**Tropospheric formaldehyde**

**Introduction**

Formaldehyde (CH$_2$O) is a major intermediate gas in the oxidation of methane and many other hydrocarbons. The lifetime of formaldehyde is short, and the photolysis reactions and reaction with OH form a major source of CO. Because of the short lifetime of several hours, the presence of formaldehyde signals hydrocarbon emission areas. Formaldehyde is important, since it is a measure of the total amount of oxidised hydrocarbons, and together with NOx quantifies the chemical ozone production. The presence of elevated levels of CH$_2$O is related to the release of hydrocarbons (e.g. ethene, isoprene, and methane) by forests, biomass burning, traffic and industrial emissions.

Formaldehyde has been retrieved from GOME observations by Thomas et al. [1998] and by Chance et al. [2000]. Within PROMOTE the retrieval of CH$_2$O is based on a combined retrieval/modelling approach which is similar to the approach for NO$_2$. The main motivation for this new approach is to provide realistic CH$_2$O distributions to the retrieval. A chemistry-transport model, driven by high-quality meteorological fields, will provide best-guess profiles of CH$_2$O, based on the latest emission inventories, atmospheric transport, photochemistry, lightning modelling and wet/dry removal processes. These model forecast fields are collocated with the GOME/SCIAMACHY/OMI observations, and the radiative transfer modelling in the retrieval will be performed based on the model trace gas profiles. The retrieval is coupled to cloud top height and cloud fraction retrievals derived from the GOME/SCIAMACHY/OMI data, and the retrieval will be coupled to high quality albedo maps.
## Product specification

### Product description

| Summary                              | Daily tropospheric CH₂O column |

### Product properties

<table>
<thead>
<tr>
<th>Parameter(s)</th>
<th>Tropospheric column of CH₂O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Error on the tropospheric column of CH₂O</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Specified as parameter</td>
</tr>
<tr>
<td>Geometric resolution</td>
<td>320 x 40 (GOME) and 120 x 30 km² (SCIAMACHY) except for high latitudes</td>
</tr>
<tr>
<td>Grid / projection</td>
<td>Orbit geometry</td>
</tr>
<tr>
<td>Spatial coverage</td>
<td>Swath 960 km (both GOME and SCIAMACHY), 14 orbits per day, global coverage in 3 (GOME) to 6 (SCIAMACHY) days. After cloud filtering global coverage is obtained in approximately 1-2 weeks.</td>
</tr>
<tr>
<td>Temporal coverage</td>
<td>Daily</td>
</tr>
<tr>
<td>Data format</td>
<td>HDF</td>
</tr>
<tr>
<td>Availability</td>
<td>semi-operational implementation in PROMOTE by end of 2005</td>
</tr>
</tbody>
</table>

### Production process

| Method/algorithm                     | SCD retrieval of NO₂ with GwinDOAS  |
|                                      | Tropospheric VCD retrieval of CH₂O based on the atmospheric model TM4, radiative transfer model DAK, and FRESCO cloud parameters. |
| Model / assimilation                 | A-priori CH₂O profile from TM4 model |

### Quality standards

| Production                          | Detailed error calculation, error checking and output range control |
| Validation                          | Comparison ground observation within ACVE validation campaign |

### Input data

| EO data                             | SCI_NL_1P, GOME lv1  |
|                                     | Combined TOMS-GOME albedo climatology |
| Other data | Meteorological data of ECMWF  
|           | GTOPO30 (topology database)  
|           | Emission databases (EDGAR, etc) |

**Optional or other specific properties (if applicable)**

| Historical archive | GOME (ERS-2) : 1996-2003  
|                    | SCIAMACHY (Envisat) : 2003-today |
| Offline/NRT       | Offline |
| Visualization standards | Images available. |

**Underlying primary user requirement(s)**

| Key requirement | Long-term dataset |
| Originator(s)   | Chemical data assimilation service |

**Table 1: Characteristics of the CH₂O product**
**System overview**

The first step in retrieving tropospheric CH$_2$O is performed by BIRA-IASB and covers the spectral fitting of SCIAMACHY Lv1 data to generate so-called slant column densities (scd’s) of CH$_2$O. These scd’s will then be matched to KNMI FRESCO cloud parameters to obtain a pre-processed input dataset for the complete retrieval scheme.

<table>
<thead>
<tr>
<th>Number</th>
<th>Origin</th>
<th>Short name</th>
<th>Description</th>
<th>Backup</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EO</td>
<td>SCI_NL_1P</td>
<td>GOME, SCIAMACHY earthshine reflectance spectra</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GOME lv1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ECMWF</td>
<td>meteo</td>
<td>Meteorological fields</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

**Table 2: Input data of the formaldehyde service**

**External interface(s)**

For the processing direct ftp-transfer of input data from ESA is provided (through AO projects). Meteorological information is provided by ECMWF via ftp-transfer. Results will be distributed through the PROMOTE web-interface.
Retrieval steps

<table>
<thead>
<tr>
<th>Number</th>
<th>Type</th>
<th>Process name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Retrieval</td>
<td>SCD data retrieval</td>
</tr>
<tr>
<td>2</td>
<td>Retrieval</td>
<td>FRESCO cloud retrieval</td>
</tr>
<tr>
<td>3</td>
<td>Retrieval</td>
<td>Tropospheric CH$_2$O column retrieval</td>
</tr>
<tr>
<td>4</td>
<td>Post-processing</td>
<td>Error calculation and averaging kernel calculation</td>
</tr>
</tbody>
</table>

Table 3: Sub-processes of the formaldehyde service

Retrieval of the Slant Column Density of CH$_2$O (process 1)

The technique used to retrieve total slant columns (SCDs) of CH$_2$O from GOME measurements is the Differential Optical Absorption Spectroscopy (DOAS). The DOAS technique was first developed in the late seventies for ground-based remote sensing [Platt, 1994]. The spectral analysis is performed using WinDOAS, a multi-purpose DOAS analysis software developed over the nineties at BIRA-IASB [Vandaele et al., 2005].

