# $H_2CO$ COLUMNS RETRIEVED FROM GOME-2: FIRST SCIENTIFIC RESULTS AND PROGRESS TOWARDS THE DEVELOPMENT OF AN OPERATIONAL PRODUCT

**Isabelle De Smedt[1], Jenny Stavrakou[1], Jean-François Müller[1], Nan Hao[2], Pieter Valks[2], Diego Loyola[2], Michel Van Roozendael[1].**

(1) Belgian Institute for Space Aeronomy (BIRA- IASB), Brussels, Belgium
(2) German Aerospace Center (DLR), Wessling, Germany

**Abstract**

Non-methane volatile organic compounds (NMVOCs) of both natural and anthropogenic origin play an important role in the formation of tropospheric ozone. Satellite observations of formaldehyde ($H_2CO$) have been shown to provide useful information to test and improve bottom-up inventories of NMVOC emissions because $H_2CO$ is an intermediate product in the oxidation of a large number of NMVOCs and its lifetime is only a few hours.

Since 1995, global distributions of $H_2CO$ columns have been retrieved from the ESA GOME and SCIAMACHY UV-Vis nadir sounders. With its improved spatial resolution and coverage, GOME-2 on METOP may provide useful new measurements extending the successful time-series established with past instruments. The retrieval of $H_2CO$ by differential absorption spectroscopy (DOAS) is challenging mainly due to the overall faintness of the $H_2CO$ signal, but also to uncertainties in the calculation of air mass factors (AMFs). Interference effects due to combined spectroscopic effects and possible instrumental problems must be carefully investigated and the consistency with past observations must be assessed. We present results from scientific retrievals of $H_2CO$ performed using 2 years of GOME-2 observations, based on the algorithms developed for the GOME and SCIAMACHY instruments.

An adapted version of this scientific algorithm is proposed for implementation in the GOME-2 trace gas operational processor developed at DLR-IMF for the O3-SAF. First results from the corresponding demonstration product are shown.

## SCIENTIFIC RESULTS

### Baseline Slant Columns Retrieval and Air Mass Factors Determination

Despite the relatively large abundance of $H_2CO$ in the atmosphere (on the order of $10^{16}$ molec/cm$^2$), the fitting of $H_2CO$ slant columns in earthshine radiances is a challenge because of the low optical density of $H_2CO$ compared to other UV absorbers. Therefore, the detection of $H_2CO$ is limited by the signal to noise ratio of the measured radiance and by possible spectral interferences due to other molecules absorbing in the same fitting interval, mainly ozone.

$H_2CO$ columns have been previously retrieved from GOME/ERS-2 and SCIAMACHY/ENVISAT. DOAS settings used for these two instruments are documented in the literature (De Smedt et al., 2008). Slant columns are fitted in the 328.5–346nm wavelength range. This choice of the fitting interval minimizes uncertainties due to a polarization anomaly affecting the SCIAMACHY spectra around 350 nm, and to a major absorption band of the $O_4$ collision complex (centred near 360 nm). Furthermore, it decreases fitting residuals in tropical areas and reduces noise over the oceans (De Smedt et al., 2008). The same retrieval approach has been applied to GOME-2 spectra. Although different fitting windows have been tested, the best results and consistency with the previous $H_2CO$ observations have been found in the 328.-5-346 nm interval. Compared to De Smedt et al. (2008), a common set of high–resolution cross-sections, convolved to the resolution of each satellite instrument, has been used to retrieve $H_2CO$ slant columns from GOME, SCIAMACHY and GOME-2 measurements. Particularly, the $H_2CO$ absorption cross sections are those of Meller et al. (2000). The detailed DOAS settings used for $H_2CO$ slant column retrievals are given in Table1.

| **Fitting interval** | 328.5-346 nm |
|---|---|
| **Absorption cross-sections** | |
| $H_2CO$ | Meller et al., 2000 |
| Ozone | Brion et al. 1998, 228°K + 243°K |
| BrO | Fleischmann et al., 223°K |
| $NO_2$ | Vandaele, 2002, 220°K |
| OClO | Bogumil et al., 2003 |
| Ring effect | 2 Ring eigenvectors generated using SCIATRAN |
| **Polynomial** | $5^{rd}$ order (6 parameters) |
| **Intensity offset correction** | Linear offset |

