GOME2 on MetOp-C

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

Final Report

Prepared by:

T. Bösch and A. Richter

With contributions from:

M. Weber, S. Noël, L. Alvarado and J. P. Burrows

University of Bremen

Institute of Remote Sensing

University of Bremen, FB1

P.O. Box 33 04 40

D-28334 Bremen

Document: **Final Report**

Issue: **Version 1.1** Issue Date: **11.03.2020**

Table of Contents

1 Executive Summary

1.1 Summary of activities

The aim of this study was to assess the quality of GOME2-C lv2 data products by comparing results with its predecessors GOME2-A/B. The following tests have been performed:

- A degradation analysis for O_3, NO_2, HCHO, OClO, BrO, SO_2, CHOCHO and H_2O
- A spectral stability test on channels 2 and 3.
- A detailed investigation of differences in NO_2 fitting results between GOME2-A, B and C
- A fitting residual analysis in the UV to evaluate the impact of residual structures on HCHO results
- A key data test for the HCHO fitting window in the UV to assess the impact on residuals and global maps.
- A high abundancy test for SO_2 on the example of the Raikoke eruption to show the performance in the UV at shortest wavelengths.

Disclaimer: Please note that the level-2 retrieval results presented in this document are not reflecting the performance of the official level-2 products available from the O3MSAF at https://acsaf.org/. Level-2 retrieval results presented here are tailored for studying the impact of the GOME-2 radiometric degradation. Operational GOME-2 level-2 products are often corrected for degradation effects, where applicable, and consequently show improved or different product stability and performance. For more details on the performance of the individual operational GOME-2 level-2 products please consult the validation reports and product quality monitoring pages available at https://acsaf.org/.

1.2 Main conclusions

- GOME2-C is very stable over the first year in terms of spectral registration, slit function and other calibration issues, in particular in channel 3. There is however a systematic oscillation in the wavelength of channel 3 which also affects the quality of the solar irradiance measurements.
- All products show a systematic dependence on subset for all GOME2 instruments, the sign and magnitude of the effect depending on the details of the fit settings.
- Channel 4 was analysed on the example of H_2O time series and shows good agreement with results from GOME2-A and B.
- Channel 3 shows reasonable and stable results for NO_2 and CHOCHO lv2 products. The impact of fitting window changes and different background spectra is similar for all three GOME2 instruments, with slightly larger residuals for irradiance background in GOME2-C.
- Channel 2 was found to create issues for the spectral range of standard HCHO lv2 results.
- Residual structures in the fitting range of HCHO were found but do not affect lv2 data strongly. The reason for these spectral features remains unclear.
- Spatial artefacts appear for the standard HCHO fitting window making the daily analysis of data impossible when not corrected or avoided.
- Good agreement was found between GOME2-A/B/C HCHO when changing to spectral ranges which are unaffected by spectral and spatial issues.

- Good agreement was found for OClO from GOME-2 and results from the other GOME2 instruments if a larger degree of the polynomial is used. For OClO, there is no indication for problems from the polarisation calibration.
- Inhomogeneous scenes (cloud edges, snow and ice edges) lead to systematic increases in the spectral shift for all products and to artefacts in the columns for OClO.

1.3 Main recommendations

- The origin of the wavelength oscillations in channel 3 should be investigated.
- The reason for larger residuals when using solar irradiance should be investigated.
- The problem in GOME2-B calibration highlighted in the section on OClO should be solved.
- The subset dependence found for all products and all GOME2 instruments should be further investigated and where possible corrected, for example by post-processing of lv2 products.
- For HCHO retrievals, some parts of the UV spectrum do always lead to spatial or spectral features which need to be corrected during the creation of lv2 products or as an additional post processing step. Correction can be done by changing the fit settings.
- Smallest wavelengths in the UV can be used for the SO_2 lv2 product and lead to similar results for GOME2-B and C but additional filtering will change the background level of low abundancy areas.

1.4 Overview in trace gas specific results

Table 1: Top level summary of GOME2-C trace gas results

Product	Summary
SO_2	works well, background offset
O_3	works well, offset relative to other GOME2- instruments
HCHO	very sensitive to wavelength settings and polarisation corrections, change in analysis settings needed compared to existing data sets
BrO	works well
OClO	works well, low polarisation dependency, lower columns than from other GOME-2 instruments, larger degree of polynomial needed
NO_2	works well, fitting window dependent offsets but not GOME2-C specific
CHOCHO	works well
H_2O	works well

2 Study setup

2.1 Introduction

This study builds on two earlier projects on evaluating degradation of the GOME2-A and GOME2-B instruments Dikty and Richter (2012) and Azam and Richter (2015), respectively. The focus of the present study lies on the following objectives:

- To assess possible effects of on-ground degradation of GOME2-C compared to the In-Orbit degradation of GOME2-A and B on lv2 data for the first year of operation.
- To compare the In-Orbit degradation of all three instruments for the first year of GOME2-C performance, 2019.
- To identify spectral residual structures related to e.g. calibration issues which might have a negative impact on lv2 products.
- To give recommendations for the use of key data and spectral ranges for the analysis to improve lv2 data and to minimize spectral residual structures.
- To give recommendations on improvements in calibration and lv2 processing to improve the lv2 data products

2.2 Definition of analysis

- The approach taken for this study is closely related to the previous degradation studies but focusses more on the improvement of lv2 data to reduce the effects of spectral, temporal or spatial issues of the instrument which might deteriorate lv2 products.
- GOME2 lv2 data products of IUP Bremen are used to investigate the impact of degradation.
- To do so, time series of GOME2-A/B and C lv2 data are created for different geo-locations and states of the instruments (first year of operation and the year 2019, where all three instruments provide data) which are compared to identify degradation and calibration effects.
- Subset depending changes of GOME2-C lv2 data are analysed and compared to GOME2-A/B.
- The effect of including various key data functions in the spectral retrievals is evaluated.
- The impact of spectral artefacts in fitting residuals on lv2 data quality is assessed.
- The results of the two GOME-2 degradation studies are used for comparison and discussion.

2.3 Selection of lv2 data products to be used

The choice of lv2 products includes all species analysed and discussed in the previous degradation studies for GOME2-A/B and additional lv2 data products as summarized in Table 2. Different channels are colour-coded. Channel 4 is analysed with respect to H_2O lv2 data. Channel 3 includes the NO_2 and CHOCHO data products. Since the light intensity is the smallest for channel 2 and possible degradation impacts more strongly at shorter wavelengths, five different species (SO_2, O_3, BrO, HCHO, OClO) absorbing in different ultraviolet spectral regions are selected in order to assess the overall data quality of lv2 products from channel 2.

Table 2: Data products and corresponding fitting windows.

Data Product	Spectral Range (nm)	Channel
SO_2	312.0 – 328.0	2
O_3	325.0 – 335.0	2
HCHO	323.5 – 361.0 330.25 – 356.25	2
BrO	334.6 – 358.0	2
OClO	345.0 – 389.0	2
NO_2	425.0 – 497.0	3
CHOCHO	433.0 – 465.0	3
H_2O	688.0 – 700.0	4

2.3.1 Selection of time periods to be analysed

So far, only one year of GOME2-C data is available. Data products presented in this study are analysed for one specific year or for one reference day. For the analysis of degradation effects, time series showing the current state of the instruments for 2019 are depicted as well as the early state of each instrument represented as the years 2007 (GOME2- A), 2013 (GOME2-B) and 2019 (GOME2-C).

The reference day of 20.03.2019 is used for specific tests in which for example the effect of adding different key data sets in the spectral retrievals is evaluated and fitting window dependences are discussed. The time period 22.06 – 30.06.2019 is analysed for SO_2 with respect to the eruption of the Raikoke volcano in Kamchatka.

2.3.2 Selection of regions to be analysed

Similar to the previous studies, a limited number of geo-location subsets are used for further analysis of time series. Regions without any sources for the respective species of interest are selected so that natural or anthropogenic variations can be excluded. This enables a good assessment of instrumental changes over time.

Two regions with a lower overall intensity are selected over the Pacific (25°S – 15°S and 150°W – 110°W) and the Sahara Desert (20°N – 30°N and 0°E – 30°E). Two further regions represent a high intensity regime comprising an “Ice-and-Snow-Box” scenario over Greenland (70°N – 75°N and 50°W – 30°W) and Antarctica (70°S – 75°S and 130°E – 150°E).

In addition to these low-abundance regions, two further geo-locations are introduced for the evaluation of HCHO and SO_2 lv2 data quality as high column scenarios. HCHO is analysed over Central Africa (70°S – 75°S and 130°E – 150°E) while SO_2 is discussed during the eruption of the Raikoke volcano in Kamchatka (30°N – 80°N and 130°E – 200°E). All selected regions are shown

as differently coloured boxes in Figure 2-1.

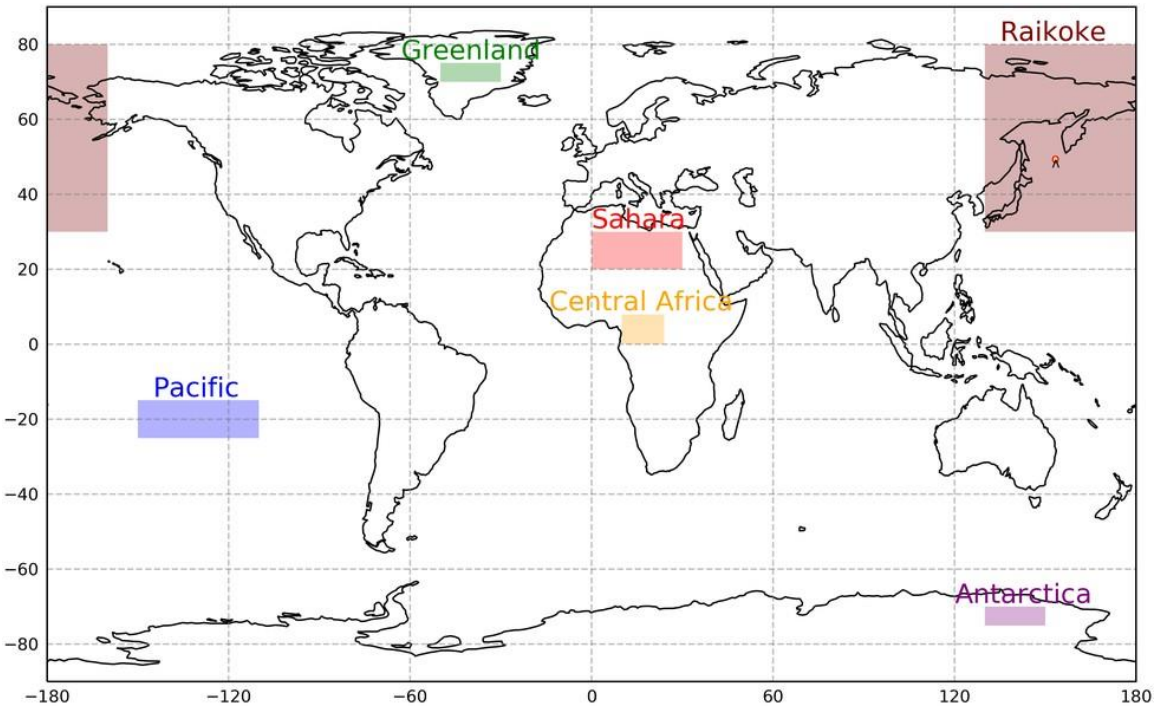

Figure 2-1: Subsets of geo-locations selected for the analysis.

2.3.3 Definition of diagnostics to be used

In this study, time series are analysed with respect to different diagnostics which are introduced in the following:

As the main quantity of interest, the vertical column density (VC) is shown for most species. It is derived from the retrieved slant column density (SC) by applying an appropriate air mass factor (AMF). The AMF values itself were calculated with the radiative transfer model SCIATRAN (Rozanov et al. 2014) assuming characteristic trace gas profiles and accounting for line of sight effects of the instrument.

The spread of the VC/SC values is quantified as the standard deviation (SD) which should stay more or less constant throughout the time series when the underlying assumption of measurement noise as being the dominant factor of lv2 data scatter is fulfilled. In addition to the spread, the mean of the corresponding distribution of data is used to find possible offsets in lv2 products.

The root mean square (RMS) of the residuals (difference between fitted X and measured μ optical depths) for a specific fitting window wavelength region is used to quantify the effect of degradation on the fit quality as residual spectral features should increase in frequency and amplitude throughout the aging of the instruments.

Table 3 introduces the equations used for the corresponding RMS of the residuals used within this study.

Table 3: Overview of different residual estimates used for GOME2 lv2 products.

Quantity	Equation	Trace gas
RMS of residuals	$RMS^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu_i)^2$	NO_2, SO_2, HCHO, BrO, OClO, CHOCHO
RMS of residuals	$RMS^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i)^2$	O_3
No residual information available		H_2O

3 Spectral stability

From the GOME2-A and GOME2-B instruments it is well known, that temperature changes of the optical bench induce systematic distortions of the spectral alignment of the diode array detectors during each orbit around the Earth and also during the orbit around the Sun. In addition to the changes in wavelength registration, there is also a change in instrument slit function. As equilibrium temperature changes throughout the lifetime of the instruments, there also is a long-term drift in both wavelength registration and instrument function.

As it is expected that similar effects occur for GOME2-C, the different aspects of spectral stability were evaluated for the first year of GOME2-C operation.

3.1 Slit function

Based on work first done for GOME, the FWHM of the instrument slit function can be determined from the solar spectra. This is achieved by starting with a high resolution Fraunhofer spectrum and using a non-linear fit to find the FWHM of the Gaussian function which when used to convolute the Fraunhofer spectrum creates the best representation of the measured GOME2-C spectrum. This procedure can be repeated for several small spectral fitting windows to also investigate the wavelength dependence of the slit function.

Some earlier work has also taken into account the asymmetry of the GOME-2 slit function by using two half Gaussian functions with different FWHM, but this was not investigated here.

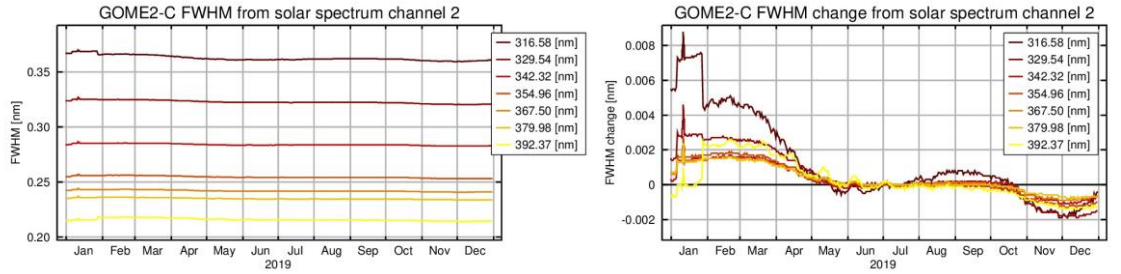

Figure 3-1: Evolution of the GOME2-C channel 2 slit function FWHM as derived from the solar spectra. Left: absolute values, right: changes relative to July 1st, 2019.

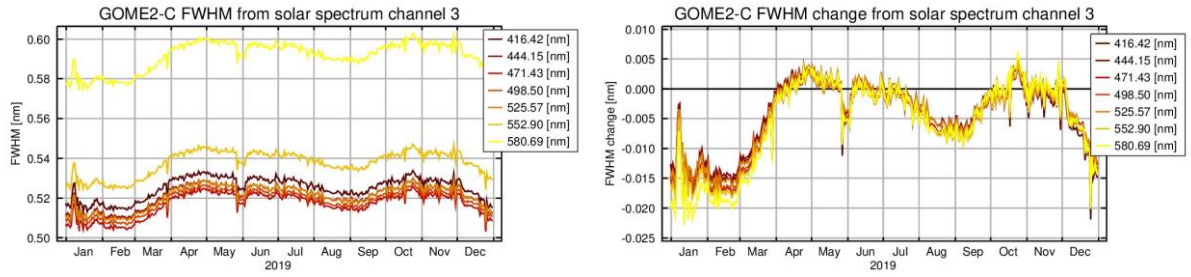

Figure 3-2: Evolution of the GOME2-C channel 3 slit function FWHM as derived from the solar spectra. Left: absolute values, right: changes relative to July 1st, 2019.

In Figure 3-1, the results of this analysis are shown for channel 2, both in absolute values and relative to the results for July 1st, 2019. As can be seen, there is a systematic variation of the slit function FWHM from shorter to longer wavelengths, with more than 0.36 nm (316.58nm) at the left edge of the detector and 0.22 nm (392.37nm) at the right edge. These values remain very stable over the first year with changes in the order of a few thousands of a nm. At the UV edge, changes are slightly larger, in particular in January when the instrument was not yet in its final configuration. Overall, the FWHM appears to decrease over time and is about 0.002 nm smaller

after one year in orbit for all wavelengths.

In Figure 3-2, results from the same analysis are shown for channel 3. Overall, there is a much smaller variation of the FWHM across the detector with values of 0.52 nm if the right edge of the detector (wavelengths larger than 550nm) is excluded. Variation of the FWHM over the year is however larger than in channel 2, and all wavelengths are affected in the same way. It should also be noted that there appears to be more day to day variability of the slit function, again affecting all wavelengths in the same way. Based on the first year of data, there is no indication for a long-term drift in FWHM.

3.2 Stability of wavelength registration

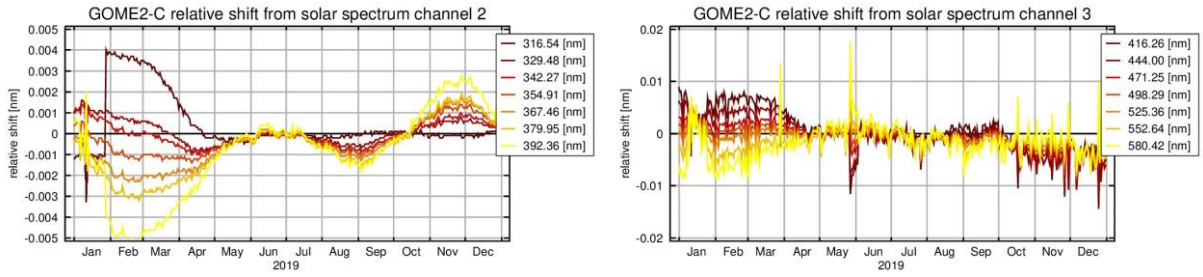

Figure 3-3: Evolution of the GOME2-C spectral registration through 2019 for channel 2 (left) and channel 3 (right). Values were retrieved from a Gaussian fit to the solar spectra and are shown relative to July 1st, 2019.

The same analysis as used for the FWHM also provides a value for the wavelength shift needed to apply to the GOME2-C spectra in order to find the best agreement with the high resolution Fraunhofer spectrum. As the IUP-UB data processor does not use the wavelength calibration information from the lv1 data but starts from an initial rough calibration, the values retrieved are shown relative to the results for July 1st, 2019. Any changes found are thus not related to calibration of GOME2-C but to physical changes in the instrument.

As shown in Figure 3-3, there is a pronounced variation of the shift over the year in channel 2, and it varies systematically with wavelength. The order of magnitude is a few thousands of a nm and the systematic spread of the results for different wavelengths indicates a moving in and out of the optical focus of the diode array. The very first wavelength in the UV appears to behave somewhat differently than the rest of the values for unknown reasons. Whether or not there is a long-term drift of the wavelength registration cannot be determined yet.

In channel 3, a somewhat similar picture is found with overall larger changes of nearly 1 hundredth of a nm but more importantly with large day to day variations, which as in the case of the FWHM appear to affect all wavelengths in a similar way. This points at a general problem with the wavelength stability of GOME2-C channel 3.

3.3 Along orbit wavelength stability

Wavelength stability can also be investigated along the orbit by applying the Fraunhofer fit described above also to the Earthshine spectra. This has already been done for GOME2-A and GOME2-B and systematic changes in wavelength registration and FWHM were found as result of the temperature changes along the orbit.

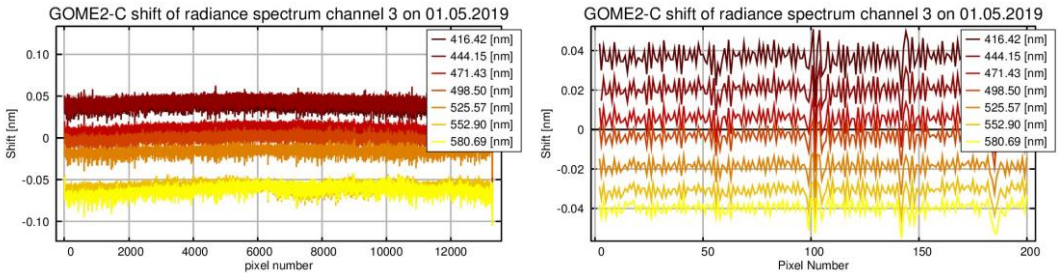

Figure 3-4: Along track changes in spectral shift, fitted for different wavelengths in channel 3 of GOME2-C for an orbit on May 1st, 2019. The left panel shows the full orbit while the right panel is a zoom into the first 200 measurements.

