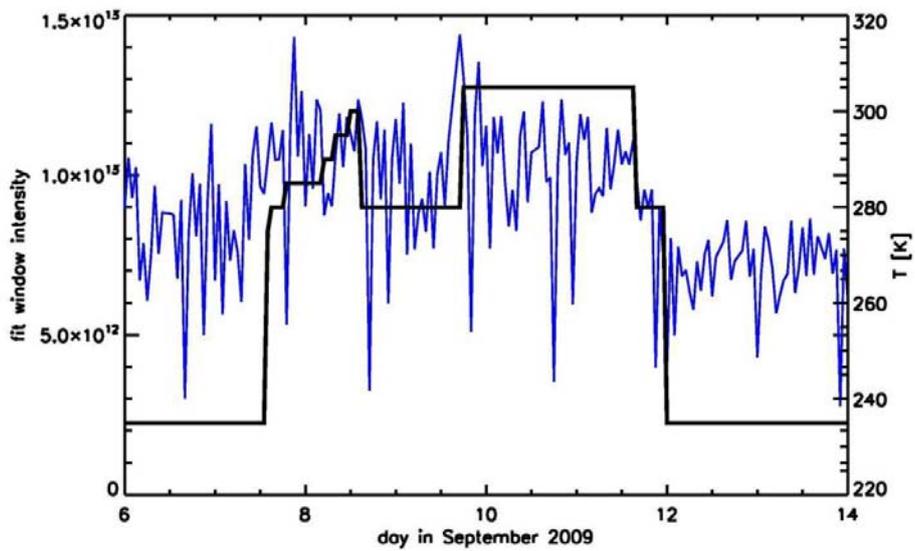


**Fig. 38:** Temporal evolution of BrO (left) and NO<sub>2</sub> (right) fit residuals (upper plot) and intensities (lower plot) during the 2<sup>nd</sup> throughput test. An hourly running mean (red and blue solid lines) was calculated using all data (pixel) between 60°S and 60°N. The sensor temperature is depicted in the middle.

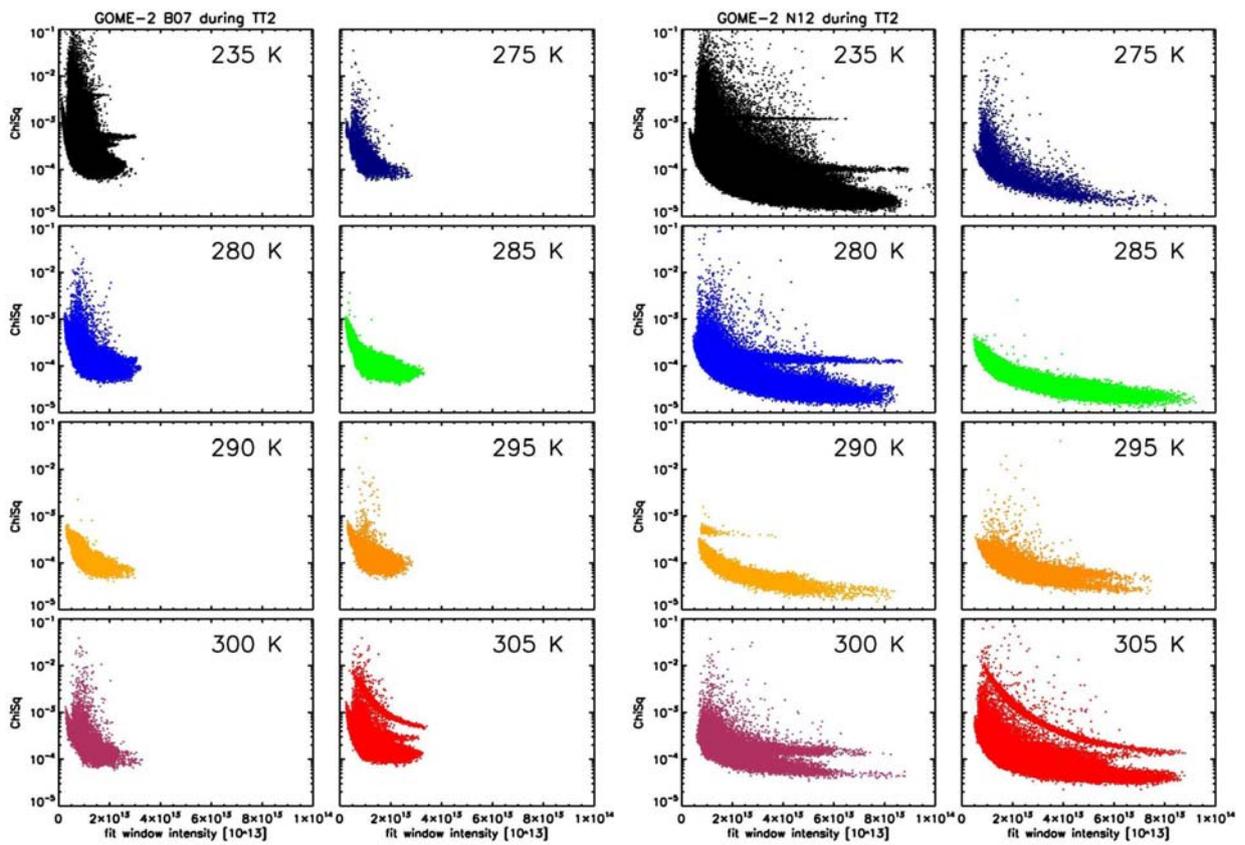
Measurements within the South American Anomaly (SAA) can be clearly identified by the increased  $\chi^2$  for 3-5 orbits once a day. Times with highly increased  $\chi^2$  are also found at the edges of the temperature step function which might be explained by key data that has not been adjusted to the new sensor temperature, yet. BrO, as a representative of channel 2, shows a stronger temperature dependent signature when compared to NO<sub>2</sub>, a representative of channel 3. For BrO  $\chi^2$  first decreases with increased temperature, but rises again when sensors were heated to 305 K. The overall throughput (here shown as fit window intensity) is better for channel 3 than it is for channel 2, as expected.