**Table 1. Baseline DOAS settings used for GOME, SCIAMACHY and GOME-2 $H_2CO$ slant column retrieval.**

In order to minimise the fitting residuals, daily radiance spectra are used as the $I_0$ reference. These are selected in a region of the equatorial Pacific Ocean where the formaldehyde column is assumed to be low. The noise on the individual $H_2CO$ measurements was found to be consistent with the ground pixel size of the instruments (approx. 40x80 $km^2$ for GOME-2, 30x60 $km^2$ for SCIAMACHY and 40x320 $km^2$ for GOME). For GOME-2, tests have been performed using daily solar irradiance spectra as reference. The results showed that the quality of the $H_2CO$ slant columns is equivalent, even if the fitting residuals are 10 to 20% higher. Therefore, it has been decided to use solar irradiances as reference in the case of GOME-2.

To reduce the impact of fitting artefacts due to unresolved spectral interferences with ozone and BrO absorptions, an absolute normalisation is applied on a daily basis using the reference sector method (De Smedt et al., 2008). The reference sector is chosen in the central Pacific Ocean (140°–160° W), where the only significant source of $H_2CO$ is the $CH_4$ oxidation. The latitudinal dependency of the $H_2CO$ slant columns in the reference sector is modelled by a polynomial, subtracted from the slant columns and replaced by the $H_2CO$ background taken from the tropospheric IMAGESv2 global chemistry transport model (Stavrakou et al., 2009a) in the same region.

Vertical columns are obtained by dividing the slant columns by air mass factors (AMFs) calculated using scattering weights evaluated from radiative transfer calculations performed with the DISORT code. A correction for cloud effects is applied based on the independent pixel approximation, but aerosols are not explicitly considered in the current version of the dataset. The ground albedo is obtained from the GOME climatology of Koelemeijer et al. (2003). Cloud information is provided by the FRESCO+ algorithm (Wang et al., 2008). For the determination of the AMFs, $H_2CO$ vertical profiles are needed a priori. They are provided by the IMAGESv2 model (Stavrakou et al., 2009a) on a monthly basis and interpolated for each satellite geolocation.

### Retrieval Improvement with GOME-2

As shown on the left panel of Figure 3, the baseline retrieval settings for GOME-2 $H_2CO$ columns offer a very good consistency with the GOME and SCIAMACHY observations in tropical regions. However, in mid-latitude regions, the GOME-2 observations show an underestimation of the $H_2CO$ columns, mainly during the winter period. The earlier GOME-2 overpass time (9.30am) compared to GOME (10.30am) and SCIAMACHY (10.00am), reduces the quality of the signal at mid-latitudes because of higher solar zenith angles. In an attempt to reduce this effect and to minimize the interference with the $O_3$ and BrO absorptions, we have applied two corrections during the DOAS fit of the slant columns:

1. The BrO slant columns are fitted in a larger wavelength interval (328.5-359 nm). The DOAS settings used in this fitting interval are the same as in Table 1. The $H_2CO$ SCD fit is performed in a second step, in the 328.5-346 nm interval, using the BrO slant column as a fixed parameter. Globally, this allows reducing the fitting residuals and the standard deviation on the $H_2CO$ slant columns. Figure 1 illustrates this effect for one particular GOME-2 orbit of April 4 2007. For this orbit, the errors on the SCD are reduced by 40% and the standard deviations on the SCD are reduced by 20%.
2. The ozone absorption cross-sections are corrected by a 2D-AMF table. This table is

wavelength and solar zenith angle dependant and has been calculated off-line with DISORT, using a typical stratospheric ozone profile. This allows reducing the fitting residuals and the $H_2CO$ slant columns underestimation at large solar zenith angles. Figure 2 illustrates the improvement of the VCD in mid-latitudes.