In the case of GOME2-C, results for channel 3 are unexpected in that a large (±0.02 nm) high frequency variability is found along the orbit as shown in Figure 3-4. When zooming in onto the detailed variation with ground pixel, it becomes clear that all wavelengths are affected in the same way, as if the whole diode array was vibrating. No such effect is found for channel 2.

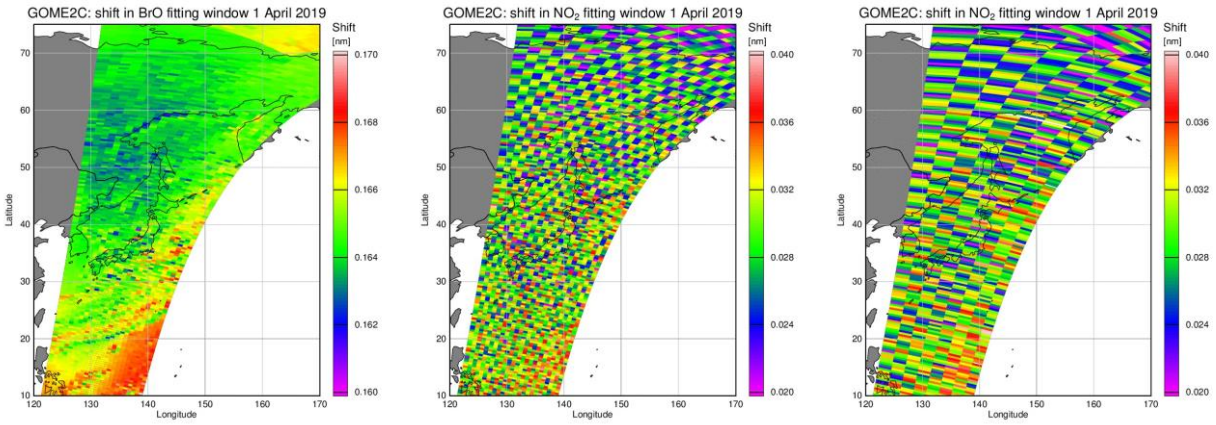

Figure 3-5: GOME2-C spectral shift of the forward scan pixels, fitted in channel 2 for the BrO retrieval (left) and in channel 3 for the NO_2 retrieval (middle).The right figure shows the results for the back scan pixels in channel 3. All data are for one orbit on April 1st, 2019.

This effect is also very pronounced when looking at spatial maps of the spectral fit retrieved in the DOAS analysis. As shown in Figure 3-5, the spectral shift is smooth over the orbit in a channel 2 retrieval (in this case BrO) with the exception of some sharp lines linked to inhomogeneous scenes (see section 5.5.4). In channel 3 however, a marked chess-board pattern is found, with higher and lower values found across track and periodically changing also along track every couple of scanlines. Similar patterns are found for forward and backward scans.

This periodic variation explains the apparent noise on the retrieved shift seen in Figure 3-4, and also the larger variability of the retrieved shift and FWHM in Figure 3-2 and Figure 3-3 if one assumes that also the irradiance measurements are affected by these periodic variations. So far, no convincing explanation for the root of this problem has been found. Mechanical vibrations appear implausible leaving electronic oscillations as a possible explanation, but how that could affect the wavelength registration of a full channel is difficult to understand.

4 Degradation and time series analysis

Time series for eight different species (H_2O, NO_2, CHOCHO, BrO, HCHO, OClO, O_3, SO_2) covering Channels 2 – 4 of GOME2-A, B and C were created to analyse temporal degradation effects. The comparison of all three instruments for the early state of the instruments (G2-A: 2007, G2-B: 2013, G2-C: 2019) and for the same measuring time (2019) reveal how In-Orbit degradation differs from the degradation effects in a safe storage environment with stable atmospheric conditions as applied to GOME2-C before launch.

A data set of lv2 products of the before introduced key parameters (VC/SC, SD, RMS), geo-locations (Pacific, Greenland, Sahara, Antarctica, Central Africa and the Raikoke region) and all three instruments is performed on 7-day means for statistical reasons.

4.1 Time series comparison of GOME2 on MetOp-A, B and C

The time series in the following sub sections (Figure 4-1 - Figure 4-14) always show a low (Pacific, Sahara) and a high intensity regime (Greenland, Antarctica) for the early state of the instruments and also for the year 2019. Each Figure shows the VC (or SC) in the first row, the scatter of the VC values represented as standard deviation in the second row, and the RMS of the residual in the last row (not available for H_2O).

In addition, two selected regions (Central Africa, Raikoke region) show high column cases for HCHO and SO_2, respectively, to assess the impact of residual structures and calibration issues in Channel 2 on lv2 data products.

All time series were filtered to remove scenarios with low sun (SZA > 88°) and poor fits (RMS > 0.01).

4.1.1 NO_2

In Figure 4-1 and Figure 4-2, time series are shown for NO_2 vertical columns, the scatter of the NO_2 vertical columns and the fitting RMS for 4 different regions (Pacific, Sahara, Greenland, Antarctica). In the left column, data for the three GOME2 instruments are compared for their respective first year of operation (2007, 2013, 2019) while in the right column the values for 2019 are compared. As the selected regions are away from anthropogenic sources, the columns represent stratospheric NO_2 and should show the stratospheric seasonality (little variation in the tropics, a maximum in summer and low values in winter hat higher latitudes) with little variation between years.

As can be seen in the figures, vertical columns between the three instruments agree quite well with a low bias for GOME2-A and a high bias for GOME2-C relative to GOME2-B. Part of these differences are slant column offsets as discussed in section 6.1.2.

The scatter of values is very similar for all three instruments when comparing the first year of operation. When looking at 2019 data, GOME2-C has lower scatter as expected while there is little difference between G2-A and G2-B results. This increase in scatter over the lifetime of the satellites is linked to reduced throughput and thus reduced SNR.

As similar result is expected for the fitting RMS, but here the results for the first year are different: GOME2-A has clearly the lowest residuals even today, while GOME2-B shows the poorest fits in the first year. Results for 2019 are similar between G2-B and G2-C indicating that some of the G2-B problems are calibration related and were fixed in more recent lv1 data versions. Overall, these differences result from calibration issues as the ones discussed in section 6.1.1.

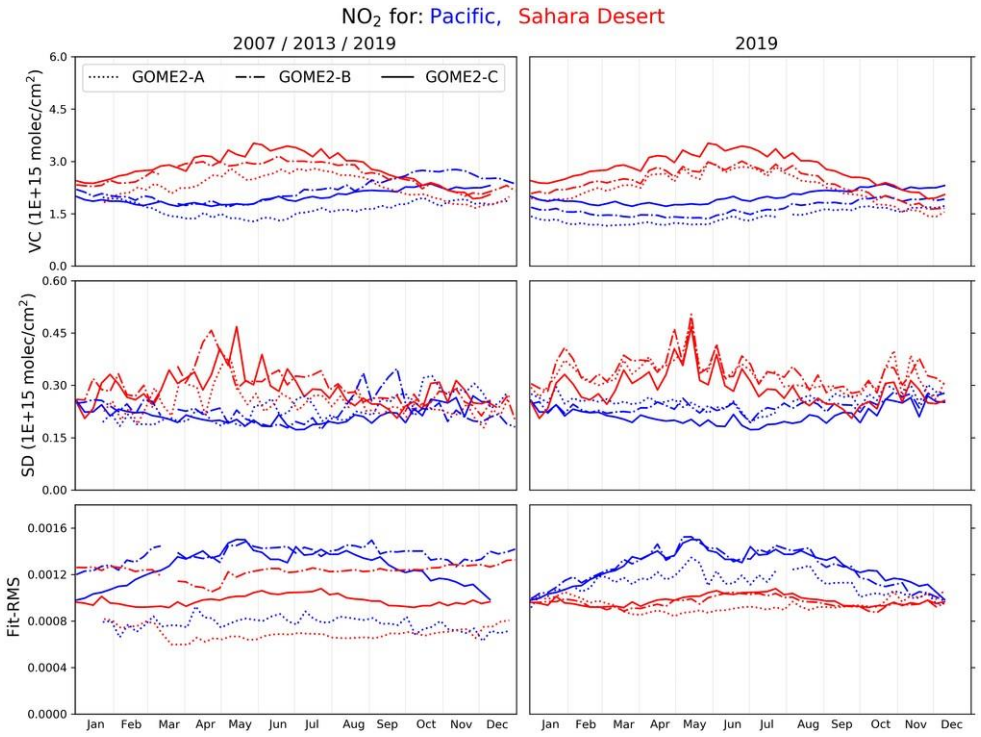

Figure 4-1: NO_2 time series of GOME2 on MetOp-A, B and C of VC, SD and Fit-RMS (top to bottom) for the early years of the instruments (G2-A: 2007, G2-B: 2013 and G2-C: 2019, left) and for 2019 (right). Different line styles show G2-A (pointed line), G2-B (dashed line) and G2-C (solid line), respectively. The time series are depicted for the Sahara Desert (red) and the Pacific (blue).

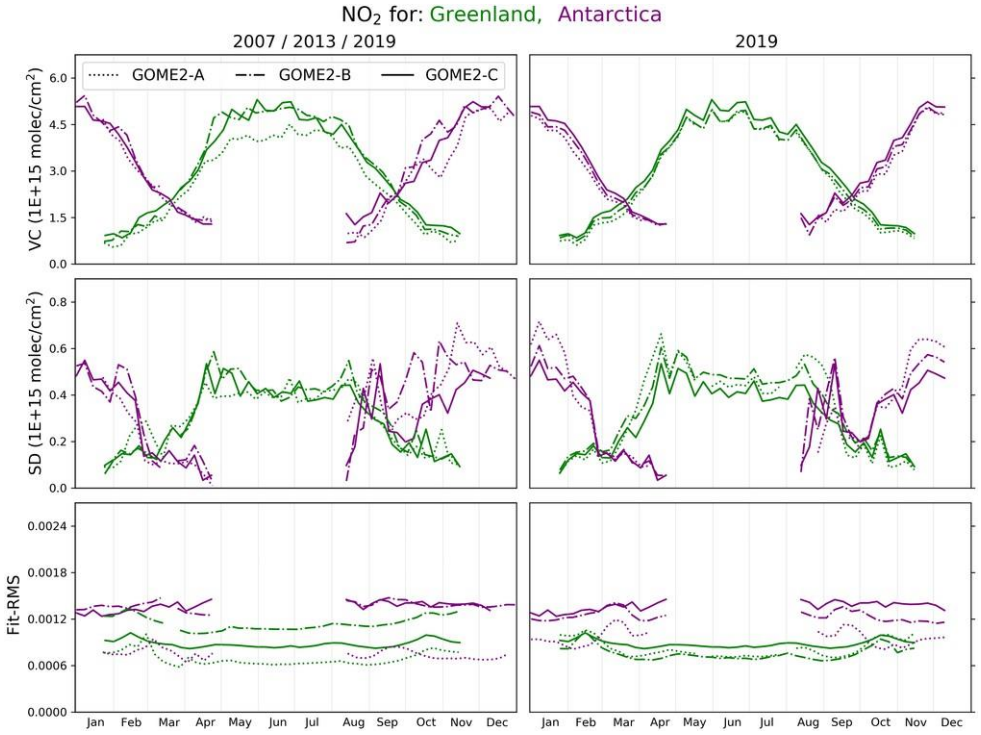

Figure 4-2: Same as Figure 4-1 but for Greenland (green) and Antarctica (purple).

4.1.2 CHOCHO

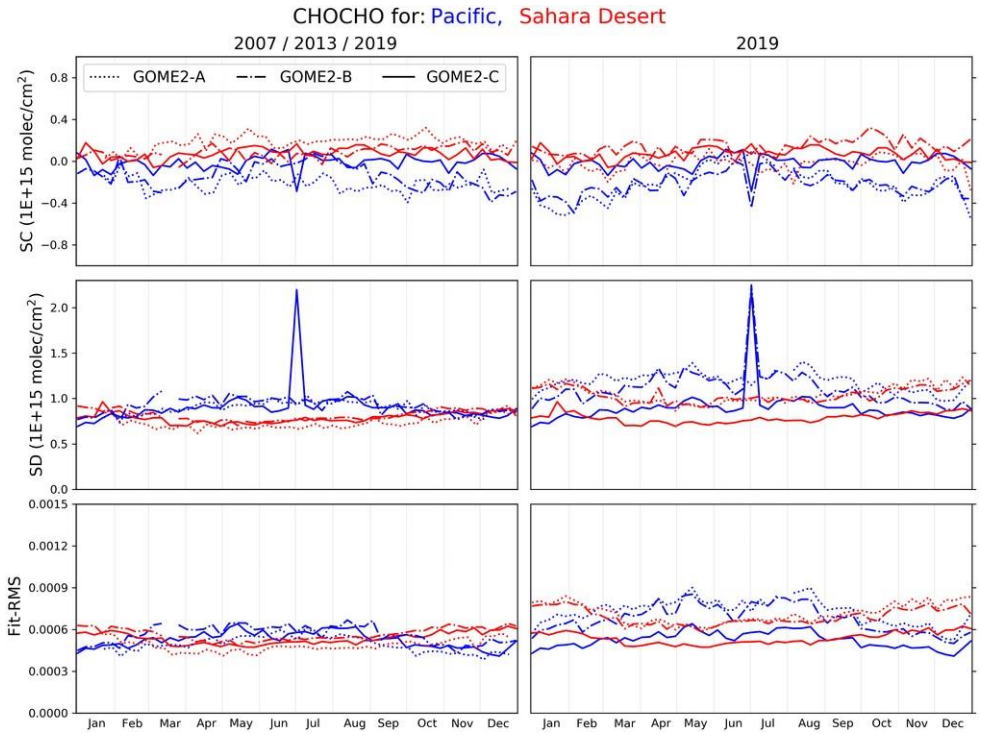

Figure 4-3: Same as Figure 4-1 but for CHOCHO.

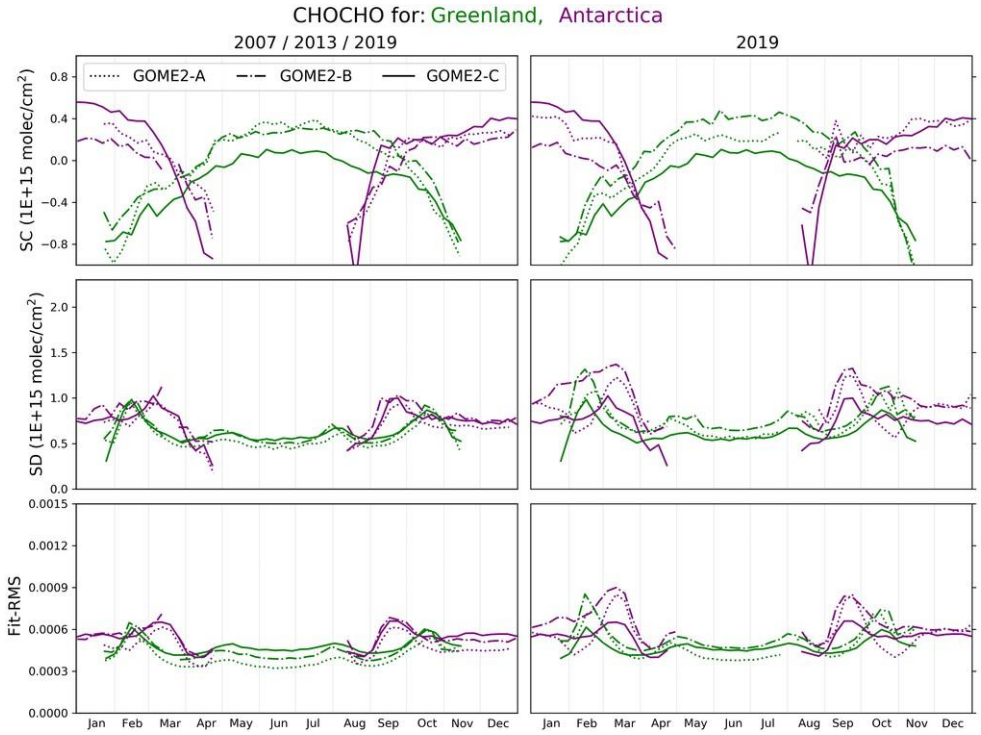

Figure 4-4: Same as Figure 4-2 but for CHOCHO.

When comparing the SC time series for 2019, good agreement is found for glyoxal (CHOCHO) over the Sahara. The time series over the Pacific are similar for GOME2-A and GOME2-B but GOME2-C shows less seasonality. Over the bright high latitude targets, there are clear differences between the time series of the 3 instruments, with inverted order in the two hemispheres. This is indicative of

hemispheric biases for GOME2-B and GOME2-C. It should be noted that values around 0 are expected at both target regions throughout the year, so any deviations from that are considered to be artefacts.

Concerning the results for the scatter of values, very good agreement is found between the three instruments for their respective first years of operation, GOME2-A having slightly better performance. Comparing the results for 2019, the degradation of both GOME2-A and GOME2-B is apparent and of comparable magnitude in spite of the large difference in time since launch. This is in agreement with other observations showing reduced degradation of GOME2-A in recent years.

For the fitting RMS, similar observations can be made as for the scatter of values.

4.1.3 BrO

The BrO vertical column time series show values close to zero for the Pacific with a slight increase at the end of all years. GOME2-B and C agree well but G2-A shows lower and over the Pacific even negative values in 2019 most likely due to the smaller swath and the larger noise of the instrument, especially in the UV. The SD and RMS curves show larger values for G2-A as well, followed by G2-B and C in 2019 as expected and good agreement between all instruments for their early state of operation. Although differences are small, GOME2-C appears to have the best performance in terms of both scatter and fitting RMS over dark scenes. The reason for the slight increase at the end of all years is unknown but only present for BrO indicating issues in the lv2 product rather than the instruments.

The time series for the geo-locations of Greenland and Antarctica show a similar good agreement with the exception of G2-A over Greenland in 2019 when the VC values are smaller in winter but higher in summer. SD and RMS are higher for G2-A when less light is present in winter and on a similar level for the remaining months during the early state of operation. In 2019, G2-B/C show good agreement and low SD and RMS as smallest curves for G2-C.

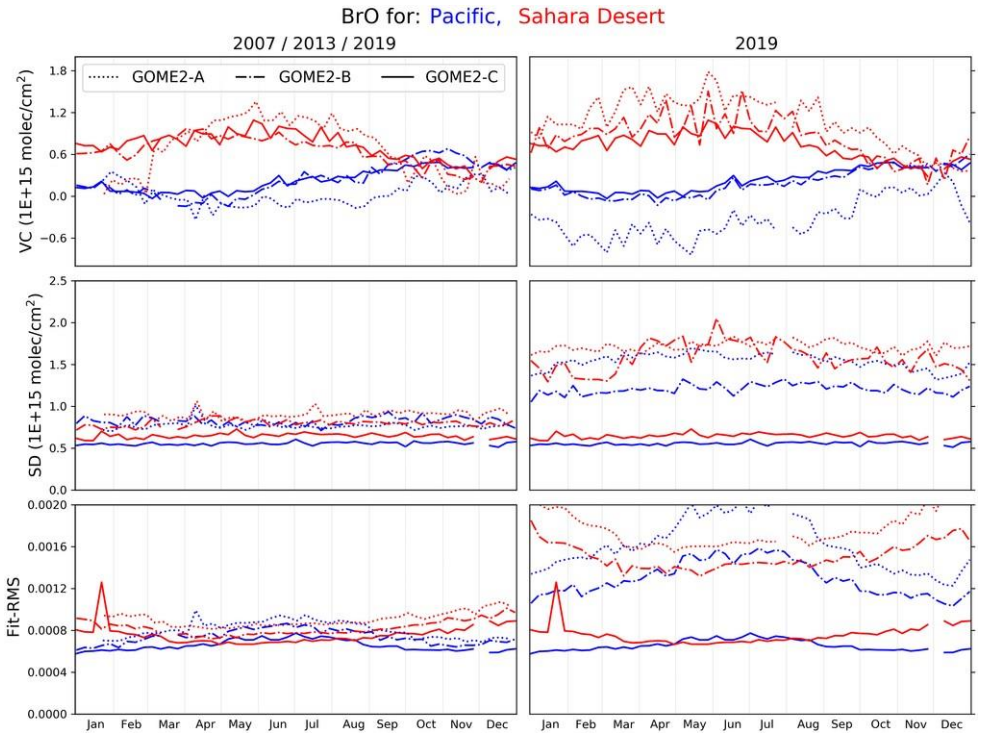

Figure 4-5: Same as Figure 4-1 but for BrO.