The correlation between the sensor temperature and the earth-shine signal as shown in Fig. 37 is also visible in Fig. 39 showing the intensity for the BrO fit window in channel 2 (zoom-in from Fig. 36, lower left plot). This indicates at a solid calibration done in the level 0 to level 2 processes.

Fig. 40 shows the  $\chi^2$ -intensity dependence for NO<sub>2</sub> and BrO additionally separated by sensor temperature for September 2009. Once again, a horizontal distribution at 235 K and 280 K and the distinct distribution at 305 K are probably due to uncorrelated key data. The general distribution follows the  $1/(\text{INT})^2$  dependence and does not change much with increasing sensor temperature. Judging from this type of depiction no favoured sensor temperature can be identified.



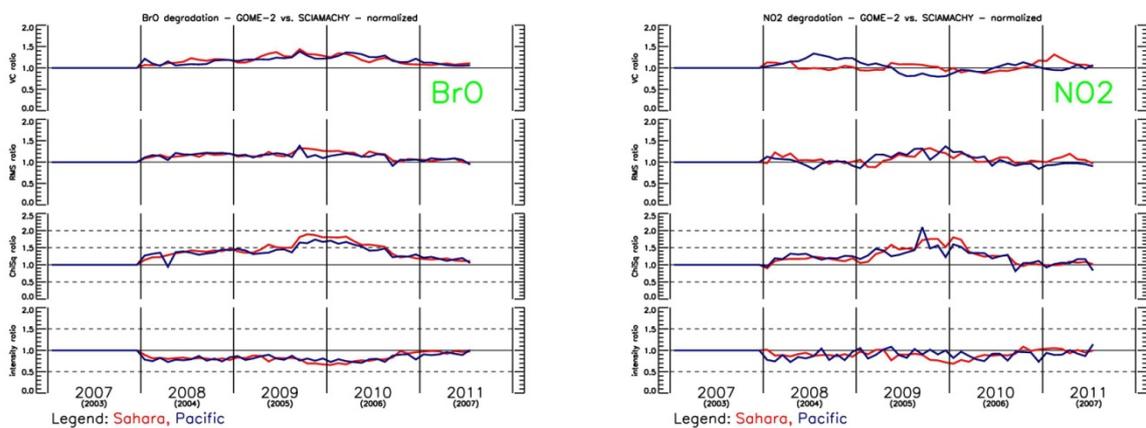
**Fig. 39** Zoom-in of Fig. 36 (lower left). BrO fit window intensity (blue) and sensor temperatures (black) are depicted.