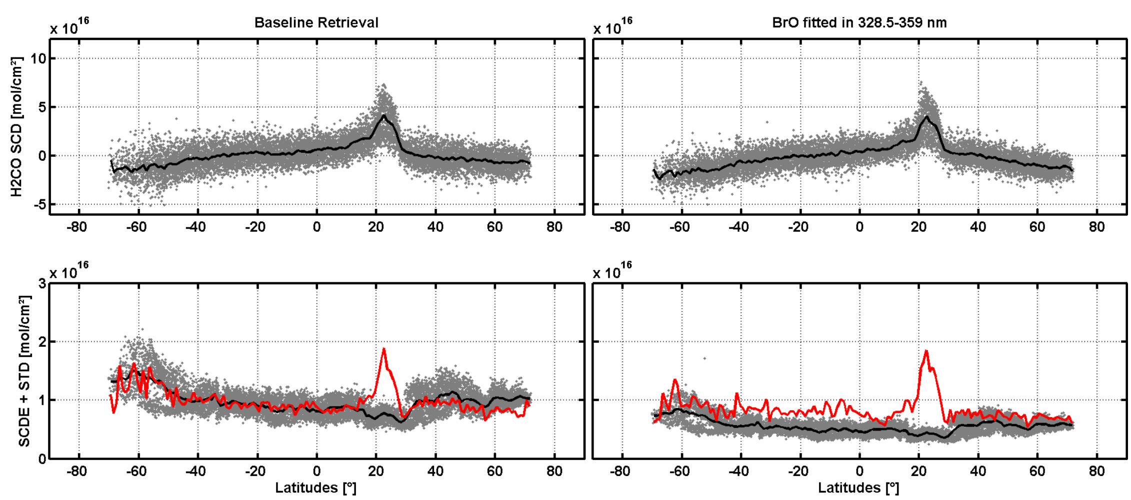


**Figure 1: $H_2CO$ slant columns (grey dots and black line, upper panels), SCD errors (grey dots and black line, lower panels) and SCD standard deviation (red line) for one GOME-2 orbit of the 4th April 2007 (orbit number: 2364). The left panel corresponds to the baseline retrieval. The right panel shows the effect of fitting BrO in the 328.5-359 nm interval.**

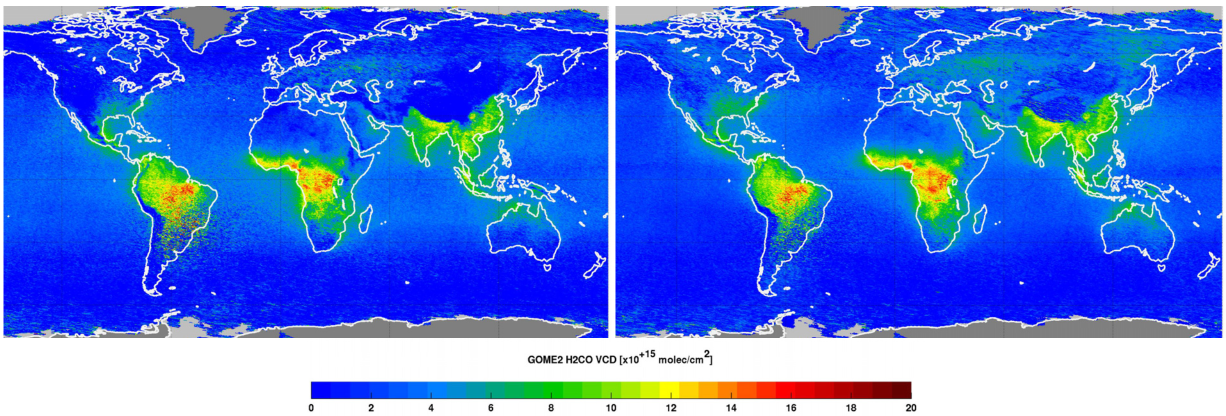


**Figure 2: Averaged GOME-2 $H_2CO$ vertical columns in 2008. The left panel corresponds to the baseline retrieval. The right panel shows the effect of the $O_3$ and BrO corrections, mainly at mid-latitudes.**

### Consistency with GOME and SCIAMACHY and with Ground-Based Observations

GOME, SCIAMACHY and GOME-2 $H_2CO$ vertical columns have been extracted above a selection of emission regions. Monthly averaged $H_2CO$ columns calculated with pixels having a cloud coverage below 40% and a solar zenith angle under 60° are presented in Figure 3. The left panels show the baseline $H_2CO$ GOME-2 retrieval, while the right panels show the results after the GOME-2 retrieval improvements. In Tropical regions like in Indonesia and Amazonia, the agreement between coincident SCIAMACHY and GOME-2 observations is already good with the baseline retrieval settings. In mid-latitude regions, the underestimation of the GOME-2 observations increases with latitude and occur not only in winter but also during the summer period in Southern China and in South-Eastern U.S. This effect is completely removed with the improved GOME-2 retrieval settings. The final agreement with the SCIAMACHY observations is very good everywhere.