Formaldehyde SCDs are retrieved in the 337.5-359 nm spectral range, after optimisation of the wavelength calibration of the GOME radiances using a NLLS fit to the Kurucz solar atlas [Kurucz, 1984; Chance and Spurr, 1997] degraded to the GOME spectral resolution. The CH$_2$O absorption cross-sections applied in the DOAS fit are the one of Meller et al. [2000] convolved to GOME resolution. The fitting also includes reference spectra for interfering species (O$_3$, NO$_2$, BrO, and the O$_2$-O$_2$ collision), and the pseudo-absorption features due to Rotational Raman Scattering (RRS) by air molecules are treated using two eigenvectors of Ring effect cross-sections [Vountas et al., 1998]. The DOAS procedure accounts also for the GOME undersampling [Chance, 1998] and a linear offset correction is applied as well as a polynomial closure term of order 5.

In order to avoid the GOME diffuser plate related artefacts [Richter and Wagner, 2001], Fraunhofer radiance spectra are selected on a daily basis in the equatorial Pacific Ocean, in a region where the formaldehyde column is assumed to be negligible. To further reduce the impact of zonally invariant artefacts, mainly due to ozone misfits in the CH$_2$O fitting window, an absolute normalisation is applied on a daily basis using the reference sector method (RSM) [Khokhar et al., 2004].

Because of the diffuser plate issue and the interference of ozone there is a need to fix the GOME retrievals to estimated background levels of formaldehyde. Background values over the Pacific ocean are dominated by the reaction of methane with OH. Methane is a long-lived gas and mixing ratio’s are known to within a few percent. Free troposphere background OH levels determine the lifetime of methane and are relatively well known, to an estimated accuracy of 20-30%. In order to fix the zero-level of the retrievals, the GOME retrievals are scaled with a latitude-dependent factor to the monthly-mean 10:00 local time CH$_2$O column values in the (-180,-150) degree longitude reference sector.
Fresco cloud detection (process 2)

Clouds influence the depth of gaseous absorption lines, particularly if the trace-gas concentration is high in the troposphere. Therefore, for accurate trace-gas column density retrievals, the presence of clouds must be taken into account. This can be done by exploiting cloud properties derived from the O₂ A band measurements from SCIAMACHY. The most important cloud parameters needed to correct trace-gas column density retrievals for the presence of clouds are cloud fraction, cloud optical thickness and cloud top pressure [Koelemeijer and Stammes, 1999]. However, it is almost impossible to derive uniquely both cloud fraction and cloud optical thickness from the measured spectral reflectivity in and around the oxygen A band. This is because cloudy scenes with the same cloud top pressure may possess different cloud fractions and cloud optical thicknesses, which give rise to nearly the same reflectivity. However, the retrieved column is mainly sensitive to the radiation level. Therefore, it is useful and necessary to introduce an effective cloud fraction, which is the cloud fraction derived from the satellite measurements, assuming an a priori chosen cloud optical thickness or cloud albedo. The cloud parameters retrieved with FRESCO consist of the effective cloud fraction and cloud top pressure. Alternatively, separation of cloud fraction and cloud optical thickness could be done using PMD information, although ambiguity remains to some extent in that approach too.

The FRESCO method was originally developed for near-real-time ozone column retrieval from GOME. The FRESCO method for GOME is described in the following papers. In [Koelemeijer et al., 2001], the FRESCO method is described together with a sensitivity study and validation using ATSR-2 data. In [Koelemeijer et al., 2002], a comparison is made between cloud top pressures and effective cloud fractions of FRESCO and ISCCP on a monthly average basis. The reader is referred to these papers for a general description of the method and validation.

Tropospheric CH₂O column density retrieval (process 3)

The retrieval approach for tropospheric columnar CH₂O is the application of an air mass factor to convert the slant column into a vertical column. For practical applications of DOAS it is important that the trace gas under investigation has a small absorption optical thickness in the predefined spectral window.

A chemistry-transport model, driven by realistic meteorological fields, provides best-guess profiles of CH₂O, based on the latest emission inventories, atmospheric transport, photochemistry, and wet/dry removal processes. These model forecast fields are collocated with the satellite observations, and the radiative transfer modelling in the retrieval is performed based on the model trace gas profile. The retrieval is coupled to cloud top height and cloud fraction retrievals derived from the satellite data (process 2).

Error analysis and kernels (process 4)

For a detailed discussion on the error analysis approach, the reader is referred to Boersma et al. [2003]. The most important error sources in the retrieval of tropospheric CH₂O columns:

- Errors in the slant column density: this includes measurement noise and DOAS fit related errors.
- Errors in the tropospheric air-mass factor: this includes terms originating from errors in the retrieved cloud fraction, cloud top height, the albedo data set and rough estimates of the profile shape related errors.

The tropospheric CH₂O products are provided with error estimates based on these error contributions. Error estimates are given on a quantitative and pixel-to-pixel basis.
Averaging kernels are provided as part of the data product. Kernels describe how the retrieved vertical column is related to the trace gas profile $x_a$, i.e. $\hat{N}_v = A \cdot x_a$. The kernels describe the sensitivity of the measurement to formaldehyde at different altitudes. The kernels are essential for detailed comparisons with chemistry-transport model fields, with independently measured profiles and or data assimilation. The elements of the averaging kernel are equal to the elements of the altitude dependent air mass factor divided by the total air mass factor.

References