**Fig. 40:** Fit window intensity dependent residuals  $\chi^2$  for September 2009 divided into bins of equal sensor temperature.

### 5.3 Changes in degradation rate after 2<sup>nd</sup> throughput test

Fig. 41 shows time series for BrO and NO<sub>2</sub>, respectively, normalized to the same month of the previous year. It is a measure of the degradation rate. Focussing on  $\chi^2$  and fit window intensity, one can easily identify the increasing loss of throughput and increase in RMS and  $\chi^2$  up to the 2<sup>nd</sup> throughput test. During the test, additional loss of throughput happened which lead to further increases in RMS and  $\chi^2$ , in particular for BrO. After the test, these values tended towards 0 again, which means that the loss of throughput has slowed down or even almost stopped in the case for NO<sub>2</sub>. Also, the increase in RMS has stopped and the increase in  $\chi^2$  slowed down (BrO) or stopped (NO<sub>2</sub>). If this trend continues, the overall effect of the 2<sup>nd</sup> throughput test might be considered to be beneficial for data quality in spite of the initial additional losses of throughput.



**Fig. 39:** Time series of BrO (left) and NO<sub>2</sub> (right) based on monthly means for the Sahara (red) and the Pacific (blue) normalized to the same month of the previous year.

### 5.4 Summary throughput test

The initial throughput gain was not proportional to temperature but rather appears to be linked to a quick instrument change during the time of the first (and modest) warm-up. The origin of this improvement in performance is not clear – it could either be linked to a gain in quantum efficiency of the detector at higher temperatures or evaporation of an absorbing layer on one of the optical components, probably the detector. While increases in quantum efficiency of diode arrays are expected at higher temperature, the rapid increase at relatively low temperatures and the loss in throughput at the end of the test which occurred already at significantly higher temperatures makes this explanation not very convincing. Therefore, a contaminating layer is currently the best explanation for the observations.

In addition, the question as to how the contaminant was able to reach the detector remains. Another possible explanation might be that the contaminant formed an absorbing layer on the optical surface opposed to the detector, which is in thermal radiative equilibrium with the detector. This optical surface is open to the optical bench and the electronic boards whereas the detector is enclosed in a tube.

## 6 Summary and Conclusions

In the following questions asked in the introduction and the beginning of this study will be tackled.

1. What is the effect of GOME-2 degradation on the accuracy (absolute values) of level 2 products?

The accuracy of the measurements can ultimately only be determined by validation with external data sources. Here, it was mainly accessed by comparison with SCIAMACHY data and by analysis of time series of areas where constant values are expected. For BrO and HCHO, large changes in absolute values are observed over time, making empirical corrections such as normalisation to a value over the Pacific mandatory. Unfortunately, no correction on the level of level 1 data could be designed so far. The quality of the fit is expressed in terms of  $\chi^2$ , i.e. the scatter of fit residuals over the respective wavelength region the trace gas retrieval is being performed for. This has an impact on the individual retrievals, but should (ideally) not lead to systematic biases in the averages of the derived columns. Results show a significant increase in  $\chi^2$  over the whole observation period of GOME-2. Predominantly trace gases retrieved in channel 2 (BrO and HCHO, but also SO<sub>2</sub>) suffer from the high loss of throughput known to be present in this wavelength region. Values for  $\chi^2$  have risen up to 5 times its original value (c.f. January 2007). Comparisons to SCIAMACHY data show a clear difference. In the same observation period (2007-2011) and during its first mission years (2003-2006), SCIAMACHY has a by far lower increase in  $\chi^2$ .