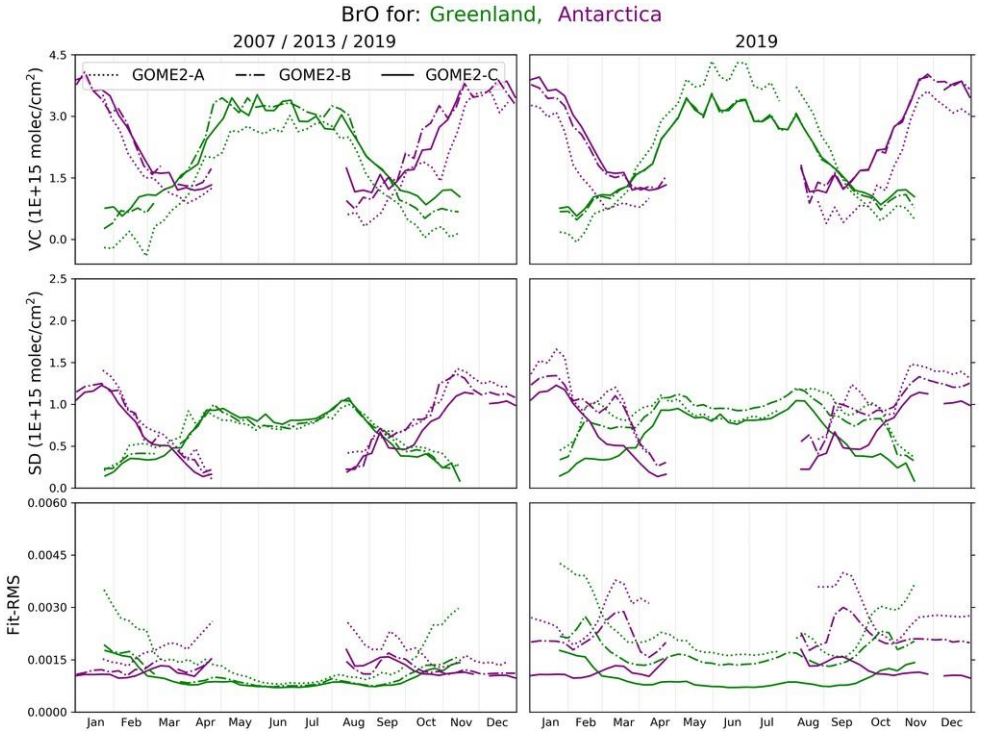

Figure 4-6: Same as Figure 4-2 but for BrO.

4.1.4 O_3

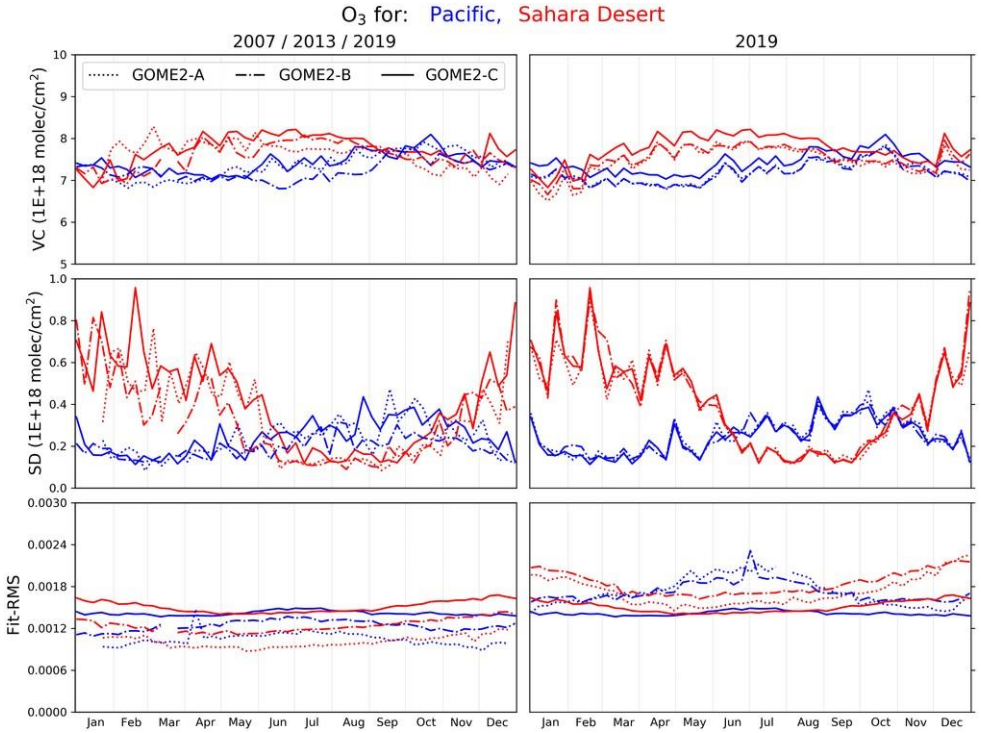

Figure 4-7: Same as Figure 4-1 but for O_3.

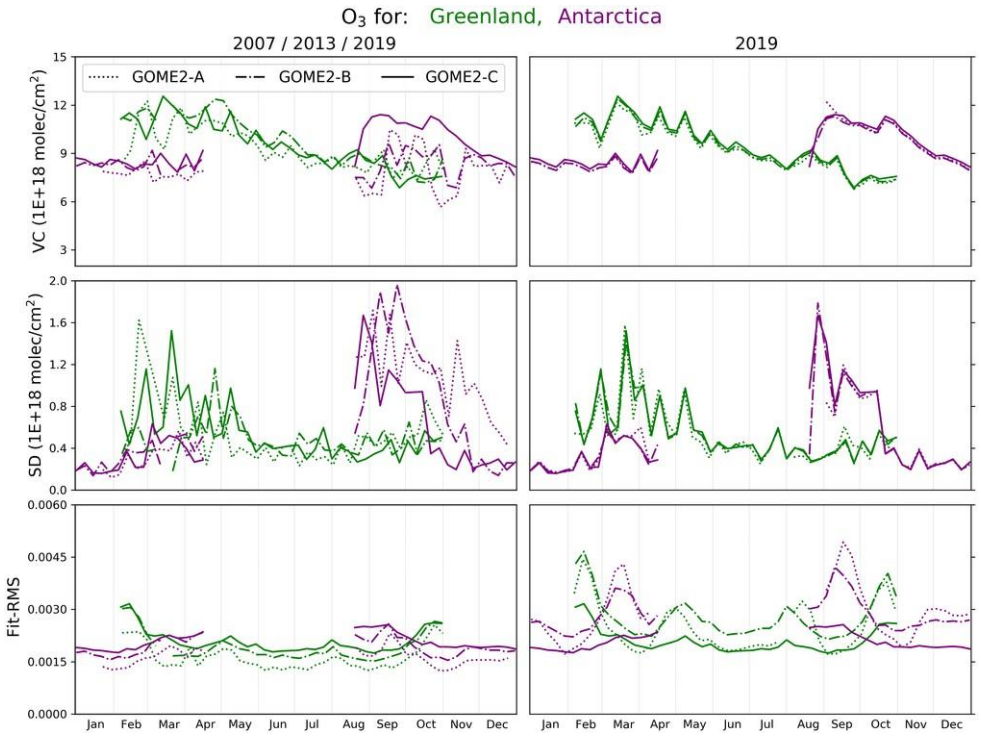

Figure 4-8: Same as Figure 4-2 but for O_3.

For O_3, it can be seen that the Fit-RMS is the smallest for GOME2-A in its first year of measurement, followed by GOME2-B and C. For parallel measurements in 2019, GOME2-C shows the smallest residuals compared to the other instruments as expected. The VC values are comparable with a similar spread of data points for all three sensors. The VC time series for parallel measurements show a nearly perfect agreement for all geo-locations but slightly larger values for G2-C for the low intensity regimes (Pacific, Sahara Desert) for unknown reasons.

4.1.5 H_2O

The agreement of VC and SD time series for all geo-locations is excellent indicating the good stability over time of channel 4 for all instruments and for strong absorbers. The early years of each instrument do also agree well with some smaller deviations which are due to varying H_2O in the corresponding years rather than to differences in the instruments. Smaller oscillations of SD have a similar magnitude for all instruments showing the small spread of VC values. Note that no residual information is available for this lv2 product. The corresponding lv2 product error is shown in this case. The errors are on a similar level but G2-C shows slightly higher values for all regions.

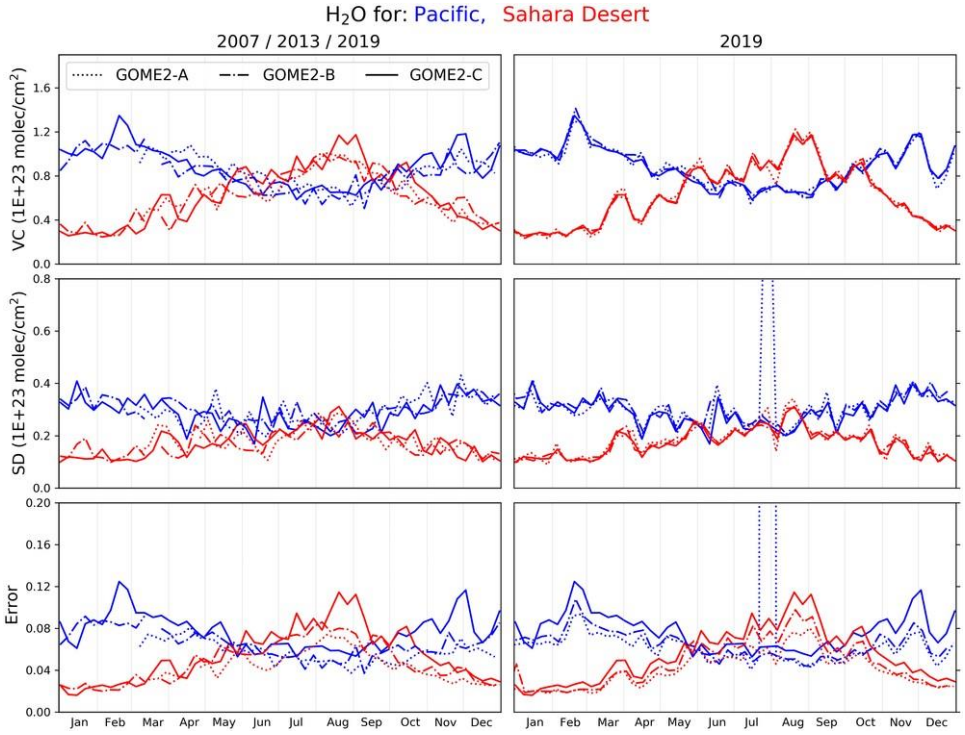

Figure 4-9: Same as Figure 4-1 but for H_2O.

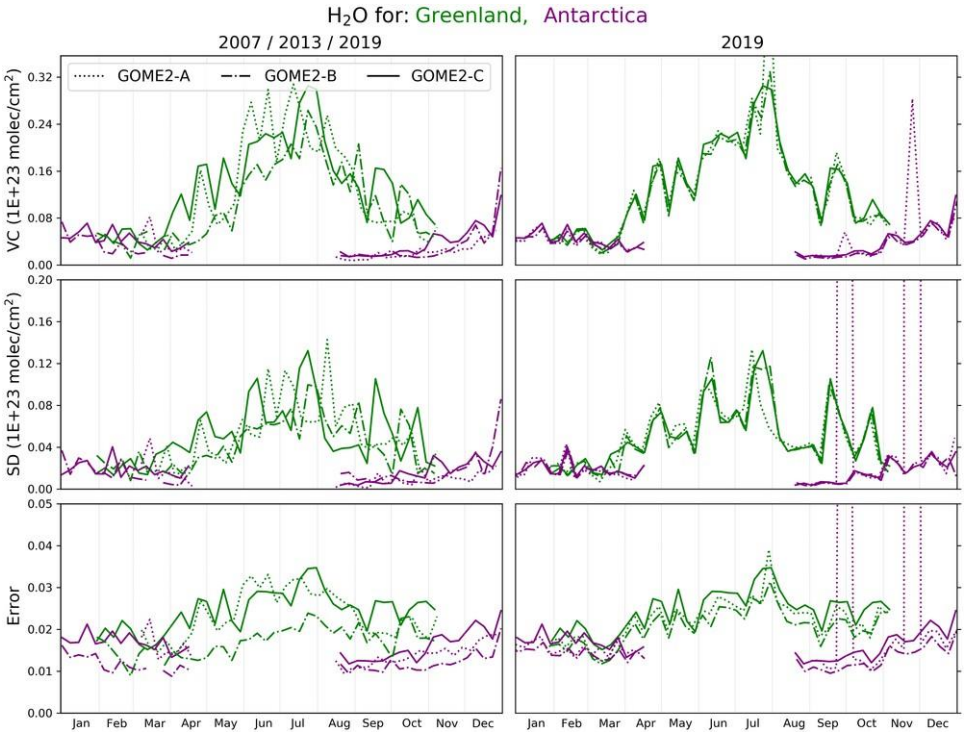

Figure 4-10: Same as Figure 4-2 but for H_2O.

4.1.6 SO_2

During the first year of operation, all three instruments show SO_2 VC time series close to zero for the Pacific and the Sahara Desert. RMS and SD are the smallest for G2-C, followed by G2-B and A. In 2019, G2-A suffers from strong oscillations indicating instabilities due to the reduced throughput propagating into the retrieval and leading to a much larger SD and Fit-RMS compared to the other instruments. Interestingly, the VC time series for Greenland seems to

have an offset of up to 100% for G2-A compared to G2-B/C in 2019 and also the early years of operation. This cannot be explained so far and might indicate issues with the DOAS retrieval rather than the instrument. However, the agreement for G2-B and C is good for all geo-locations with a slightly smaller SD and RMS for G2-C, compared to G2-B. The overall tendency is that larger deviations of the VC values from zero and higher SD and Fit-RMS can be seen when less light is available in winter.

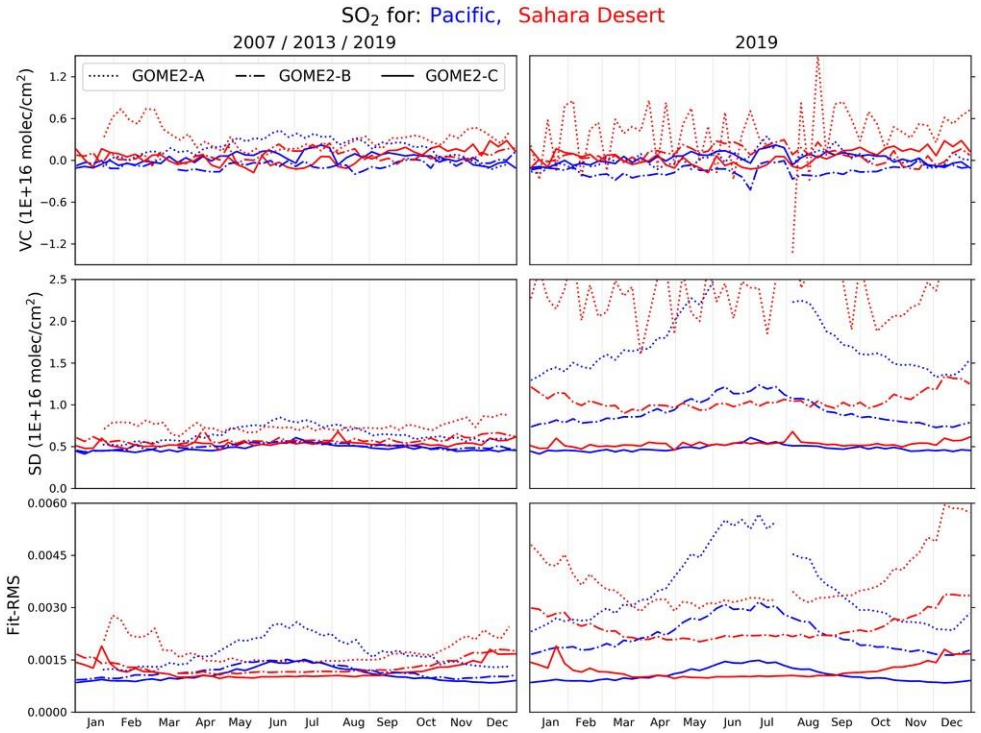

Figure 4-11: Same as Figure 4-1 but for SO_2.

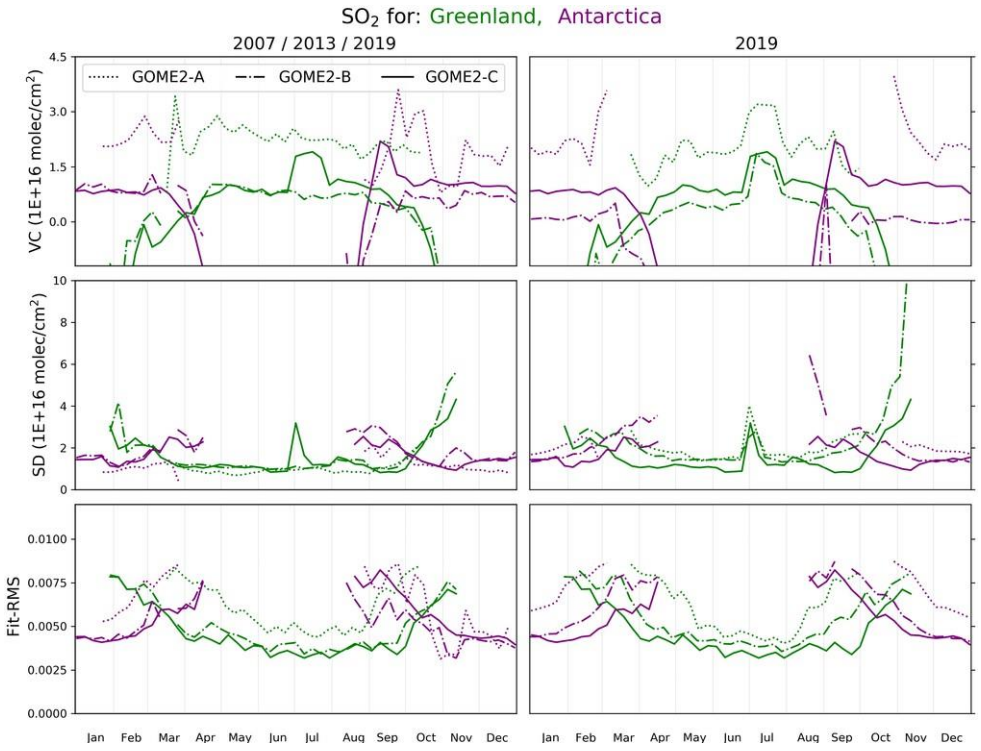

Figure 4-12: Same as Figure 4-2 but for SO_2.

4.1.7 HCHO

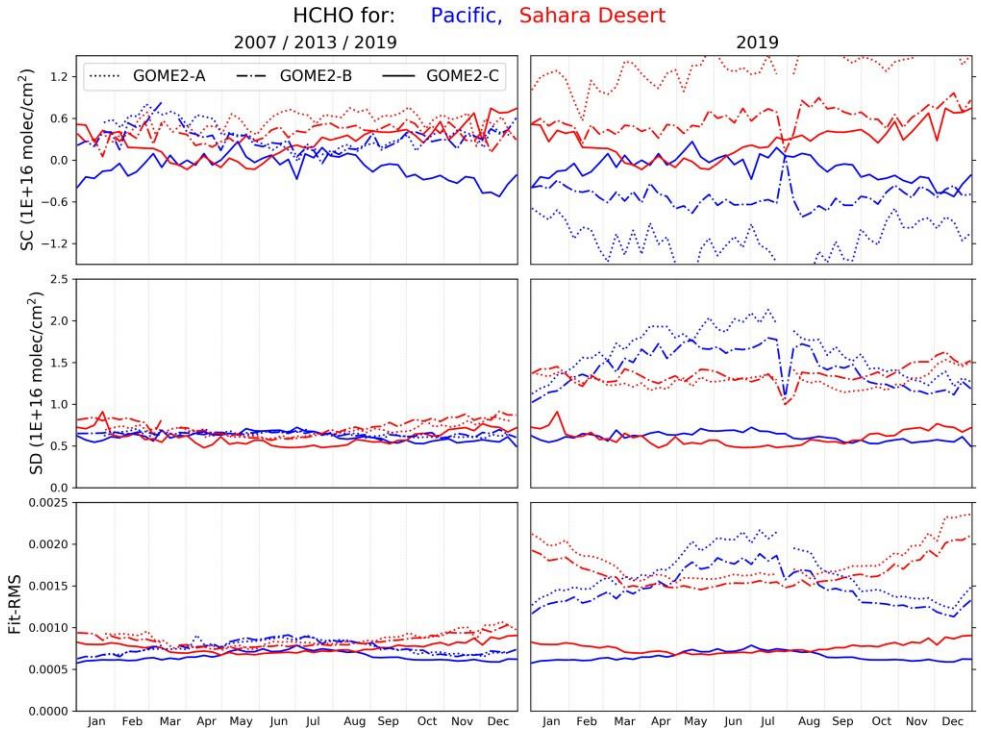

Figure 4-13: Same as Figure 4-1 but for HCHO.