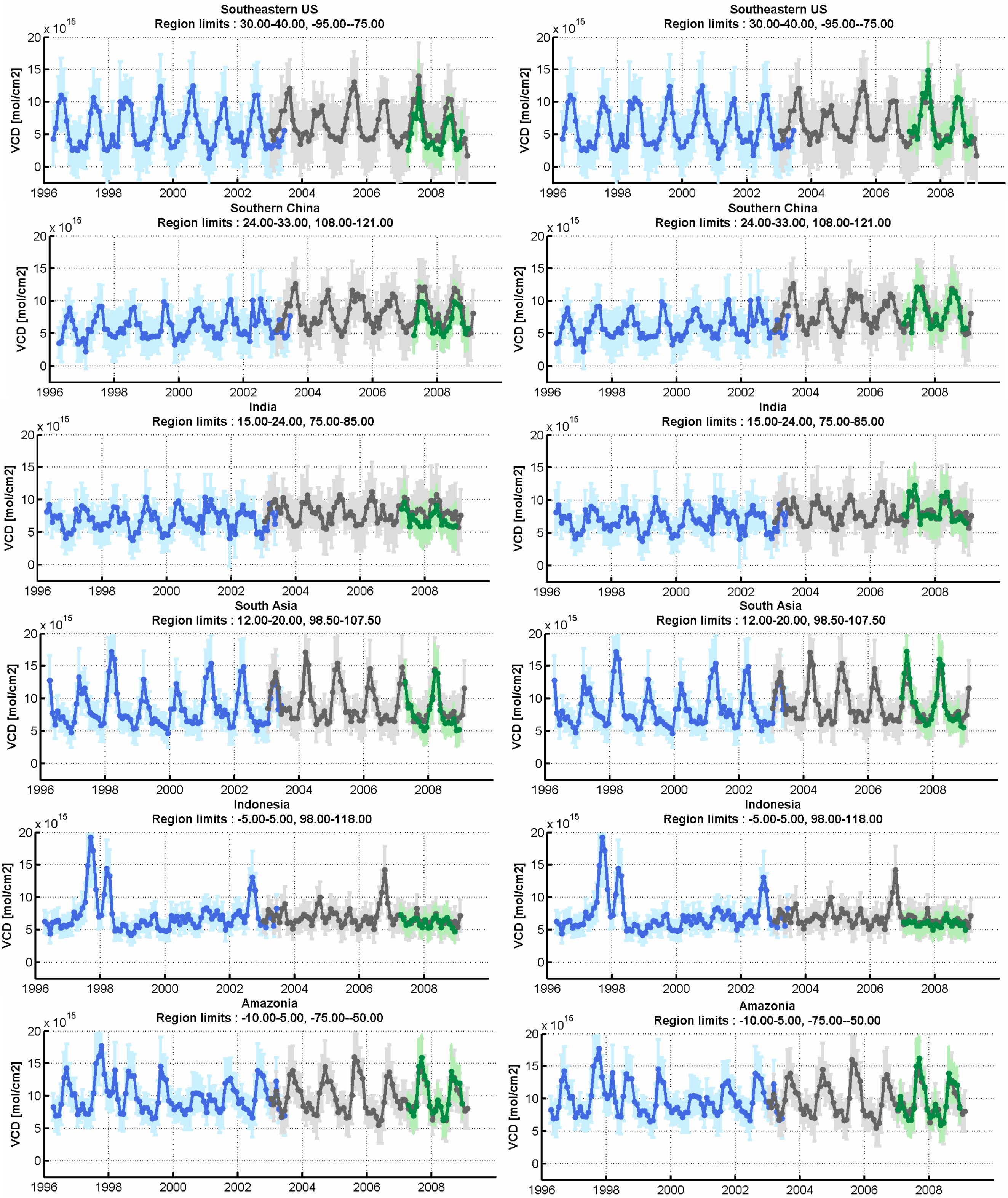


**Figure 3: Regional monthly averaged $H_2CO$ VCD retrieved from GOME (in blue), SCIAMACHY (in grey) and GOME-2 (in green). The left panels show the results obtained with the baseline retrieval settings. The right panels show the results obtained with the $O_3$ and BrO corrections.**

A MAX-DOAS instrument developed at BIRA-IASB was installed in Beijing (China) from June 2008 until April 2009. Figure 4 presents SCIAMACHY and GOME-2 monthly mean vertical columns extracted in a radius of 200 km around Beijing. These are compared to monthly averaged MAX-DOAS tropospheric columns measured at the satellite overpass time. The left panel presents the baseline GOME-2 $H_2CO$ retrieval. The GOME-2 VCDs are underestimated compared to SCIAMACHY. The right panel shows the GOME-2 DOAS retrieval improvements with the $O_3$ and BrO corrections. The agreement with SCIAMACHY results is largely improved. The satellites observations are generally in good agreement with the MAX-DOAS measurements, expected in June when we find a large discrepancy between the satellite and the ground-based observations. Future work will focus on the application of the different averaging kernels of the measurements to take into account the differences in sensitivity to the $H_2CO$ distribution, and on the effect of aerosols on the satellite air mass factors. Aerosols properties can be derived from the MAX-DOAS instrument (Clémer, 2009).