2. What is the effect of GOME-2 degradation on the precision (scatter) of level 2 products?

The precision of level 2 products has been investigated calculating the standard deviation (RMS) of all vertical columns within a chosen box of geo-locations. Results indicate that the loss of throughput leads to a systematic increase in the scatter of retrieved vertical columns. This increase is most pronounced for BrO. When compared to results from SCIAMACHY data, the increase in RMS for HCHO, O<sub>3</sub>, and H<sub>2</sub>O is roughly in the same order of magnitude, larger for NO<sub>2</sub> and much larger for BrO. After the 2<sup>nd</sup> throughput test, the increase in RMS has slowed down or even stopped, stabilizing the results.

3. Is the degradation dominated by throughput loss or are there also systematic spectral structures linked to instrument changes or degradation related calibration deficiencies?

The loss of throughput plays a vital role in the retrieval of trace gases. Especially the retrieval of trace gases with minor absorption features such as BrO and HCHO heavily rely on strong earth-shine signals to overrule the shot and detector noise of the instrument. Comparisons of the fit window intensity (INT) with SCIAMACHY show a much more rapid loss of throughput for GOME-2. This is the dominating factor for the reduction in fitting quality ( $\chi^2$ ). Yet, results from section 4.2 (Intensity dependencies) of this report strongly suggest that another reason must also play a role for the elevated level of fit quality decrease observed for the trace gases such as BrO and HCHO. NO<sub>2</sub> retrieved in channel 3 does not show these elevated levels as strongly. The year-to-year increase in  $\chi^2$  at the same fit window intensity cannot be explained with a loss in throughput alone. Results from a BrO retrieval done with earthshine background spectra show that the nature and origin of this systematic component can be found in the solar measurements. The systematic component

vanishes for earthshine background spectra used and the overall reduction in  $\chi^2$  is visible. In addition, the 2<sup>nd</sup> throughput test is also linked to the year-to-year decrease in fit quality. A sudden decrease can be seen before and after the 2<sup>nd</sup> throughput test. However, no correction approach could be advised so far. Whether or not this result helps to reduce the RMS (which is of direct relevance for data users) cannot be predicted.

4. Are there possibilities to correct for degradation effects on GOME-2 level 2 products?

First of all this question has to be separated into two parts:

- (i) What possibilities exist to ensure better level 2 products from data from the current flight model of GOME-2 (i.e. FM3)?
- (ii) What decisions have to be made to ensure better level 2 products from data that will be collected with upcoming flight models of GOME-2 (i.e. FM2 + FM1)?

This first question remains. Whether or not degradation effects as observed with the current flight model of GOME-2 can be dealt with the help of including additional calibration steps into the level 1A to 1B processor cannot be fully answered. Clearly, the dominant part of the degradation is explained by throughput loss and cannot be improved. The remaining systematic component appears to be most important for the  $\chi^2$  which is of secondary importance for data users. There is hope that this can be improved in the future, but what the impact on the quality of the columns is remains to be determined. The most practical correction scheme at this point is to use a normalisation of the data over a clean Pacific region as already applied to GOME, in some cases also SCIAMACHY data, and tested for GOME-2 BrO and NO<sub>2</sub>.

As for the remaining two flight models still hardware solutions have to be found, one of which is to make sure that contaminants within the instrument, once evaporated, can escape the enclosure into space.

5. What happened with GOME-2 level 2 products during the 2<sup>nd</sup> throughput test, and what can we learn from these results?

This question is closely linked to the questions 3 and 4. During the 2<sup>nd</sup> throughput test the performance of the instrument briefly increased, i.e. the throughput increased while parameters like the dark signal increased as well. The improvement could be linked to a gain in quantum efficiency of the detector at higher temperatures on the one hand or the evaporation of an absorbing layer on an optical component. Among both possibilities and judging from the behaviour of observed parameters an absorbing layer on an optical component is more convincing. The results also show that improvements for the upcoming flight models can still be implemented.

The rate at which the degradation occurred, especially for the accuracy of level 2 products (i.e.  $\chi^2$ ), slowed down after the second throughput test. For NO<sub>2</sub> it has reached a rate comparable to the one observed in SCIAMACHY data. Unexplained remains the increase for trace gases retrieved in channel 2 (BrO and HCHO). They still show some sign of degradation in the fit quality. Nonetheless, the level of throughput loss in channel 2 has reached critical levels.