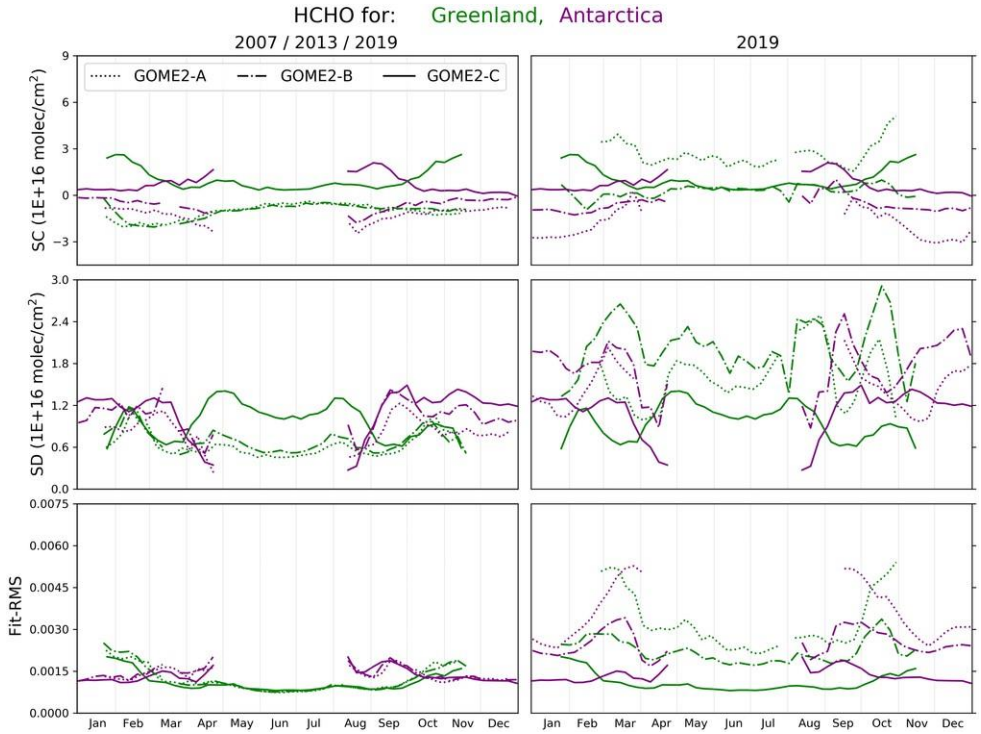

Figure 4-14: Same as Figure 4-2 but for HCHO.

The HCHO VC time series show a good agreement for the early state of the instrument with similar values for GOME2-A and B for all regions and a small offset of G2-C compared to G2-A/B. Since all three instruments show VC values around zero, the agreement can be considered as good for this lv2 product. In 2019, all instruments differ strongly for the Pacific, the Sahara

Desert and Antarctica while there is a better agreement for G2-C and B over Greenland. Since G2-A shows larger offsets than G2- B, this could be explained by less throughout and the higher noise level of the instruments in the UV. However, these offsets were not found for BrO in a similar spectral range indicating further issues. Interestingly, the offsets are negative for the Pacific and Antarctica while they are positive for the Sahara Desert. The SD and Fit-RMS time series for the early state agree quite well for all regions with smaller differences between G2-C and A/B for the SD values over Greenland. In 2019, SD and Fit-RMS differ more strongly with G2-C as lowest curve for all regions. Further analyses in Section 4 deal with the inconsistent behaviour of HCHO VC time series in 2019.

4.1.8 OClO

Retrievals of OClO have proven to be very sensitive to degradation and to noise. They also are prone to offsets which usually are corrected in the lv2 retrievals. Here, we use uncorrected OClO retrievals with daily solar irradiance backgrounds as this will highlight any possible degradation effects. As OClO is only present in the atmosphere under certain conditions (activated polar vortex, some volcanic eruptions, some lee waves in high latitudes), most of the results should be close to 0 and high values should mainly be found in Antarctica in early spring.

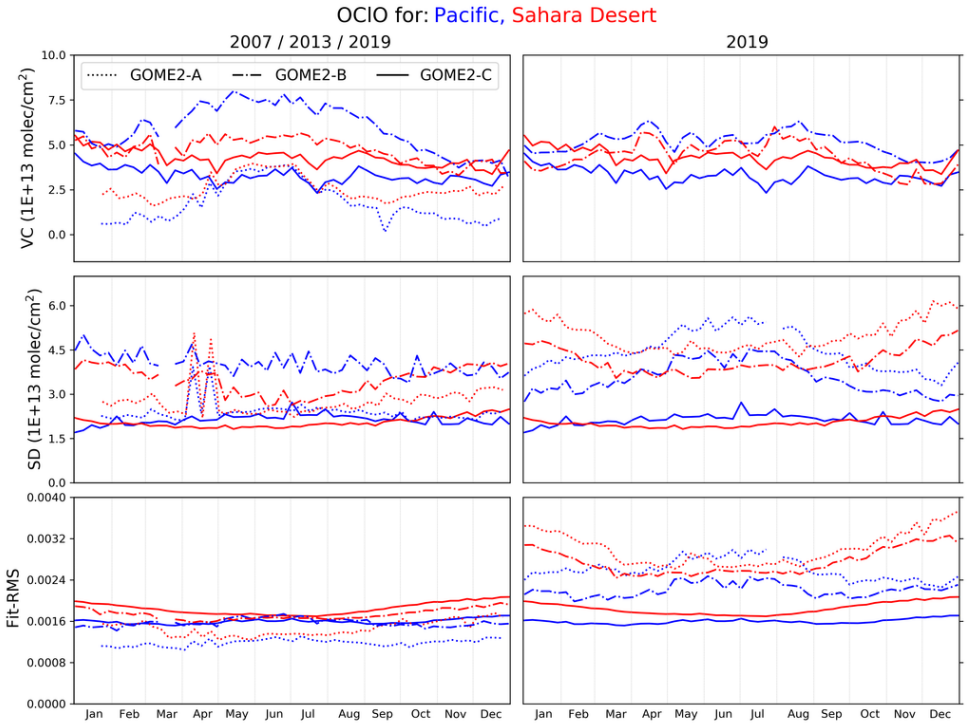

Figure 4-15: Same as Figure 4-1 but for OClO.

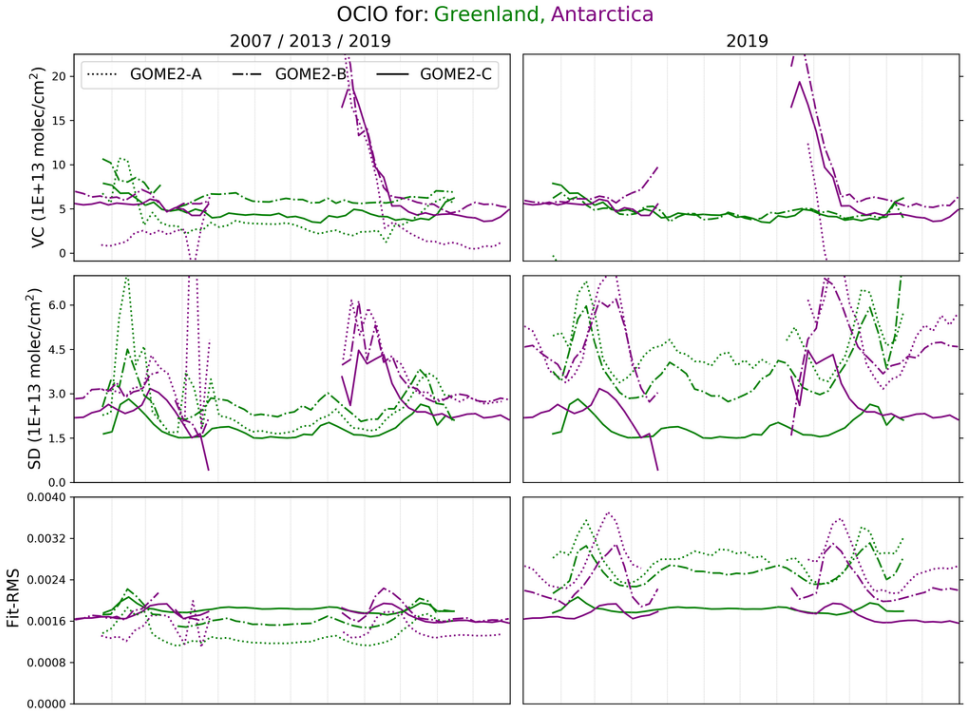

Figure 4-16: Same as Figure 4-2 but for OClO.

As can be seen in Figure 4-15, significant and instrument specific offsets exist in the OClO vertical columns over Sahara and Pacific both in 2019 and in the first year of operation. GOME2-A data, if not corrected, have such a large offset that they are off scale in the figure. Both for scatter and RMS the expected results are found with GOME2-C as the newest instrument having lowest values and GOME2-A the largest. In the first year of operation, the three instruments behaved more similar, with the scatter in GOME2-B data being larger than for the other instruments and GOME2-A having the lowest fitting RMS. The latter result is in line with those for several other trace gases.

Vertical OClO columns over Greenland and Antarctica are more consistent (larger values in polar spring in Antarctica). The scatter and RMS behaves similar in these scenarios as over the Pacific and Greenland, again showing that GOME2-A had the lowest RMS in the first year of operation while GOME2-C has the lowest scatter of the vertical columns.

5 Analysis of issues in Channel 2

GOME2-C channel 2 was found to have spectral features which need to be analysed in order to solve possible inconsistencies of Ch.2 lv2 products between GOME2-A/B and C. These spectral features are discussed together with global lv2 maps on the example of HCHO. Here, G2-C showed spatial and, to some extent, temporal differences compared to G2-A/B. Different key data sets and fit settings were applied and discussed on their impact on HCHO lv2 results and fitting residuals.

Due to the inconsistencies in mid-channel 2 spectral regions and because of the offsets in SO_2 time series, further analyses of high SO_2 abundancies in low intensity regimes for the smallest spectral region of Ch.2 are discussed in a separate section.

Retrievals of OClO from GOME2 instruments are known for their sensitivity to polarisation calibration, and therefore these GOME2-C retrievals are evaluated in more detail and compared to results from GOME2-B. Results are reported in section 5.5.1.

5.1 Issues of HCHO lv2 products for GOME2-C

Figure 5-1 shows the mean residual of IUP Bremen's standard lv2 HCHO product over a region in the Pacific for one day (top sub plot) for both GOME2-C and GOME2-B. Large peaks can be identified at ~358nm for GOME2-C which cannot be seen in GOME2-B fits. These peaks are independent from the day of analysis and from the size of the selected region in the Pacific. The separation of mean residuals by subset (bottom sub plot in Figure 5-1) reveals that some parts of these peaks are subset independent while others remain only for the easterly part of each swath. However, further subset and wavelength depending spectral issues can be seen in the lower part of this figure. East/West oscillations appear strongly at the wavelengths ~328nm and ~338nm with further smaller E/W oscillations in other spectral regions. Wavelength depending oscillations can be identified for larger subsets at 325nm - 328nm and for smaller subsets at 345 -358nm. There are also some subset dependent features in GOME2-B data, but there number is smaller and the amplitude lower.

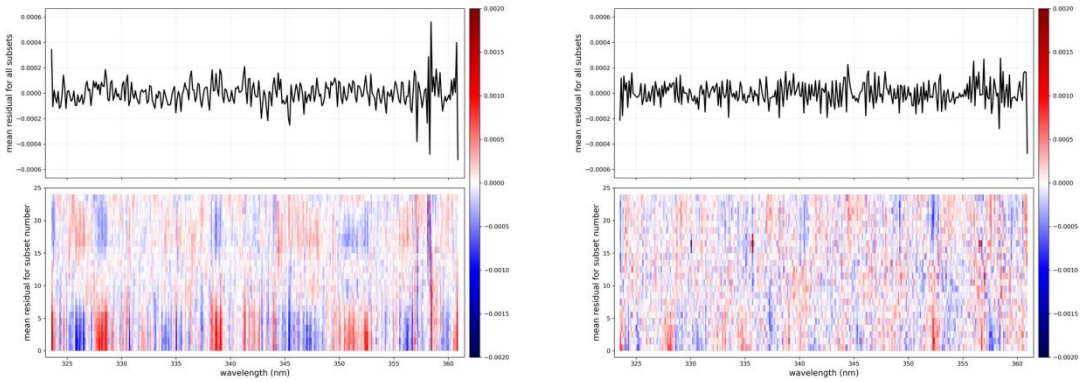

Figure 5-1: Mean residual of the HCHO lv2 product for a region over the Pacific for one day as all- subset mean (top) and subset-depending mean (bottom). The left panel is for GOME2-C, the right one for GOME2-B.

Figure 5-2 shows the global map of HCHO lv2 values for GOME2-C (left) and B (right) for the same day as above. Several spatially distributed issues can be identified for G2-C. Large HCHO values appear in regions where no abundancy is to be expected in winter (e.g. East side of North America, mid of Europe). Areas with negative values appear more frequent at the easternmost part of each orbit leading to a clear E/W (or subset) depending difference. Stripe-like features can be identified for more or less each orbit close to the mid of the orbit going from North to

South. Large values are also present for G2-C at higher latitudes and large SZA. Some of these effects are even more pronounced for fit settings without ETA and ZETA cross sections (Figure 5-3) indicating that the effects might be linked to polarization calibration.

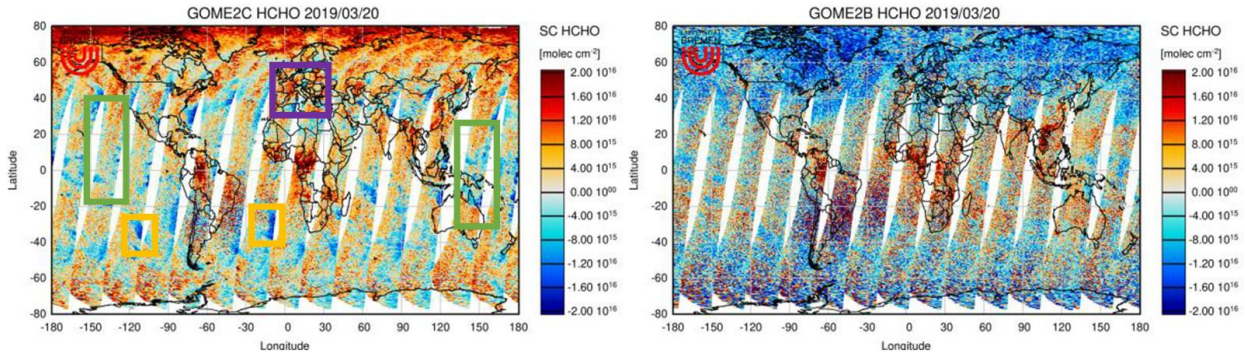

Figure 5-2: Global distribution of HCHO for one day (20.03.2019) for GOME2-C (left) and B (right).

In order to better understand the differences in the global maps and the residual structures, the impact of different fit settings and key data included in the DOAS fits is analysed in the following two subsections.

Note that the key data results (Section 5.2) might differ when applied to other fitting windows while the fitting window changes (Section 5.3) might vary when additional cross sections are applied to the fit. However, these tests give an impression on how to optimize G2-C lv2 results and if it is possible to create consistent lv2 products for Channel 2 of all GOME2 instruments.

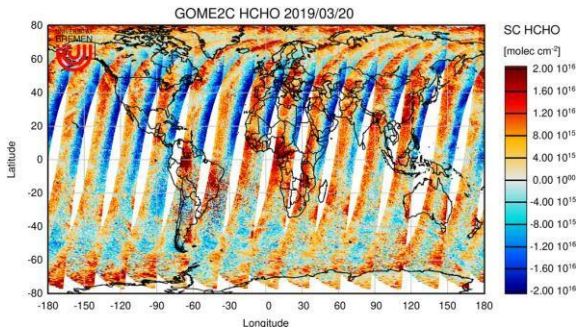

Figure 5-3: Global distribution of HCHO for one day (20.03.2019) for GOME2-C without including ETA and ZETA in the fit.

5.2 Evaluation of different key data fits for GOME2-C HCHO

In order to analyse residual and spatial issues of HCHO lv2 data, different key data were included in the standard DOAS fit (323.5 – 361nm) as additional absorbers. In these tests, ETA and ZETA were not included. Results are evaluated in comparison to the standard fit including ETA and ZETA.

5.2.1 HCHO residual structures when including different key data

To reveal how the different key data sets impact on the residual structures of the HCHO lv2 product (compare Figure 5-1), tests were made with the standard HCHO fitting window, without ETA and ZETA but with each available key data function as single polarization correction cross section in the DOAS fitting routine.

Figure 5-4 shows the fitting residual responses of three different key data function examples representing the range of possible solutions of all performed tests. In general, none of the key data functions can correct for all residual structures. Some improve the east-west oscillation while others decrease the wavelength depending oscillations or the spectral peak feature around 358nm. Furthermore, wavelength and subset depending oscillations sometimes appear (e.g. for POL_GAMMA, 345 – 355nm) which are not as pronounced as in Figure 5-1 for the standard HCHO lv2 product with ETA and ZETA applied within the fit. Further tests (not shown) which involved the standard fit settings including ETA and ZETA and an additional key data set led to an improvement of the residuals but no clear improvement on a global map. As an anticipation of the following subsection it should be clarified that the residual issues as e.g. the East-West oscillations do not directly propagate into spatially distributed issues seen on global maps. On the other hand, a clear independence of residual and spatial structures was also not found.

Note that similarly created key data function tests for GOME2-B did not lead to such strong dependencies (not shown in this document).

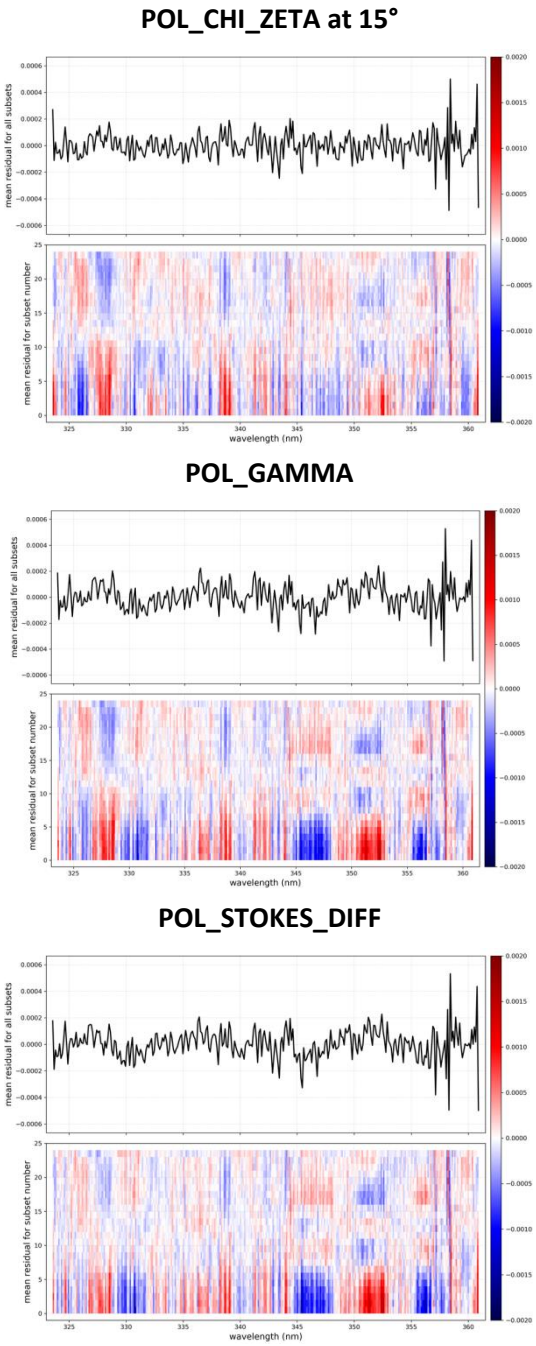

Figure 5-4: Mean residual of a HCHO test fit with one key data function applied in the fit for a region over the Pacific for one day as all-subset mean (top) and subset-depending mean (bottom).

5.2.2 HCHO global maps when including different key data

In the following section, global maps of HCHO SC values for 20.03.2019 are shown. For each map, only one key data function is included in the fit from which ETA and ZETA were removed, similar to the test in the previous section.