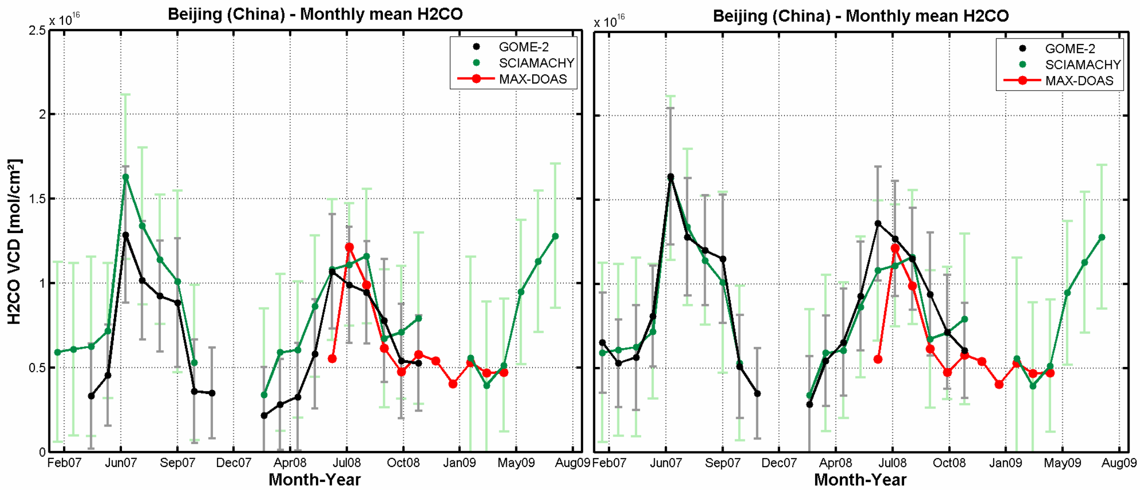


**Figure 4: Monthly averaged $H_2CO$ VCD retrieved from SCIAMACHY (in green) and GOME-2 (in black) in a radius of 200km around Beijing compared with the observations of the MAX-DOAS instrument (in red). The left panel shows the results obtained with the baseline retrieval settings. The right panel shows the results obtained with the $O_3$ and BrO corrections.**

## IMPLEMENTATION IN THE GOME-2 OPERATIONAL PROCESSOR

An adapted version of this scientific algorithm has been proposed for implementation in the GOME-2 trace gas operational processor. Verification and initial validation of the GOME-2 $H_2CO$ total column product, produced by the GOME Data Processor (GDP) version 4.4 operated at DLR on behalf of EUMETSAT, has been performed.

The baseline DOAS settings for the retrieval of $H_2CO$ slant columns have been chosen to be implemented in the GDP 4.4 (see Table 1 for detailed DOAS settings). For the air mass factor determination, the same algorithmic steps and a priori profiles from the IMAGESv2 model are used in the GDP 4.4. However, scattering weights are evaluated with the LIDORT 3.3 radiative transfer model and the ground albedo is obtained from the combined TOMS/GOME climatology (Boersma et al, 2004). In the GDP 4.4, the cloud information (cloud fraction, cloud top-pressure and albedo) are provided by the OCRA and ROCINN algorithms.

For verification purposes, comparisons between the BIRA-IASB and the GDP 4.4 results were performed on a limited set of GOME-2 orbits. Results of the slant column comparisons are illustrated in Figure 5. As can be seen, a good level of agreement was obtained, demonstrating the consistency between the two slant column fitting algorithms. After the reference sector correction, the offset between the slant columns is smaller than $5x10^{14}$ molec/cm², which is well under the noise on the slant columns.

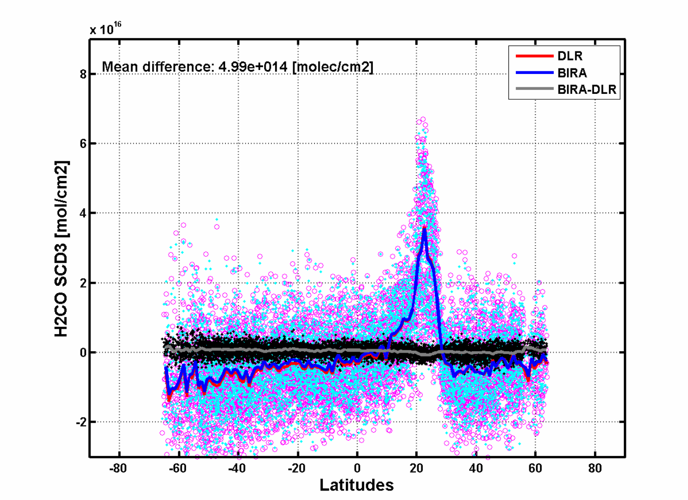


**Figure 5: Comparison of $H_2CO$ slant columns retrieved from GDP 4.4 and from the BIRA-IASB scientific algorithm, after the reference sector correction, for one GOME-2 orbit of the 4$^{th}$ April 2007 (orbit number: 2364).. The black dots correspond to the difference between GDP and BIRA-IASB $H_2CO$ SCDs.**

The comparison of the clear-sky AMF is illustrated on the first panel of Figure 6. The difference between the albedo climatology leads to a systematic difference of a little less than 10%, resulting in higher vertical columns in the GDP 4.4. When using the same albedo climatology, the mean difference is greatly reduced, proving the consistency of the weighting functions calculated respectively with DISORT and LIDORT. The second panel of figure 6 shows the comparison of the total AMF from the BIRA-IASB scientific algorithm (using FRESCO) and the GDP 4.4 (using OCRA-ROCINN). There is a relatively good agreement between the two data sets, the mean difference between the AMF being 0.15± 0.2. Tests have shown that using the same cloud product reduces the mean difference to 0.001± 0.1.

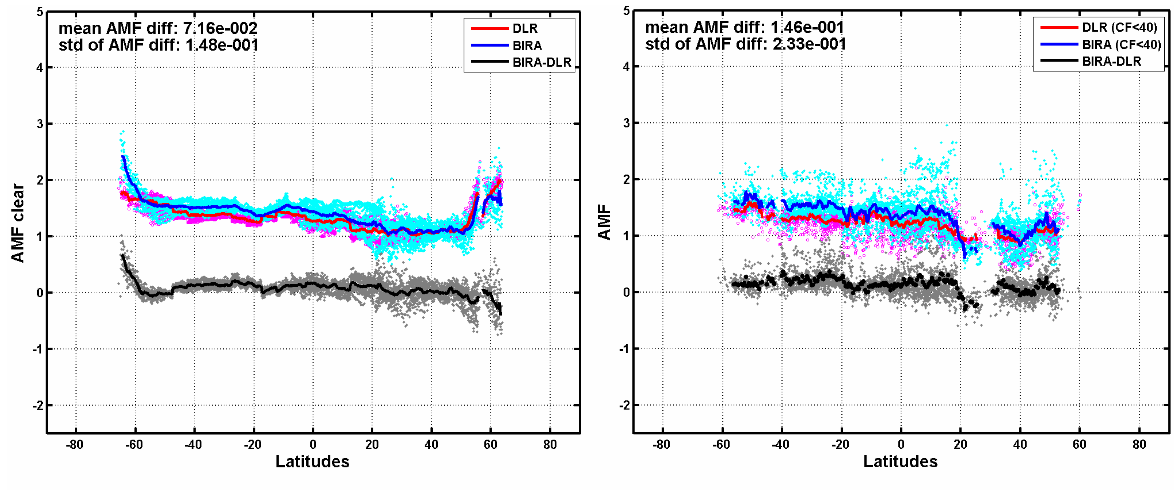


**Figure 6: Comparison of AMF calculated in GDP 4.4 and in the BIRA-IASB scientific algorithm for the same GOME-2 orbit of the 4$^{th}$ April 2007. The first panel shows the clear AMF and the second panel shows the total AMF, for cloud fractions below 40%.**

Figure 7 shows the correlations between the monthly averaged $H_2CO$ columns retrieved from the baseline scientific GOME-2 algorithm and from the GPD 4.4 GOME-2 algorithm. The correlation coefficients are very good in all regions (higher than 0.9, excepted in Indonesia where we find 0.76). However, the GDP 4.4 $H_2CO$ vertical columns tend to be larger than the baseline scientific retrieval (of about 15% in average). These differences between the scientific and the GDP 4.4 vertical columns, starting from very consistent slant columns, show the impact of the cloud product, and to a lesser extend, of the albedo climatology, used in the calculation the air mass factors. More validation of these external data is needed to improve the accuracy of the satellite observations.