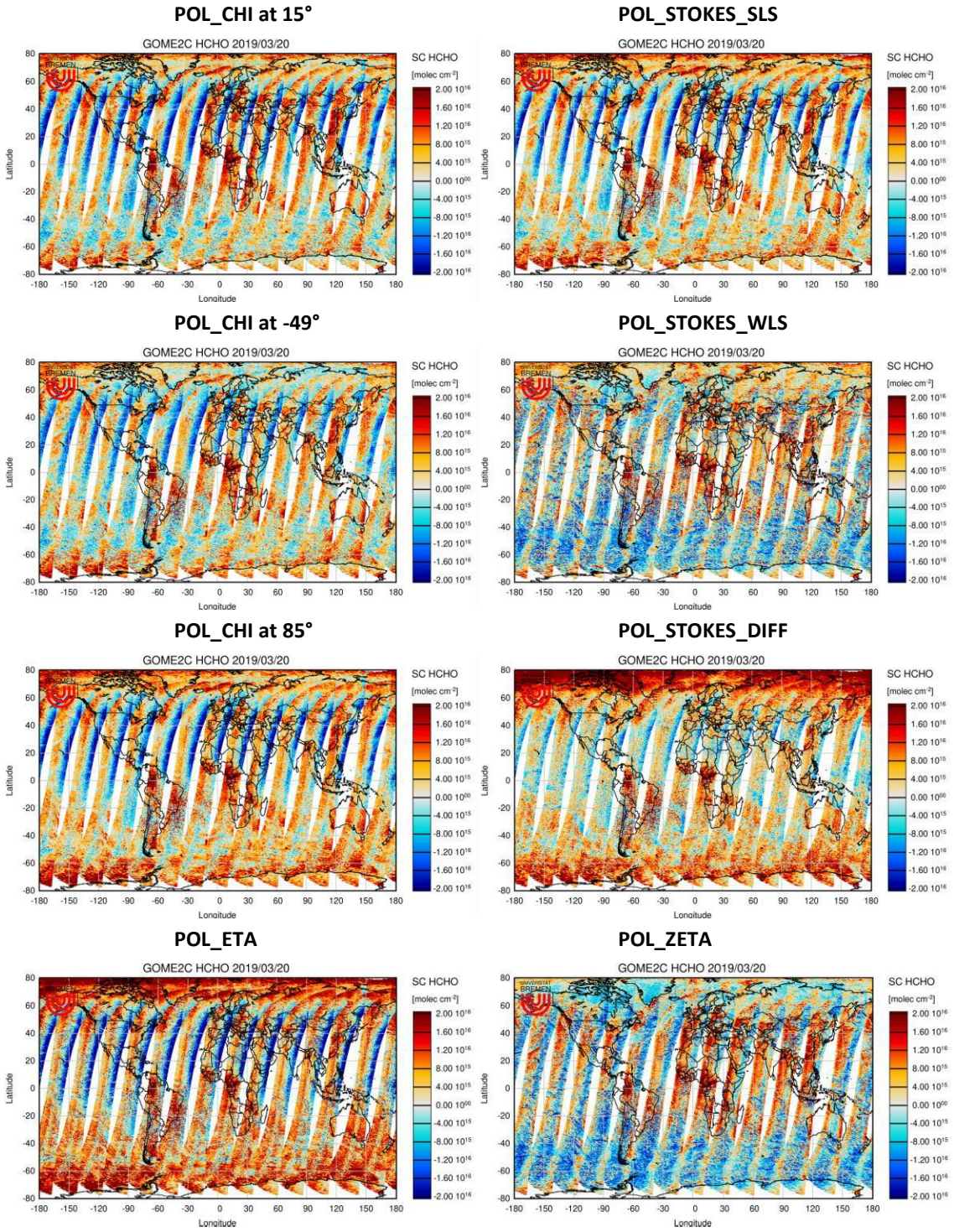

Figure 5-5: Global maps of HCHO SC values created with different key data functions used in the DOAS fit.

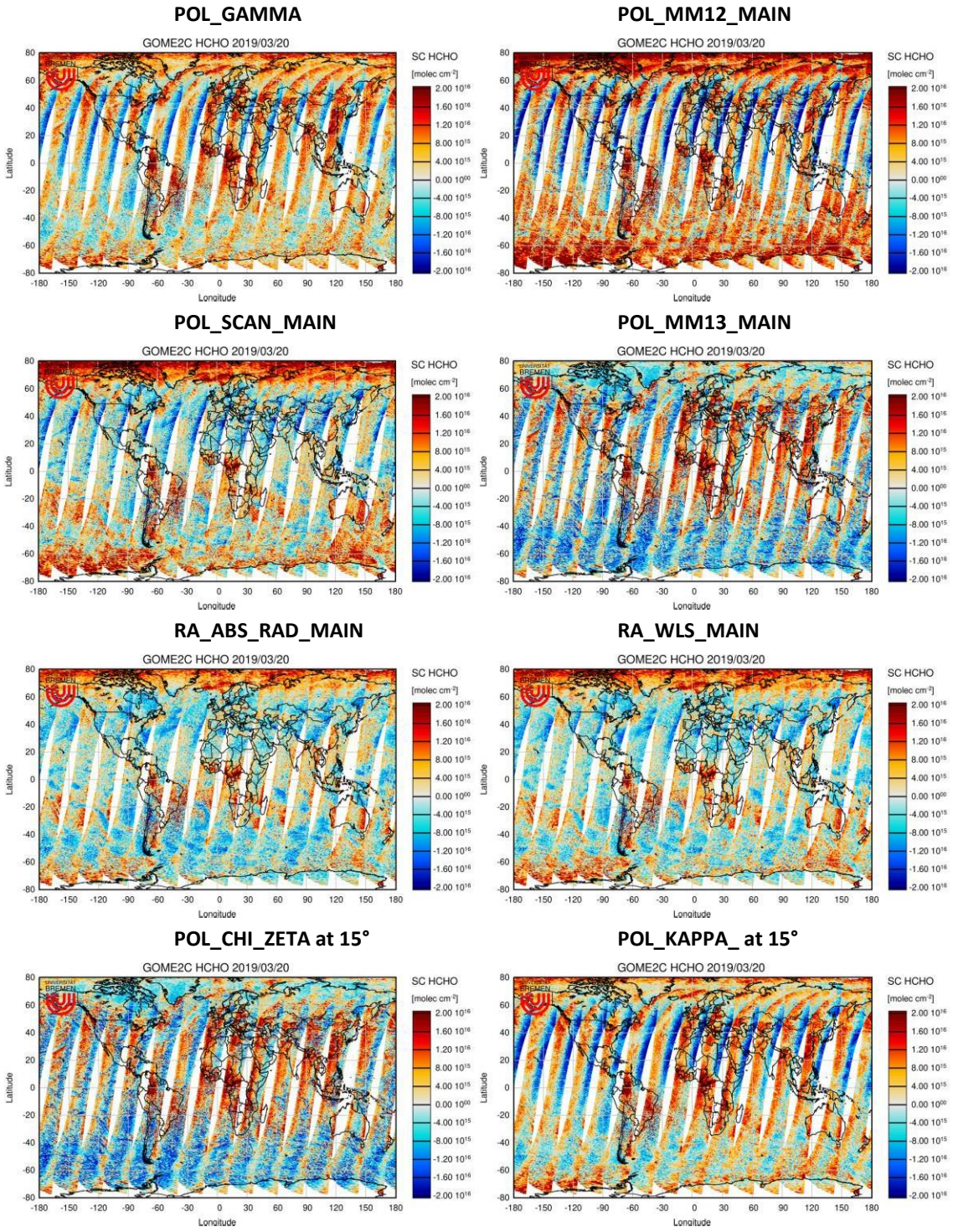

Figure 5-6: Global maps of HCHO SC values created with different key data functions used in the DOAS fit.

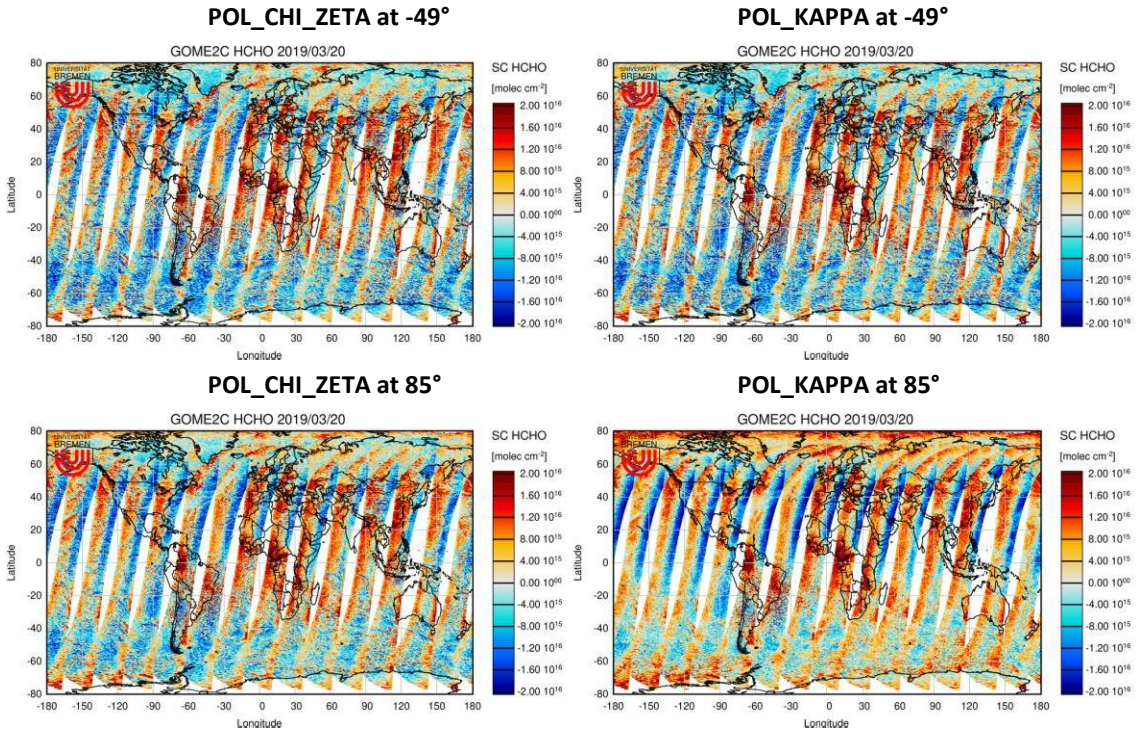

Figure 5-7: Global maps of HCHO SC values created with different key data functions used in the DOAS fit.

It can be seen that the clear East-West dependency is present for nearly all key data functions. For POL_SCAN_MAIN, POL_STOKES_DIFF and POL_STOKES_WLS, the E/W dependency is not as pronounced as for the other cross sections but stripes are visible around the centre of each orbit. The only cross sections which do not show an E/W trend or stripes are RA_ABS_RAD_MAIN and RA_WLS_MAIN. However, with these cross sections, higher values can be identified at the outermost subsets of each orbit. In addition to these points, most key data produce further issues such as strong overestimations in higher (or lower) latitudes or negative values indicating a reduced background level which might also decrease the maximum HCHO values in high abundancy areas. A combination of different key data functions might solve some of these spatial issues but the overall tendency of GOME2-C as being an instrument which needs further corrections during the fit or in an additional post-processing step for channel 2 needs to be highlighted. Note that similarly created key data function tests for GOME2-B did not lead to such strong dependencies (not shown in this document).

5.3 Wavelength dependence of HCHO SC values

In the previous section, different key data functions have been applied to the DOAS fit to solve residual and spatially distributed issues in the HCHO lv2 product by keeping the DOAS fitting window constant at 323.5 – 361nm. Here, the fitting window itself is changed to solve under- or overestimations of the product as seen in the time series of HCHO over the selected regions (see Figure 4-13 and Figure 4-14) by keeping the remaining settings (e.g. cross sections) constant.

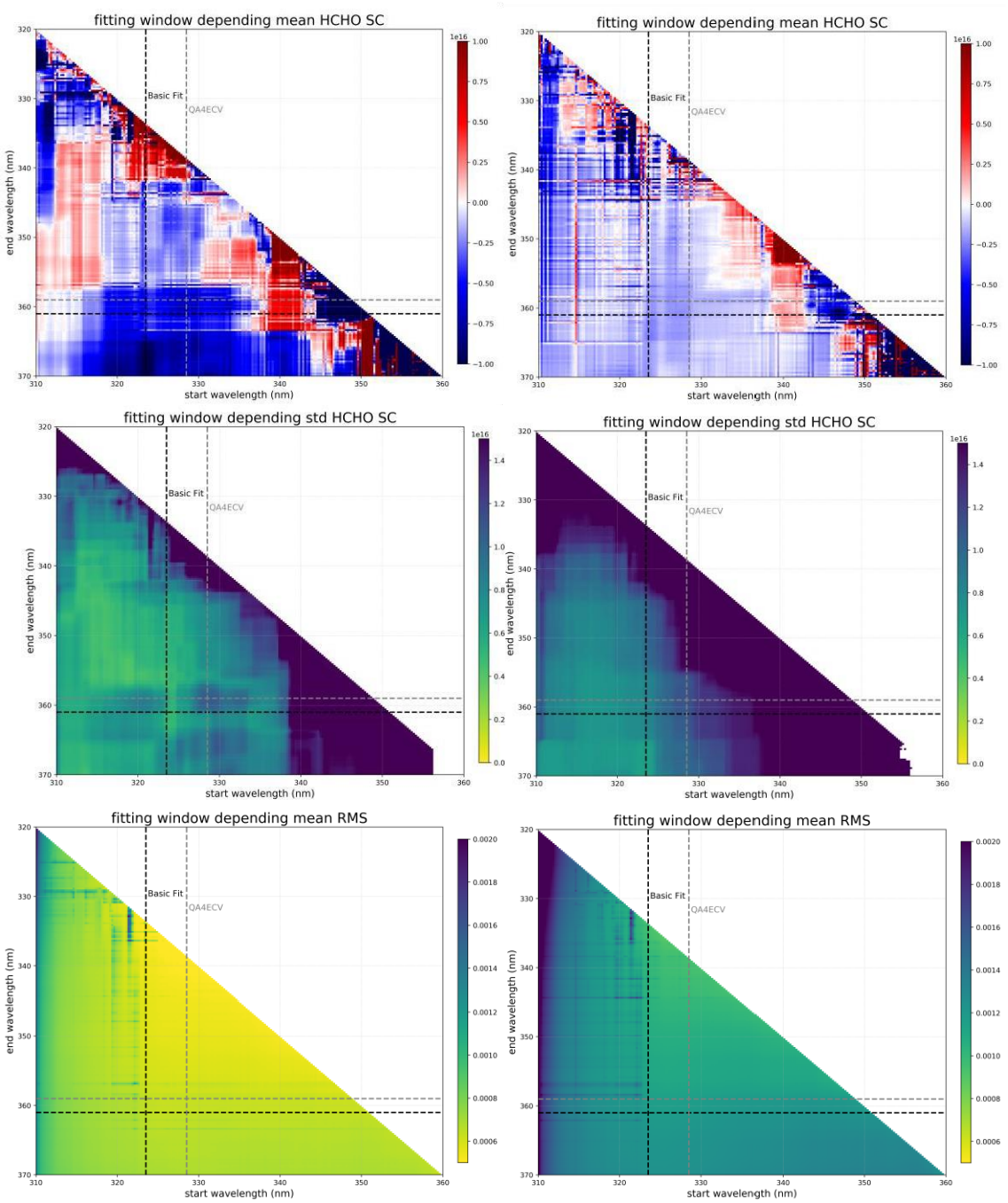

Figure 5-8: DOAS fitting window test over the Pacific. The left column shows GOME2-C results, the right column show G2-B. The top row depicts the mean HCHO SC while the mid row shows the standard deviation of HCHO SC values. The bottom row shows the fitting RMS.

Figure 5-8 shows the dependence of HCHO SC, SD and RMS on the start and end wavelength which have been changed consecutively while repeating the DOAS fitting process and leaving all other settings constant. The left-hand side shows GOME2-C results, while the right-hand side depicts values for GOME2-B. Comparing the mean HCHO SC values over the Pacific (top row), one can identify a much stronger fitting window dependency for G2-C than for G2-B for many combinations of start and end wavelengths. A large area of negative values appears for higher end wavelengths for G2-C. In this area, the end wavelengths for IUP Bremen's standard HCHO lv2 product (black dashed line) and the official QA4ECV product (grey dashed line, see De Smedt et. al 2017) are located. This issue cannot be explained and from this figure it is suggested to apply end wavelengths smaller than 358nm for HCHO when no further (subset depending) correction cross sections are used within the DOAS fit.

The variation of HCHO SC values (mid row) shows a rather smooth transition from large to small start wavelengths while it is more variable for G2-C. However, the standard deviation values themselves are usually larger for G2-B indicating the higher noise level due to the stronger (in orbit) degradation of the instrument. This does not explain the larger wavelength depending variability of the standard deviation subplot for G2-C and indicates again spectral issues.

The fitting window RMS shows a similar behaviour for both instruments with large values for the smallest start wavelengths due to the growing impact of O_3 absorption. Horizontal and vertical lines with high RMS often correspond to residual structures but sometimes appear around strong Fraunhofer absorption lines (e.g. 323nm).

Note that this particular test on the wavelength dependence can only be used to discuss possible improvements for one set of parameters (cross sections, polynomial coefficients, …) only. Results might change when additional correction cross sections are applied.

5.3.1 Comparison of two different HCHO fitting windows

Results above imply that smaller fitting windows might lead to a better agreement between the different GOME2 instruments as the HCHO SC changes for G2-C vary strongly with wavelength but do show reasonable mean values around zero for end wavelengths smaller than 358nm. GOME2-B, on the other hand, does not show strong wavelength dependencies which might support that a change of the fitting window leads to a better agreement between the instruments.

In Figure 5-9, time series and global maps are shown for the standard fitting window (323.5 – 361nm) and a smaller one (330.25 – 356.25nm). Further parameters such as polynomial degree or cross sections are kept constant.

The first figure shows on the left hand side the results from the time series discussion in the previous chapter while the right hand side introduces results for the smaller fitting window. In addition, the last sub-plot depicts a high abundancy region in Africa showing the performance of both fits over a region with a larger HCHO load.

The comparison of time series in the first row (Pacific and Sahara) shows a clear improvement when using the smaller fitting range. The HCHO SC values are closer to zero and a better agreement can be seen for all curves with the exception of GOME2-A for the Pacific. The SD curves are slightly higher indicating that there is a larger scatter of values in the analysed area and a slightly stronger seasonal dependency. The Fit-RMS is on a similar level. The discrepancy between G2-B/C and A is even more pronounced when looking at the mid sub-plots of this Figure which show the results for Antarctica and Greenland. Here, GOME2-A shows an even larger

offset compared to the other instruments. GOME2-B and C, on the other hand, agree slightly better with the small fitting window. Since the curves are closer together during summer in the NH, this might indicate instability due to less light and for large SZA.

323.5 – 361nm 330.25 – 356.25nm

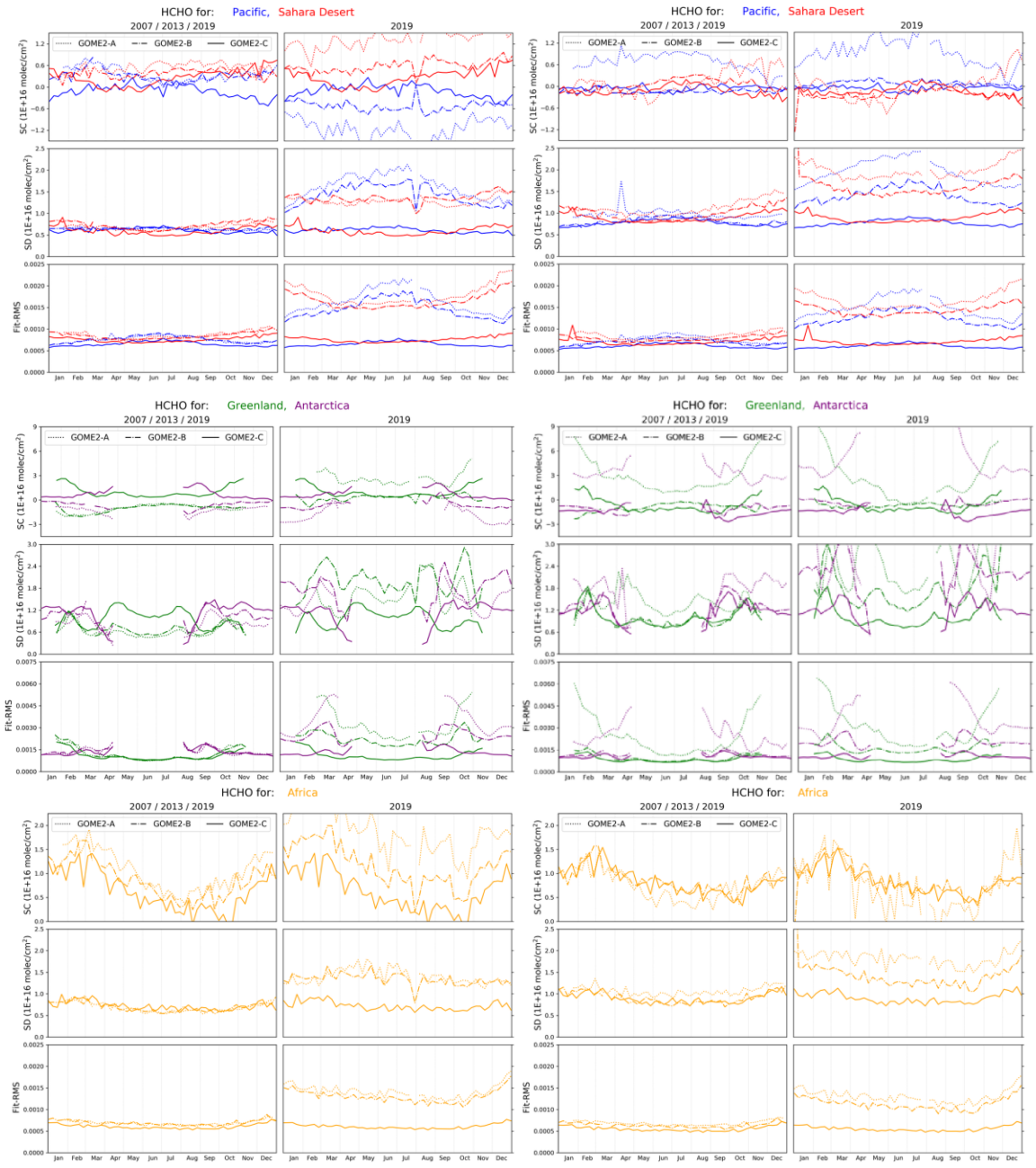

Figure 5-9: Comparison of the standard (left) and a small (right) fitting window for HCHO results. The two top sub-plots show low abundancy regimes (the Pacific and the Sahara desert, top; Greenland and Antarctica, middle). The bottom row shows results for a high abundancy regime in Africa.