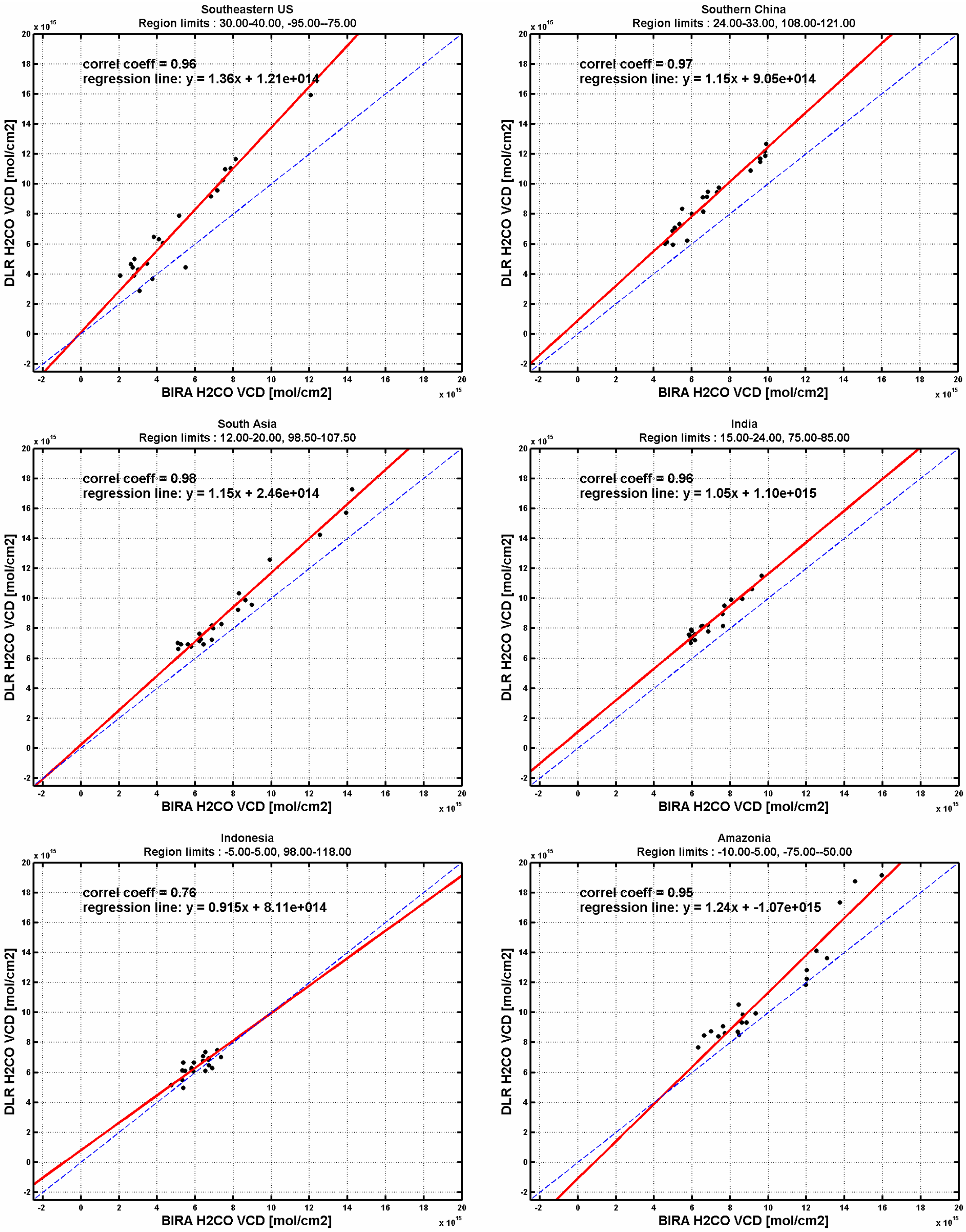


**Figure 7: Correlations between the baseline Scientific GOME-2 retrieval and the GPD 4.4 GOME-2 retrieval of $H_2CO$ columns for the same regions as Figure 3.**