SD and Fit-RMS curves show deterioration with higher values compared to the larger fitting window but also between each instrument. The last sub-plot of Figure 5-9 depicts results for a high abundancy area in Africa. The VC curves show a much better agreement for the small fitting window indicating a general better performance when enough light and HCHO is present (in

contrast to results for Greenland and Antarctica). The variability of SC values (SD curve) and the Fit-RMS are slightly larger for the small window but do not show a drastic deterioration which cannot be explained by less spectral sampling points within the DOAS fit.

323.5 – 361nm 330.25 – 356.25nm

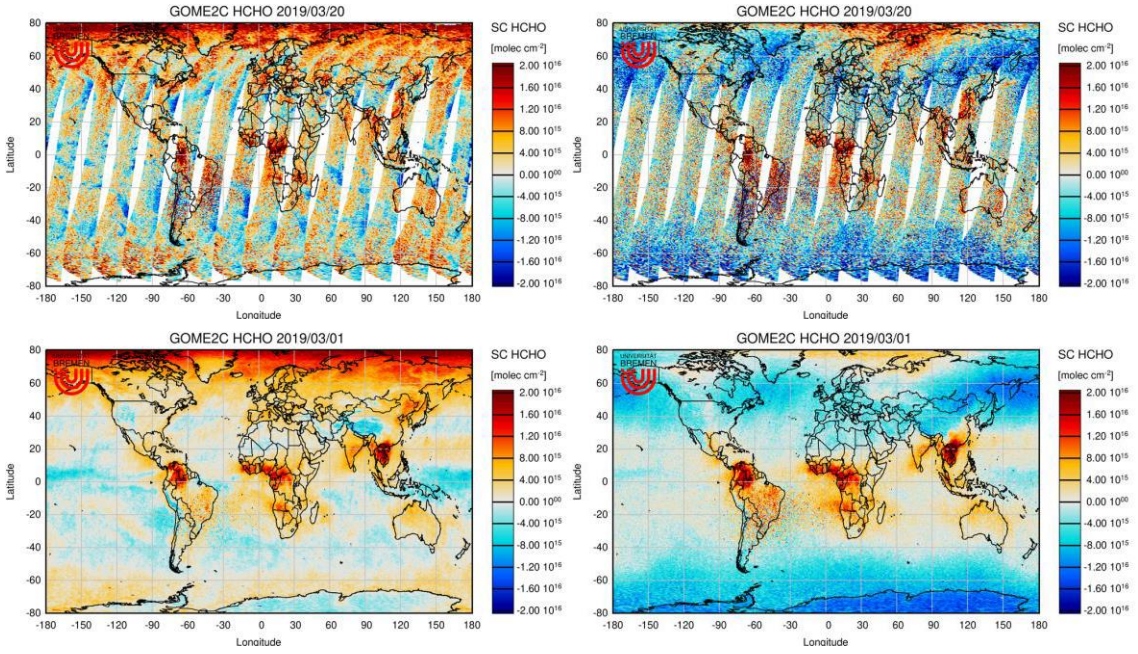

Figure 5-10: Global maps of HCHO SC for the standard (left) and the small (right) DOAS fitting window and for one day (top) and a monthly mean (bottom).

Figure 5-10 shows a comparison of global maps for both fitting windows on a daily basis (top) and as a monthly mean map (bottom). Many of the spatial issues are gone with the smaller window. No stripes can be seen on the right-hand side and the large E/W dependence with high values in the West and lower SC in the East is not as pronounced as in the large fitting window. There is still a large area with unrealistic high HCHO values above northern Russia but the northern hemispheric overestimation seems to be solved. However, the smaller fitting window suffers from many negative values at high latitudes for both hemispheres. For mid latitudes and around the equator, the map looks reasonable and similar to the version of GOME2-B (compare Figure 5-2). The monthly mean maps have significantly improved from large to small fitting window as the strong positive and negative oscillations in the Pacific are gone and values are closer to zero. The high HCHO abundancy areas in Africa look similar but a westward outflow around -20°S is more pronounced with the smaller window. Again, the high overestimations for the NH are solved but there appears to be too many negative values for the southern Hemisphere above Antarctica.

In general, the test in this section has shown that comparable time series and data products for GOME2-A, B and C can be created but the spectral and spatial issues in G2-C show the need for new settings for the creation of lv2 products in the UV. Further possible solutions are additional (correction) cross sections or post processing as e.g. a destriping correction on lv2 basis.

5.4 Analysis of lv2 data in low intensity regimes for high trace gas abundances on the example of SO_2

The time series in the previous chapter have shown that relative offsets of up to 100% can be identified for SO_2 VC between GOME2-A and B/C. Furthermore, the VC time series of GOME2-A is noisier with larger SD and RMS values, compared to the other instruments. In this section, these findings are assessed on their impact on high SO_2 abundancies on the example of the eruption of the Raikoke volcano in Kamchatka starting in June 2019.

The left-hand side of Figure 5-11 shows SO_2 for the second day of the eruption for all three instruments. In this figure, black areas depict values smaller than zero. It can be seen that the general spatial agreement between the instruments is good but the GOME2-A results are noisier, compared to GOME2-B/C. The right-hand side of the figure shows the corresponding frequency distribution of SO_2 values in the analysed area. The width of the depicted distribution is the widest for GOME2-A indicating the larger spread of values in the maps. However, the maximum SO_2 values inside this area are in good agreement, with $5.3x10^{18}$ molec/cm^2, $5.1x10^{18}$ molec/cm^2 and $5.4x10^{18}$ molec/cm^2 for G2-A, B and C respectively. The spatial distribution as well as the maximum values indicate that there is a good general agreement for high abundancy regimes. The only drawback is the larger spread of values the older the instrument which can be explained by less throughput in the UV spectral region due to degradation.

Figure 5-12 depicts the time series of SO_2 over the Raikoke region for all three instruments for 2019. Throughout the year, the instruments show a clear offset (left, light lines) with G2-A larger than G2-B and G2-C for the mean of the region except for the time of eruption, which supports the findings in the previous chapter. The maximum value (left, dark lines) of this region shows no offset, neither in the low abundancy time period nor during the time of eruption. The strong decrease of all curves (but especially G2-B) in the early and late months is due to the large SZA for this high altitude region which further decreases light and therefore fitting quality.

The right-hand side of this figure shows the unfiltered time series and confirms that the instruments agree well for high abundancy regimes. In contrast to the filtered maximum curves, an offset can be identified for the dark lines as well. The mean values for GOME2-B and C agree better but an offset for G2-A is also visible. The increase in background level when no filtering is applied for the maximum curves and for all instruments shows the larger level of noise and larger uncertainties (Fit-RMS) in higher latitudes when less light is available.

As a conclusion, less throughput for all instruments and the smaller swath width for G2-A in particular are the most probable explanations for the larger spread of values and the offsets. However, the change in the background level could also be explained by not fully optimized DOAS fit settings. In contrast to the HCHO findings above, the SO_2 lv2 products are comparable and the performance of the instruments for high abundancy areas is good.

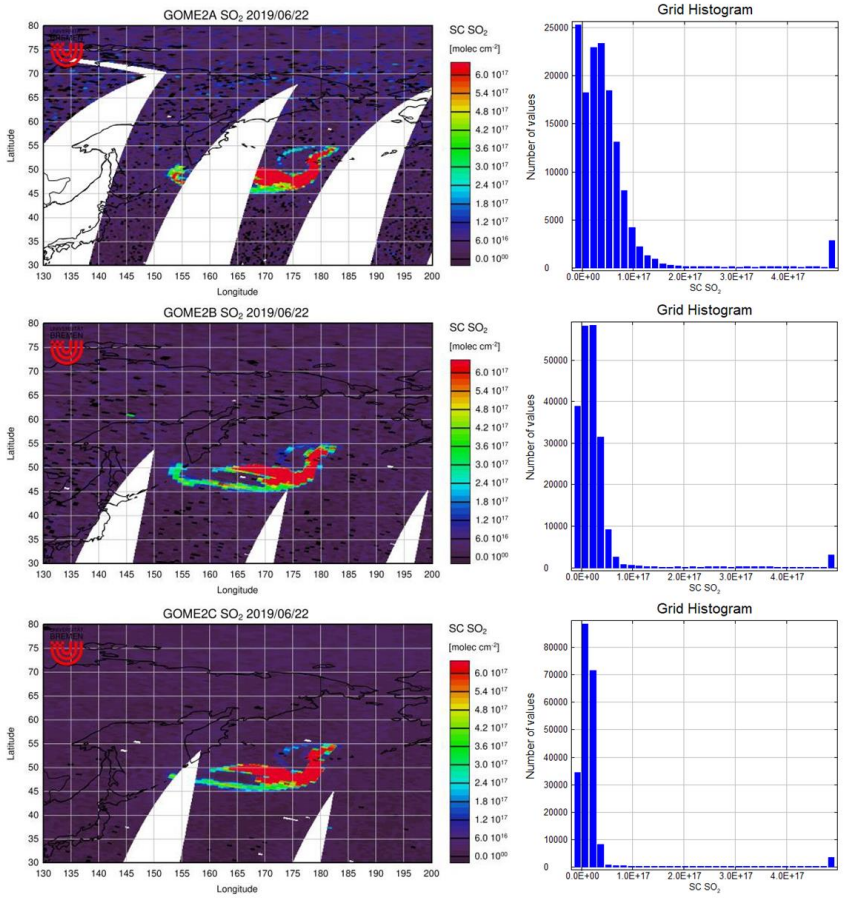

Figure 5-11: SO_2 maps of the Raikoke eruption (left) and histograms of SO_2 SC values in this area (right) for GOME2-A/B/C (top to bottom).

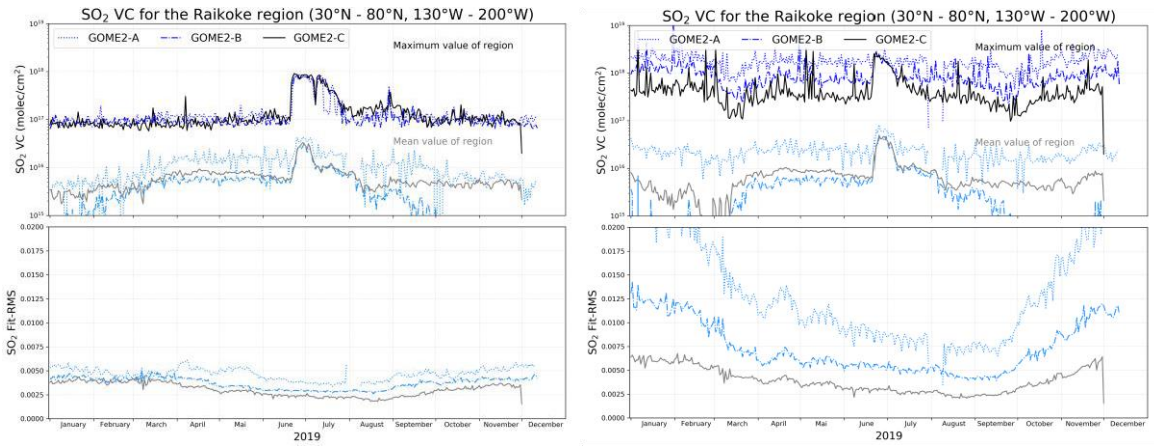

Figure 5-12: SO_2 time series of GOME2 on MetOp-A, B and C of VC and Fit-RMS (top, bottom) for 2019. Different line styles show G2-A (pointed line), G2-B (dashed line) and G2-C (solid line), respectively. The time series are depicted for the Raikoke region (Kamchatka) with RMS > 0.01 filtered (left) and without filter (right). Dark colours show the maximum value of the region; light colours show the mean value.

5.5 Evaluation of OClO retrievals

5.5.1 Impact of polarisation calibration

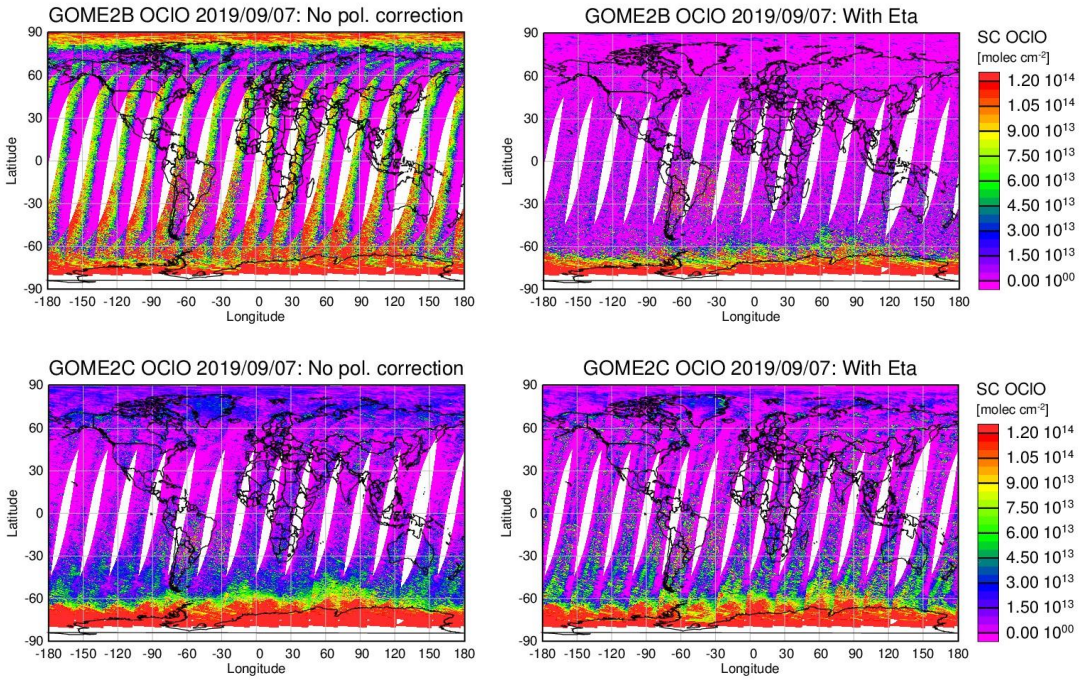

Figure 5-13: OClO SC for September 7, 2019. The top row shows GOME2-B results, the bottom row GOME2-C data. The left column is for results without including eta and the right column for data with eta included in the fit.

When applying the fit settings used for GOME2-A and GOME2-B to GOME2-C data, good general agreement is found in the geographical distribution and the absolute values with some specific differences as discussed in the next sections. For GOME2-A and GOME2-B, including the Eta key data function in the fit was strictly necessary to create reasonable results, and in the case of GOME2-A data, additional empirical cross-sections were included to reduce artefacts. In the case of GOME2-C, this is not necessary, and indeed, OClO SC results show less artefacts in the retrieval without polarisation correction (see Figure 5-13). This is indication for a good polarisation calibration, at least in the spectral window used for OClO retrievals.

5.5.2 Impact of subset

Although the GOME2-C OClO retrievals look good at first glance, averaging of all data from September 2019 by subset reveals that there is a systematic dependence on subset with an amplitude of up to $1x10^{13}$ molec cm^{-2}. The origin of this dependence is unclear but must be linked to calibration issues. In the retrievals, an Earthshine background spectrum was used which was created by averaging radiance measurements over the equatorial Pacific taken on the same day. This background spectrum was not separated by subset as can also be done but creates ambiguities at larger SZA where pixel size increases and subset counting is no longer consistent. It is expected that a subset dependent background spectrum would remove at least the additive part of this subset dependence.

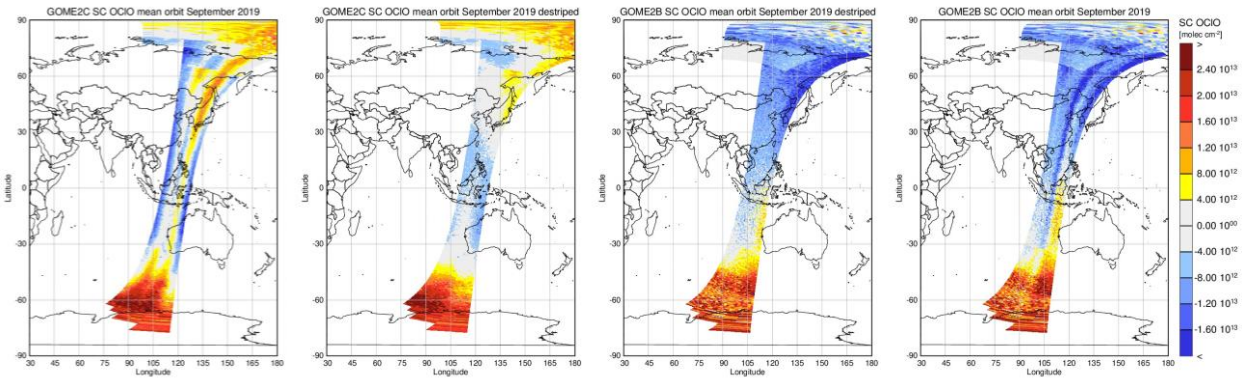

Figure 5-14: GOME2-C mean orbit dependence of OClO for September 2019. From left to right: GOME2-C without destriping, GOME2-C with destriping, GOME2-B with destriping, GOME2-B without destriping. For GOME2-C, the retrieval with polynomial of degree 5 is shown. Both retrievals use a Pacifc background spectrum.

In order to investigate if the effect is indeed additive, a post destriping was applied to the slant columns by removing the subset dependent mean slant columns as determined over the equatorial Pacific. The results are shown in Figure 5-14 for both GOME2-C and GOME2-B. As can be seen, both instruments show pronounced dependence on subset in the mean orbit without destriping, and quite smooth behaviour with subset after destriping. This indicates that the largest part of the observed effect is indeed additive. It is also apparent that GOME2-C data are not only less noisy (as expected for a new instrument) but also show less variation along the orbit towards Northern latitudes, where no OClO is expected (grey values). There is a hint of a problem in the upper right part of the GOME2-C orbit, a region known for difficulties with the polarisation correction, but the effect is small and overall, GOME2-C appears to have good calibration in the OClO fitting range.

5.5.3 Comparison with GOME2-A and GOME2-B data

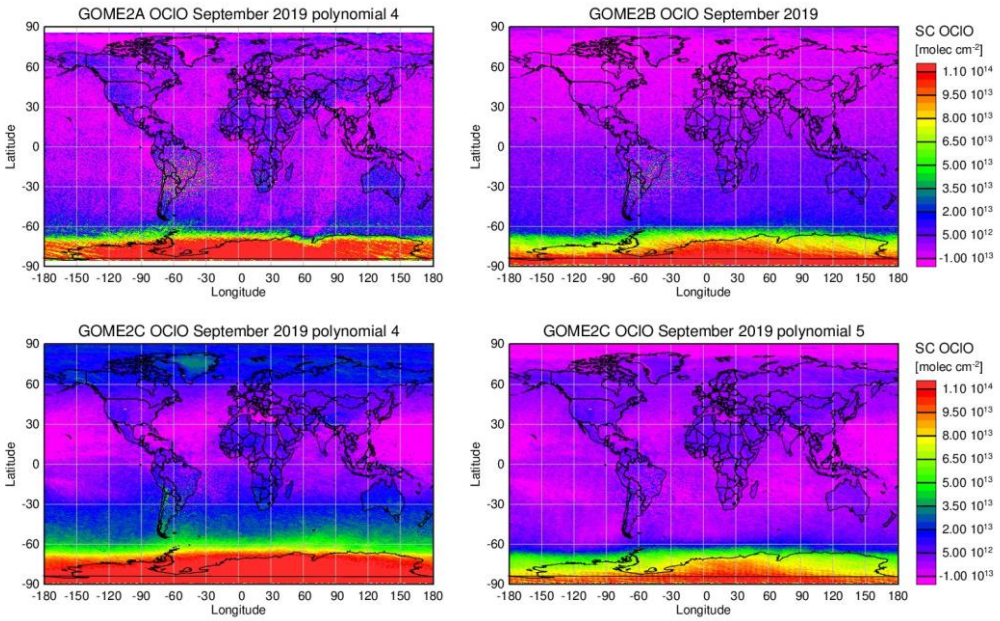

Figure 5-15: OClO SC for September 2019 from GOME2-A (top left), GOME2-B (top right) and GOME2-C (bottom) using two different degrees of the polynomial in the DOAS fit.

As discussed in the previous sections, GOME2-C OClO retrievals appear to perform well. It is therefore interesting to compare the results to those of GOME2-A and GOME2-B (see also section 4.1.8). As shown in Figure 5-15, there is overall good but not perfect agreement between GOME2-A and GOME2-B OCLO columns for September 2019. If the same retrieval settings are used for GOME2-C, values over the activated Polar vortex are larger than from the other two instruments, and in addition there is a marked sea-land contrast and indication for a SZA dependence. This can be much improved by increasing the degree of the polynomial in the DOAS fit from 4 to 5. With these settings, the sea-land contrast is much reduced, the SZA dependence nearly gone but OClO values over Antarctica are now also lower than from the other two instruments. More work will be needed to create fully consistent OClO retrievals from all three GOME2 instruments.

5.5.4 Impact of inhomogeneous scenes

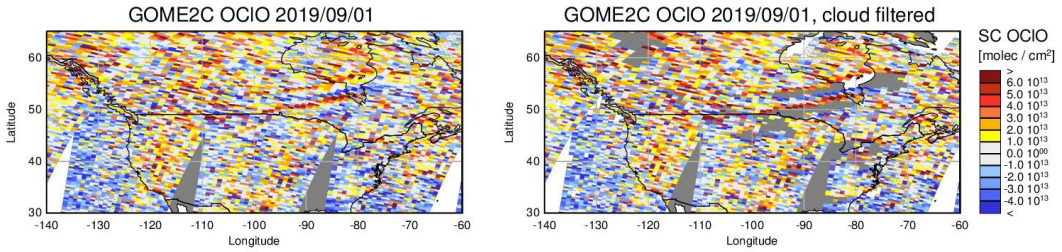

Figure 5-16: Zoom into GOME2-C OClO slant columns for September 1st, 2019. Left: all data, right: after removal of the brightest scenes (clouds).

It is well known that inhomogeneous scenes with large intensity gradients can lead to changes in the instrument function, leading to increased residuals in the retrievals, mainly for imaging spectrometers but also for scanning instruments such as GOME2 (Noël et al., 2012). In first approximation, the effect is a spectral shift which can be compensated in the fit by the shift and stretch terms included.

In many GOME2-C OClO retrievals, sharp lines of very high and also of negative OClO are apparent where no OClO is expected. Closer inspection reveals that these lines coincide with the edges of bright clouds which are regions of large intensity changes (compare Figure 5-16). The same patterns are also visible in maps of the spectral shift fitted, both in GOME2-C and GOME2-B data of the same scene (Figure 5-17). This confirms that the OClO artefacts are indeed linked to inhomogeneous scenes.

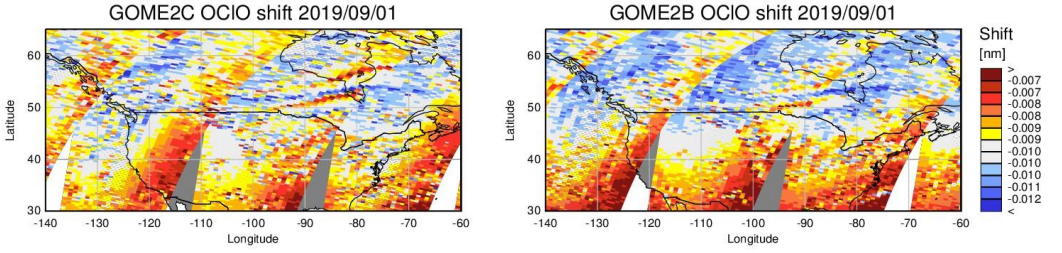

Figure 5-17: Spectral shift retrieved in the OClO fit for September 1st, 2019. Left: GOME2-C, right: GOME2-B

In the case of clouds, such high and low OClO values usually average out when averaging over several days as is often the case for OClO. However, in some regions with persistent clouds or at the edge of snow or ice (Greenland, snowy mountains, the sea ice edge) artefacts are visible

even in monthly averages of GOME2-C OClO retrievals. Such an artefact can for example be seen at the west side of Greenland in the GOME2-C image in the bottom right panel of Figure 5-15.

In the case of GOME2-B, a similar effect is found in the spectral shift, but there is no indication of artefacts in the OClO columns themselves, neither in individual observations, nor in monthly averages. Why this difference exists between the two instruments is not clear at this point.

5.5.5 Impact of background spectrum

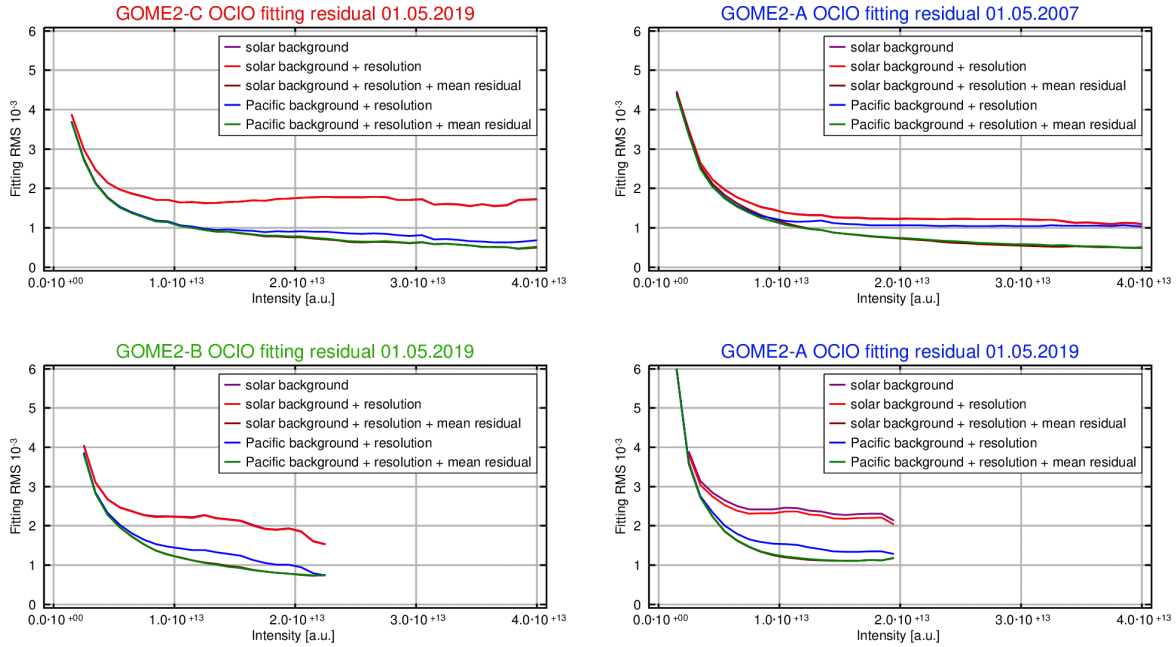

Figure 5-18: Comparison of OClO fitting residuals for 5 different fit settings (see text for details).Results are for GOME2-C (top left), GOME2-B (bottom left), GOME2-A (bottom right) and GOME2-A in the first year of operation (top right). Only measurements with short integration times are included

The fitting RMS can be used as an indicator of the quality of the calibration of the lv1 data. This assumes that the DOAS fit itself is perfect which of course is not the case. However, in the case of a good fit, the RMS should be dominated by photon shot noise and thus reduce as intensity increases. This is not the case in the presence of systematic residuals which can result from calibration issues. In Figure 5-18, the results of such an analysis are shown for an orbit over the Pacific for all three GOME2 instruments on May 1st, 2019. In addition, a similar orbit is shown for GOME2-A in 2007, the first year of operation. This can directly be compared to the GOME2-C results for 2019. Several different retrieval settings are included in each panel using either the solar irradiance as background (red, lilac and brown) or a Pacific earthshine (blue and green). With the exception of the first fit, all retrievals include a resolution correction. Two cases (brown and green) also include the mean residual retrieved from a previous fit of the same orbit.

As can be seen, the resolution correction has basically no impact on the results with the exception of the current GOME2-A data where a small improvement can be achieved. This is in contrast to results for NO_2 (see section 6.1.3).

There is a clear difference in the RMS between fits using the irradiance or a pacific earthshine as background, the latter resulting in much better fits. The only exception are GOME2-A data in 2007, where the quality of the fits with an irradiance are clearly better than for other

instruments but the earthshine retrieval is not as good as for GOME2-C. This indicates that a) the solar irradiance of GOME2-A had a better calibration in 2007 than those of GOME2-B and GOME2-C and b) that there was an issue with the earthshine spectra in GOME2-A which is not present for the other instruments.

When also including the mean residual, all instruments show very similar behaviour of RMS as a function of intensity, indicating that this leaves mainly the shot noise based RMS. It is interesting to note that for GOME2-C, the difference between a fit with Pacific background with and without including the mean residual is quite small, indicating that there are little systematic parts in the residuals of these fits, which could be interpreted as good calibration.

These results should be compared to those found for a similar analysis for NO_2 as discussed in section 6.1.3.

5.5.6 Long-term stability

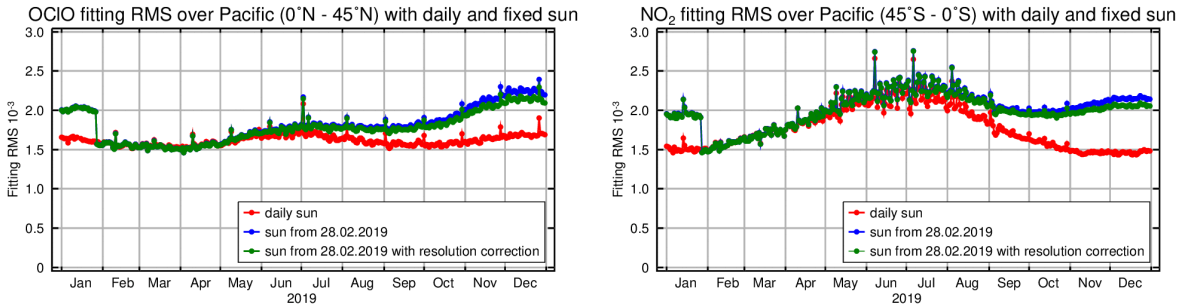

Figure 5-19: GOME2-C OClO fitting residuals over the Pacific for the Northern Hemisphere (0° - 45°N, left) and the Southern Hemisphere (0° - 45°S, right). Results from three different retrievals are shown: The standard evaluation with a daily solar irradiance (red), an analysis with a fixed solar irradiance from February 28, 2019 (blue) and a fit with the same fixed irradiance but including a resolution correction function.

As GOME2-C ages, it will not only suffer from reduced throughput but will also develop wavelength dependent changes of sensitivity and slit-function. Such changes can be monitored by investigating the fitting residual as a function of time in a DOAS analysis where not the daily solar irradiance is used (which cancels most degradation effects in the radiances) but a fixed irradiance spectrum. The results are shown in Figure 5-19 for one year of OClO fits, separated by hemisphere. While until end of May the RMS of a fit with daily and with fixed irradiance are comparable, curves separate and while the fit with daily irradiance is as good in December 2019 as it was in January 2019, the retrieval with fixed irradiance shows an RMS increase by about 30%. Including a resolution correction for a possible change in slit function does not help much, indicating that changes in slit-function are not the driver of the observed increase in fitting RMS. This is in contrast to the situation for NO_2 as discussed in section 6.1.4.

5.5.7 Excursion: GOME2-B calibration issues

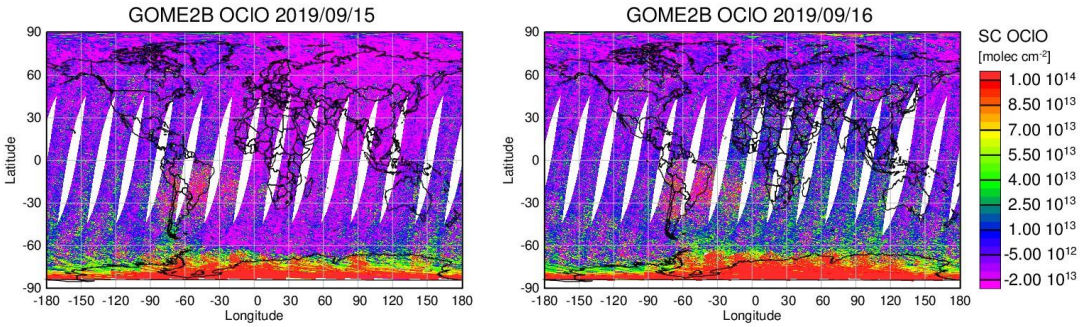

Figure 5-20: Two days of GOME2-B OClO showing an offset for some orbits. A Pacific Earthshine spectrum has been used as background.

When comparing GOME2-C and GOME2-B OClO columns with Earthshine background spectra, an unexpected artefact appeared in GOME2-B spectra which is illustrated in Figure 5-20. On some days, there is a change in OClO column around 0° longitude, the location where the daily solar irradiance saved in the orbital files changes. This effect is found on many but not on all days. It is however not present when using the daily solar irradiance as background spectrum.

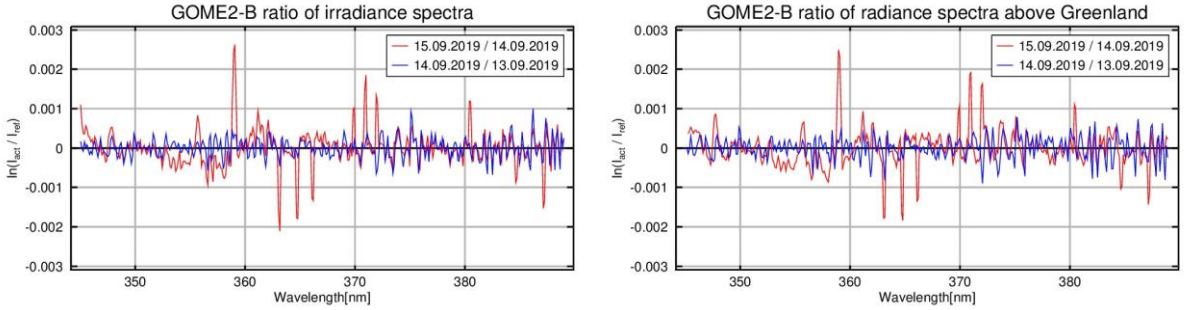

Figure 5-21: Optical depth ($ln(I_1 / I_2)$) of GOME2-B solar irradiance measurements (left) and radiance measurements over Greenland (right) for three days in September 2019. For clarity, a polynomial has been subtracted from the ratios.

Inspection of the solar irradiance files showed that on days exhibiting this behaviour, the solar irradiance changes relative to the previous day with a clear spectral pattern. This is illustrated in the left panel of Figure 5-21. Surprisingly, the same change is found for radiances taken over Greenland when comparing data from the same days (see right panel of Figure 5-21). This explains why the problem cancels when using daily solar irradiance measurements as background.

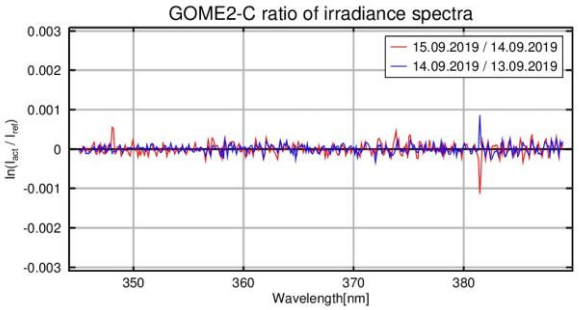

Figure 5-22: As Figure 5-21 left but for GOME2-C

This is a relatively large effect and clear indication for a calibration problem affecting both radiance and irradiance spectra of GOME2-B. For comparison, the solar irradiance ratios for GOME2-C have also been computed and are shown in Figure 5-22. Clearly, no such spectral structures are apparent, the ratio showing much lower values overall. This is also true for all other valid GOME2-C irradiance measurements. This calibration problem thus appears to be present only in GOME2-B data.

It is interesting to evaluate the long-term evolution of this phenomenon. This is done in Figure 5-23, where for all of the year 2019, the logarithm of consecutive irradiance measurements is shown. For each day, a polynomial in wavelength has been subtracted to remove broadband changes. As can be seen, the high frequency changes appear on many but not on all days, and in February and March, there is an extended period without changes. This illustrates the somewhat random pattern of appearance of this effect.

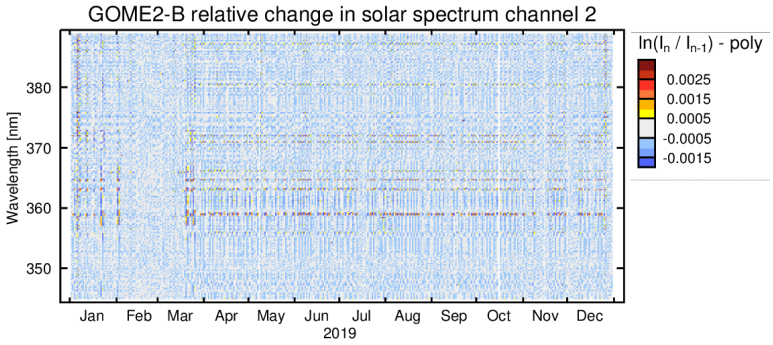

Figure 5-23: Relative change of the GOME2-B solar irradiance in the fitting window of OClO during 2019. Shown is the logarithm of the ratio of consecutive solar irradiances $ln(I_n / I_{n-1})$. To remove broadband effects, a polynomial in wavelength was subtracted for each day.

In Figure 5-24, another way of looking at the change in solar irradiance is plotted, where each irradiance measurement is divided by the same single solar irradiance measurement taken early in the mission. Again, a polynomial in wavelength was subtracted for each day to remove the effects of broadband degradation. As can be seen, high frequency variations in the solar irradiance appear early on in the mission but increase in amplitude over time. There also is a seasonal pattern in the amplitude. The problem did thus not appear suddenly at a given point in time linked for example to a change in calibration but rather slowly developed over time, indicating that it is the result of long-term degradation of GOME2-B.

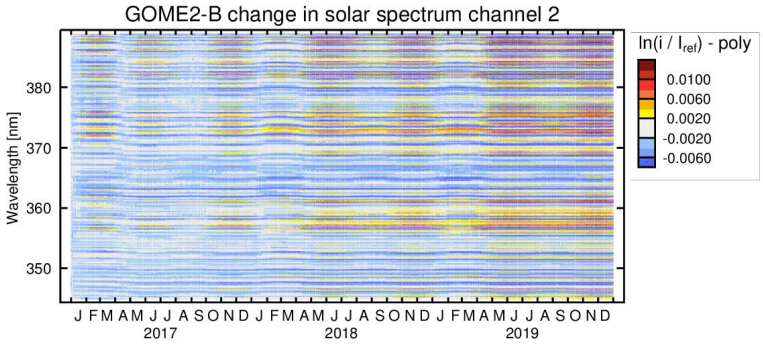

Figure 5-24: Long-term evolution of GOME2-B solar irradiance in the OClO fitting range. Plotted is the $ln(I / I_{ref})$ where I_{ref} is one of the first GOME2-B solar spectra. To remove the effect of broadband degradation, a polynomial in wavelength was subtracted for each day.

6 Analysis of issues in channel 3

6.1 Evaluation of NO_2 retrievals

6.1.1 Overall quality of retrievals

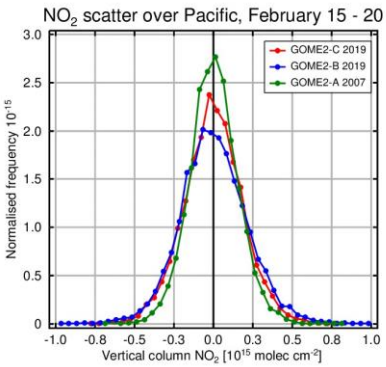

Figure 6-1: Scatter of NO_2 VC values over the equatorial Pacific compared between GOME2-C (red) and GOME2-B (blue) in February 2019. Also shown are the GOME2-A results (green) but for 2007, the first year of GOME2-A operation. Mean values are forced to 0 in order to allow easy comparison.