## CONCLUSIONS

The retrieval of $H_2CO$ columns from GOME-2 measurements has been developed at BIRA-IASB. Baseline settings, based on the experience gained from the GOME and SCIAMACHY retrievals, have been set up for GOME-2. In Tropical regions, the GOME-2 $CH_2O$ observations show a good consistency with the GOME-SCIAMACHY time-series. However, at mid-latitudes, the earlier GOME-2 overpass time reduces the quality of the signal and the GOME-2 $H_2CO$ columns are underestimated. This effect has been attenuated in the scientific product after the implementation of the $O_3$ and BrO corrections during the DOAS fit. The satellites observations are generally in good agreement with the MAX-DOAS measurements in Beijing. Future work will focus on taking into account the different column averaging kernels, and the effect of aerosols on the satellite air mass factors. The scientific product is available on the TEMIS web site (www.temis.nl). The operational product is under development at DLR-IMF for the O3-SAF. The baseline DOAS settings and a modified version of the AMF calculation have been implemented in the GDP 4.4 algorithm. The verification of this new GOME-2 O3M-SAF product has been performed. Although the GDP 4.4 $H_2CO$ columns generally tend to be slightly higher than those of the adapted scientific product, both data products show a very good level of consistency when considering the typical uncertainties on $H_2CO$ retrieval. Our study indicates that cloud effects have an important impact of the resulting $H_2CO$ data product. Future research work will focus on better assessing these effects in relation to the cloud parameters used as an input for the trace gas retrieval. For further validation of the $H_2CO$ satellite observations, more validation sites are needed, as well as longer observation campaigns.

## REFERENCES

Bogumil, K., *et al.*, Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: Instrument characterization and reference data for atmospheric remote sensing in the 230-2380 nm region, J. Photochem. Photobiol. A.: Photochem., 157, 167-184, 2003.

Brion, J., *et al.*, J., Absorption spectra measurements for the ozone molecule in the 350-830 nm region, J. Atmos. Chem., 30, 291-299, 1998.

Brion, J., *et al.*, J., High-resolution laboratory absorption cross section of O3. Temperature effect, Chem. Phys. Lett., 213 (5-6), 610-512, 1993.

Clémer, K. et al., Simultaneous retrieval of tropospheric aerosol extinction and $NO_2$ vertical profiles from MAXDOAS measurements in Beijing, Proceedings of the 8th International Symposium on Tropospheric Profiling 19–23 October, Delft, The Netherlands, Edited by A. Apituley, H.W.J. Russchenberg, W.A.A. Monna, ISBN 978-90-6960-233-2, 1999.

De Smedt I, *et al.*, M., Twelve years of global observation of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. Atmos. Chem. Phys., 8 (16), 4947-4963, 2008.

Fleischmann, O. C., *et al.,* New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. Photochem. Photobiol. A, 168, 117–132, 2004.

Koelemeijer, R. B. A., *et al.*, A database of spectral surface reflectivity in the range 335-772nm derived from 5.5 years of GOME observations, J. Geophys. Res., 108(D2), 4070, doi:10.1029/2002JD002429, 2003.

Meller, R. and Moortgat, G. K., Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323K in the wavelength range 225–375 nm, J. Geophys. Res., 105(D6), 7089–7101, 2000.

Stavrakou, T., *et al.*, A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, Atmos. Chem. Phys., 9, 20 1037–1060, 2009a.

Stavrakou, T., *et al.,* Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006, Atmos. Chem. Phys., 9, 3663–3679, 25 2009b.

Vandaele, A.C., *et al.*, High-resolution Fourier transform measurement of the $NO_2$ visible and near-infrared absorption cross-section: Temperature and pressure effects, J. Geophys. Res. 107, D18, 4348, doi:10.1029/2001JD000971, 2002.

Wang, P., *et al.*, M., FRESCO+: an improved $O_2$ A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8, 6565-6576, 2008.

### Acknowledgments

The scientific retrieval of $H_2CO$ from GOME, SCIAMACHY and GOME-2 measurements has been developed in the framework of the ESA-DUE/TEMIS, ESA-GSE/PROMOTE and EU-FP6/AMFIC projects. The SCIAMACHY and GOME-2 $H_2CO$ retrieval are now part of the EU-FP7/MACC project. The implementation work of the scientific $H_2CO$ retrieval algorithm in the GOME-2 trace gas operational processor developed at DLR-IMF has been supported by the O3-SAF.