As a first test of NO_2 retrievals on the GOME2-C data, standard NO_2 fits were performed using three typically used fitting windows (425 – 450 nm, 425 – 497 nm, 430 – 465 nm). All of these retrievals resulted in good quality fits, albeit with slightly different results. While the fitting RMS is a useful indicator of the goodness of the fit, the actual scatter of the retrieved vertical columns over a clean region, where no systematic variation in NO_2 columns is expected, is a more interesting quantity for the users (see also results presented in 4.1.1). As an example, the histogram of NO_2 deviations from the mean VC over the Pacific is shown for a few days in February 2019. As expected, GOME2-C shows less noise than GOME2-B which already suffers from some degradation related decreases in throughput. However, it is also clear that GOME2-A had even less scatter in its first year of operation, indicating that GOME2-C NO_2 data in 2019 are not as good as GOME2-A data were in 2007.

6.1.2 Impact of fitting window

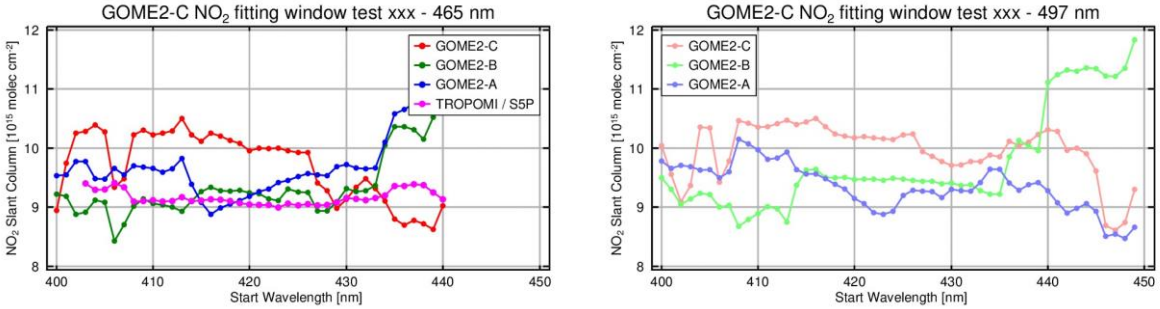

Figure 6-2: NO_2 slant column as retrieved from the GOME-2 instruments in different fitting windows for one orbit over the Pacific (30°N – 60°N) on May 1st, 2019. Shown are the fitting window starting wavelengths on the x-axis for an end wavelength of 465 nm (left) and 497 nm (right). Colours represent different instruments: GOME2-A (blue), GOME2-B (green), GOME2-C (red) and TROPOMI (pink).

When comparing GOME2-C results with those from GOME2-A and GOME2-B retrieved with identical fit settings for the same days it became apparent, that there were systematic offsets between these data. Closer inspection revealed that these differences depend critically on the exact wavelengths chosen for the DOAS retrieval. This is similar to the findings for HCHO discussed in section 5.3. How large this effect is, is illustrated in Figure 6-2 for a series of different starting wavelengths for two often used fitting window end wavelengths (465 and 497 nm). From the figure one can see that

- Variations are of the order of 10 – 30% depending on instrument and wavelength
- Results for the same start wavelength vary with end wavelength
- Each instrument shows its own peculiarities

These results indicate that offsets between NO_2 results of different instruments have to be expected and can only be minimized empirically based on experience. This is actually well known in the DOAS community and the reason for the use of many different fitting windows for DOAS retrievals depending on instrument and scientific research topic.

It is interesting to note that TROPOMI on S5P (see pink line in Figure 6-2) appears to be less affected by this problem as a much smaller dependence of the NO_2 SC on fitting window is found. One possible explanation is that the fact that each of the 450 TROPOMI viewing directions has its own specific characteristics is averaged in this figure, reducing the overall effect. This is linked to the well-known stripes in OMI and TROPOMI retrievals which result from small offsets between SCs of individual rows and are usually removed by an empirical post-fitting correction called destriping.

In summary, offsets are observed between the NO_2 columns retrieved from the different GOME-2 instruments, and these values depend on the exact wavelengths used for fitting. There is however no indication that GOME2-C behaves different or worse than the other two GOME-2 instruments in this respect.

6.1.3 Impact of background spectrum

From experience with other UV-vis instruments it is known, that using the solar irradiance spectrum as background spectrum in the DOAS fits leads to larger fitting residuals than using a reference spectrum generated from Earthshine measurements over the Pacific. Several effects contribute to this result, including

- Incomplete compensation of the wavelength shift induced by the Doppler shift in the irradiance measurements
- Differences in the light path within the instrument (diffuser) when using the irradiance measurement
- Reduction of the optical depth induced by rotational Raman scattering if an Earthshine background is used
- Change in spectral resolution along the orbit which may lead to resolution differences between irradiance and Earthshine measurements.

It is therefore interesting to compare NO_2 fits using irradiance and Earthshine background spectra with respect to the resulting fitting RMS.

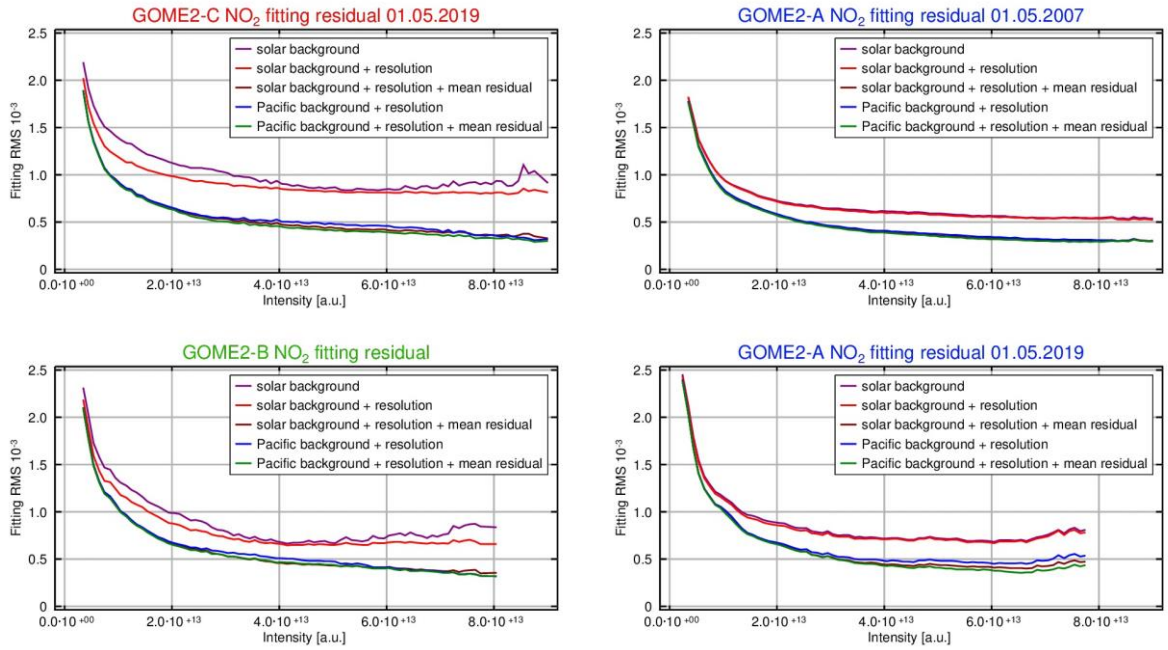

Figure 6-3: Comparison of NO_2 fitting residuals for 5 different fit settings (see text for details).Results are for GOME2-C (top left), GOME2-B (bottom left), GOME2-A (bottom right) and GOME2-A in the first year of operation (top right). Only measurements with short integration times are included

In Figure 6-3, the fitting residual of the standard NO_2 fit in the 425 – 497 nm fitting window is shown as a function of measured intensity. Ideally, the RMS should be dominated by photon shot noise, and thus be proportional to $1/(I^2)$. As can be seen in the figure, a strong decrease in fitting RMS with increasing intensity is indeed observed. The four panels of Figure 6-3 show results for the three GOME-2 instruments for May 1st, 2019 and, for comparison, also GOME2-A results for the same day in 2007, the first year of operation. In each panel, five different analysis settings are included:

- The standard fit using a solar irradiance as background.
- The same fit but including a spectral resolution change correction in order to investigate if spectral resolution differences contribute to the RMS.
- The same fit, but including in addition the mean residual of the fit, removing the systematic component of the residual, arguably introduced by differences between Earthshine and irradiance measurements.
- The standard fit with resolution correction (see below) but using a Pacific Earthshine as background spectrum.
- The same fit, but including the mean residual of the fit, removing the systematic component in the residual.

The resolution correction mentioned above is a simple additional cross-section included in the fit which is created by convolving a high resolution Fraunhofer spectrum with two slit functions of slightly differing width and taking the logarithm of the ratio of the two resulting spectra. This cross-section can compensate for the effect of small differences in spectral resolution between measurement and background spectrum in the DOAS fit (see Azam and Richter, 2015).

For an ideal instrument, all fits should result in the same RMS. In real instruments, where small

differences exist between Earthshine and irradiance measurements (see above), a difference exists between the retrievals with Earthshine and those with irradiance background spectrum. As shown in Figure 6-3, this is clearly the case in all four panels. For GOME-2 in its first year of operation (top right), this is the only relevant effect: The resolution correction does not impact on the results, and including the mean residual in the fit with Earthshine background also does not improve the results. The difference between Earthshine and irradiance fit can be completely removed by including the mean residual in the fit with irradiance background. Thus, for GOME2-A in its first year of operation, the only relevant factor was the difference between irradiance and radiance measurements.

Comparing this with the results for GOME2-A in 2019, the curves have moved to lower intensities as expected from degradation, but the RMS has also increased for the same intensity. In addition, small improvements become apparent when including the mean residuals in the Earthshine fit, indicating that differences between radiance and irradiance are no longer the only relevant effect. Resolution changes do not appear to play a role.

Comparing GOME2-C results with those from the first year of GOME2-A operation shows three important differences:

- All RMS values are larger for GOME2-C than for GOME2-A in the first year
- Including the resolution correction improves the fits
- Including the mean residual in the fit with Earthshine background improves the fits

Comparing GOME2-C results to GOME2-B data for 2019 shows that both instruments behave in a similar way, but GOME2-B NO_2 fits have lower RMS although the instrument is already in its seventh year of operation. In particular, the fits with solar irradiance as background are better than for GOME2-C.

These results indicate that GOME2-C has lower quality spectra in the NO_2 fitting region, and that differences between radiance and irradiance are larger than for the other two GOME-2 instruments. The larger RMS could be linked to the wavelength oscillations discussed in section 3.3, and this could also contribute to the larger RMS when using a solar irradiance. There also appears to be an issue with resolution changes, but this can be seen in GOME2-B data as well.

6.1.4 Long-term stability

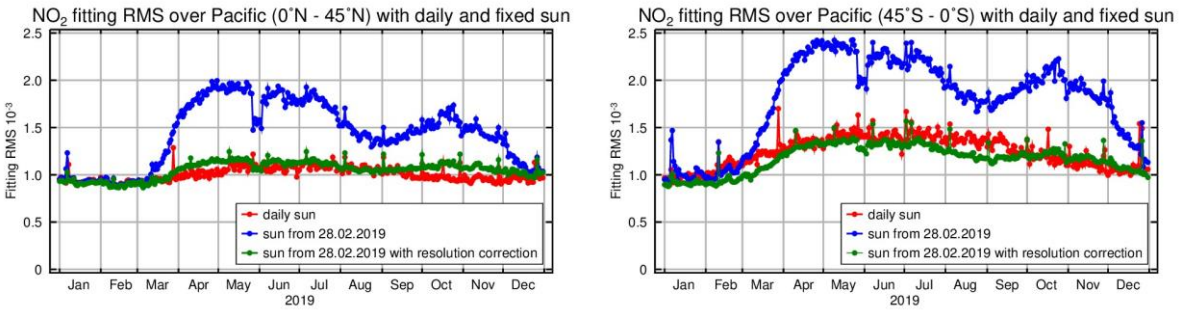

Figure 6-4: GOME2-C NO_2 fitting residuals over the Pacific for the Northern Hemisphere (0° - 45°N, left) and the Southern Hemisphere (0° - 45°S, right). Results from three different retrievals are shown: The standard evaluation with a daily solar irradiance (red), an analysis with a fixed solar irradiance from February 28, 2019 (blue) and a fit with the same fixed irradiance but including a resolution correction function.

As any space borne instrument, GOME2-C will age in space, and this will lead to degradation of

its optical and electronic performance. A large part of that is compensated in the DOAS analysis which always uses a recent measurement as background spectrum I_0. However, in addition to long-term degradation, also other effects such as seasonal changes of instrument temperature, changes in the geometry of Sun observations and also periodic changes of the solar irradiance can contribute to changes in the observed radiances and irradiances.

It is therefore interesting to compare the results of retrievals using different background spectra: the daily solar irradiance or an arbitrarily selected fixed one. When using the fixed background spectrum, instrumental changes should result in reduced fitting quality, indicating degradation and seasonal changes as listed above. In Figure 6-4, the results of such an analysis are shown for NO_2 fits on the first year of GOME2-C data, averaged over the Pacific for tropical and mid-latitudes in the Northern and Southern Hemisphere, respectively. RMS values for three different retrievals are shown: the standard analysis using the daily irradiance, a fit using the solar spectrum of February 28 as background spectrum for all days and the same fit, but including a correction for spectral resolution changes. The latter was included as analysis of the instrument function (section 3.1) which showed a systematic variation of the slit function over the first year.

As shown in Figure 6-4, all three retrievals have similar fitting RMS at the beginning, but over time, the fit using a fixed solar irradiance deteriorates relative to the one with daily irradiance before the two results approach each other again at the end of the year. The patter of the difference is similar to the pattern of slit function changes shown in Figure 3-2, and it is therefore not surprising that including the resolution correction in the fit brings the RMS values into close agreement. Small differences remain with lower RMS for one or the other fit depending on season and hemisphere, arguably as a result of along orbit resolution changes. It is interesting to note that at the end of the year, fitting RMS values are back to where they started in January, indicating that there is no effect of degradation.

These results imply that at least for NO_2, the GOME2-C instrument was very stable over the first year, the only variation being systematic resolution changes which can be fully compensated by including a simple resolution correction in the fit. This is different from the results for OClO which is fitted in channel 2 (see section 0).

6.1.5 Dependence on subset

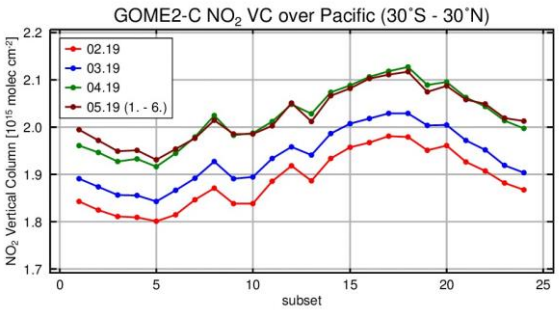

Figure 6-5: Subset dependence of GOME2-C NO_2 VC retrieved over the tropical Pacific for February to May 2019.

The GOME2 instruments are scanning instruments, using a moving mirror to direct light from different across-track scenes into the spectrometer. This is in contrast to imaging instruments such as OMI and TROPOMI, which have no moving mirror but image different across track scenes on different parts of a large 2D-CCD detector. In such setups, slight differences between

the individual pixels on the CCD can lead to offsets between different viewing directions, known as stripes (see also discussion in section 6.1.2). For GOME2-C, no such effect is expected, unless row specific calibration issues would introduce row specific problems. Such problems are known from the polarisation correction, but they usually have a rather smooth behaviour with row. It was therefore a surprise when analysis of GOME2-C NO_2 retrievals showed persistent systematic differences between results for different rows. As shown in Figure 6-5, across-track differences of the order of $2x10^{14}$ molec cm^{-2} are found (10%) which are remarkable constant over time. A small variation in stratospheric NO_2 VC is expected across track for the large GOME2 swath with larger values in the west (larger subset numbers), but the dependence on subset should be nearly linear and the magnitude clearly smaller than observed.

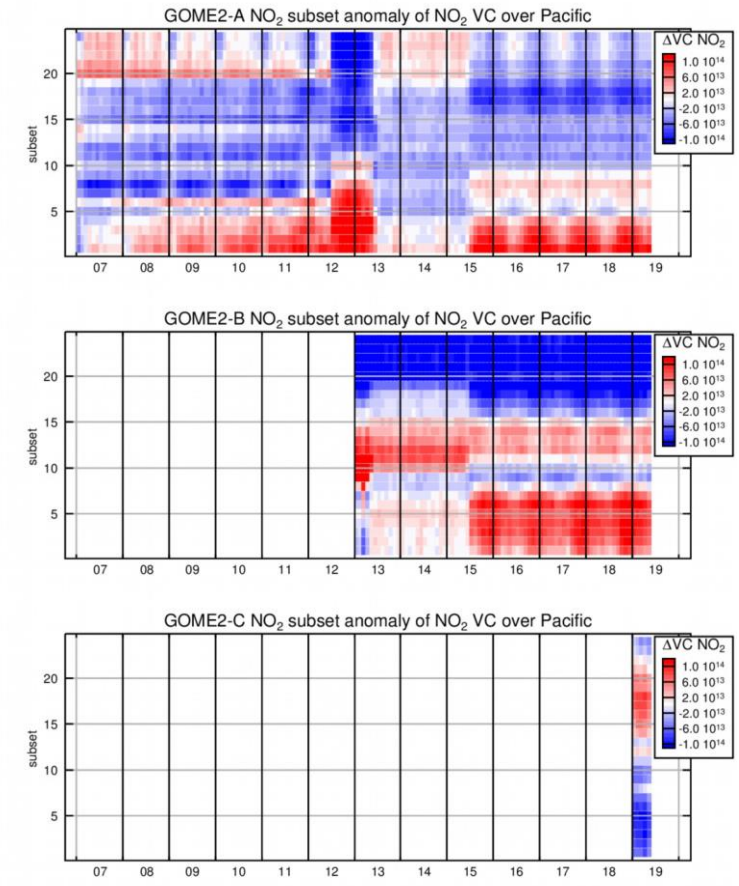

Figure 6-6: Subset depending anomaly of the NO_2 VC over the Pacific for GOME2-A (top), GOME2-B (middle) and GOME2-C (bottom).

Comparison to results from GOME2-A and GOME2-B showed that these also show pronounced NO_2 variations with subset, however with opposite dependence on the subset value. In order to evaluate this in more detail, the full time series of all three satellite instruments was evaluated, and the results are presented in Figure 6-6. As can be seen, the magnitude of the subset dependence is similar in all three instruments; the patterns have some seasonality but are otherwise quite stable over time with the exception of the first year of GOME2-A operation and several sharp changes which can be traced back to changes in lv1 data version.

Summarizing these observations, all GOME2 instruments show systematic and quite large subset dependent offsets in the NO_2 columns which appear to be linked to the scan angle dependent calibration. GOME2-C is no exception to this but there is also no indication that it is more

affected by this than the other GOME2 instruments.

7 References

- Azam, F. and Richter, A., GOME2 on MetOp, Follow-on analysis of GOME2 in orbit degradation, Final Report, EUM/CU/09/4600000696/RM, University of Bremen, 28.09.2015
- Dikty, S. and Richter, A., GOME-2 on MetOp-A, Support for Analysis of GOME-2 In-Orbit Degradation and Impacts on Level 2 Data Products, Final Report, ITT 09/10000262, University of Bremen, 18.01.2012
- Noël, S., Bramstedt, K., Bovensmann, H., Gerilowski, K., Burrows, J. P., Standfuss, C., Dufour, E. and Veihelmann, B.: Quantification and mitigation of the impact of scene inhomogeneity on Sentinel-4 UVN UV-VIS retrievals, Atmos. Meas. Tech., 5(6), 1319–1331, doi:10.5194/amt-5-1319-2012, 2012.
- Rozanov, V. V., Rozanov, A. V., Kokhanovsky, A. A., and Burrows, J. P.: Radiative Transfer through Terrestrial Atmosphere and Ocean: Software Package SCIATRAN, J. Quant. Spectrosc. Ra., 133, 13–71, https://doi.org/10.1016/j.jqsrt.2013.07.004, 2014.
- De Smedt, I., van Geffen, J. , Richter, A., Beirle, S., Yu, H., Vlietinck, J., Van Roozendael, M., van der A, R., Lorente, A., Scanlon, T., Compernolle, S., Wagner, T., Eskes, H. , Boersma, F. , Quality Assurance for Essential Climate Variables, Product User Guide for HCHO (Version 1.0), 2017