EUMETSAT
AC SAF
ATMOSPHERIC COMPOSITION MONITORING

ALGORITHM THEORETICAL BASIS DOCUMENT

&

PRODUCT USER MANUAL

GOME-2 NO_2 and H_2O Level 3 Climate Products

Prepared by: M. Grossi and P. Valks German Aerospace Centre

Products:

Identifier	Name	Acronym
O3M-87	Level-3 total and trop. NO_2 (GOME-2A & B)	MXG-DS-TCDRNO2
O3M-88	Level-3 total H_2O (GOME-2A & B)	MXG-DS-TCDRH2O

Reporting period: GOME-2/MetOp-A Jan 2007 – Aug 2017

GOME-2/MetOp-B Jan 2013 – Aug 2017

Input data versions: GOME-2A Level 2 GDP 4.8

GOME-2B Level 2 GDP 4.8

document type	AC SAF ATBD
revision	1/D
date of issue	March 2017
products	MXG-DS-TCDRNO2 MXG-DS-TCDRH2O
identifier	O3M-87, O3M-88
product version	1.0

distribution

Function	Organisation
AC SAF	DLR

document change record

Issue	Rev.	Date	Section	Description of Change
1	A	19.09.2016	all	Creation of this document
1	B	15.12.2016	all	Several updates after PCR
1	C	31.03.2017	all	Document name updated to reflect that the content is both ATBD and PUM
1	D	06.11.2017	all	Several updates after DRR

EUMETSAT SATELLITE APPLICATION FACILITY ON ATMOSPHERIC COMPOSITION MONITORING (AC SAF)

Background

The need for atmospheric chemistry monitoring was first realized when severe loss of stratospheric ozone was detected over the Polar Regions. At the same time, increased levels of ultraviolet radiation were observed.
Ultraviolet radiation is known to be dangerous to humans and animals (causing e.g. skin cancer, cataract, immune suppression) and having harmful effects on agriculture, forests and oceanic food chain. In addition, the global warming - besides affecting the atmospheric chemistry - also enhances the ozone depletion by cooling the stratosphere. Combined, these phenomena have immense effects on the whole planet. Therefore, monitoring the chemical composition of the atmosphere is a very important duty for EUMETSAT and the world-wide scientific community.

Objective

The main objectives of the AC SAF is to process, archive, validate and disseminate atmospheric composition products (O_3, NO_2, SO_2, BrO, HCHO, H_2O and OClO), aerosol products and surface ultraviolet radiation products utilising the satellites of EUMETSAT. The majority of the AC SAF products are based on data from the GOME-2 spectrometers onboard Metop-A and Metop-B satellites.

Another important task of the AC SAF is the research and development in radiative transfer modelling and inversion methods for obtaining long-term, high-quality atmospheric composition products from the satellite measurements.

Product categories, timeliness and dissemination

Data products are divided in two categories depending on how quickly they are available to users:

Near real-time products are available in less than three hours after measurement. These products are disseminated via EUMETCast, WMO GTS or internet.

- Near real-time trace gas columns
    - O_3, NO_2, HCHO, SO_2
- Near real-time ozone profiles
    - coarse and high-resolution
- Near real-time absorbing aerosol indexes
    - from main science channels and polarization measurement detectors
- Near real-time UV indexes
    - clear-sky and cloud-corrected

Offline products are available in two weeks after measurement and disseminated via dedicated web services at EUMETSAT, FMI and DLR.

- Offline trace gas columns
  - O_3, NO_2, SO_2, BrO, HCHO, H_2O and OClO
- Offline ozone profiles
  - coarse and high-resolution
- Offline absorbing aerosol indexes
  - from main science channels and polarization measurement detectors
- Offline surface UV

More information about the AC SAF project, products and services:

http://acsaf.org/

AC SAF Helpdesk: helpdesk@acsaf.org

CONTENTS

ACRONYMS AND ABBREVIATIONS

AC SAF	SAF on Atmospheric Chemistry Monitoring
AMF	Air Mass Factor
BRDF	Bidirectional Reflectance Distribution Function
DLR	German Aerospace Centre
DOAS	Differential Optical Absorption Spectroscopy
ECV	Essential Climate Variable
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FDCR	Fundamental Climate Data Record
FMI	Finnish Meteorological Institute
GCOS	Global Climate Observing System
GDP	GOME Data Processor
GOME	Global Ozone Monitoring Experiment
H_2O	Water Vapour
IMF	Remote Sensing Technology Institute
ITCZ	Intertropical Convergence Zone
LOS	Line Of Sight
NIR	Near-infrared
NO_2	Nitrogen Dioxide
REMSS	Remote Sensing System
RMSE	Root Mean Square Error
SAD	Scan Angle Dependency
SCD	Slant Column Density
SSD	Service Specification Document
SSMIS	Special Sensor Microwave Imager Sounder
SZA	Solar Zenith Angle
TCDR	Thematic Climate Data Record
TCWV	Total Column Water Vapour
UPAS	Universal Processor for UV/VIS Atmospheric Spectrometers
VIS	Visible
VCD	Vertical Column Density

1. INTRODUCTION

1.1 Purpose and scope

The purpose of this document is to present the theoretical basis and the user manual of the GOME-2 NO_2 and water vapour Level 3 climate products. These products contain global monthly mean total and tropospheric NO_2 and water vapour columns. The "Level" terminology is used to denote broad categories of GOME-2 data products: Level 0 (L0) denotes raw spectral channel counts, Level 1B (L1B) denotes calibrated and geo-located radiances, Level 2 (L2) denotes orbital-swath science products, and finally Level 3 (L3) denotes global-gridded science products.

1.2 GOME-2 NO_2 and H_2O Level 3 climate products

Satellite instruments provide essential means of obtaining observations of the climate system from a near-global perspective. Therefore the future of the global climate observing system depends critically upon a major satellite component. A key requirement for climate data records is long term continuity, consistency and stability which can be best achieved in stable operational environments such as provided by the AC SAF. Following recommendations from the Global Climate Observing System (GCOS), satellite data contributing to the determination of long-term records must be part of a system implemented and operated so as to ensure that these data are accurate and adequately homogeneous. Requirements on satellite climate data records have been established by GCOS in their supplements 107 and 154 (GCOS, 2006, 2011). In particular GCOS defines a set of Essential Climate Variables (ECVs) for climate monitoring and studies, including both NO_2 and water vapour.

NO_2 plays a key role in air quality and atmospheric chemistry. It is an air pollutant affecting human health and ecosystems, and an important ozone precursor. In addition, NO_2 is involved in ozone depletion processes in the stratosphere and it is important for climate change studies, because of the indirect effect on the global climate (Shindell et al., 2009). The knowledge of the effective distribution of the total column water vapour (TCWV) is fundamental for weather monitoring as well as for the evaluation of climate models. Advancing in understanding of variability and changes in water vapor is vital, especially considering that, in contrast to most other greenhouse gases, the H_2O distribution is highly variable.

Following the definition adopted by EUMETSAT in its rolling Climate Service Development Plan (CSDP), we distinguish between Fundamental Climate Data Records (FCDRs) and Thematic Climate Data Records (TCDRs). While FCDRs derived from GOME-2 Level-1b data are produced directly by EUMETSAT, the generation of TCDRs derived from GOME-2 Level-2 and -3 data fall under the responsibility of the AC SAF. The monthly L3 NO_2 and water vapour column products from GOME-2 are generated with consolidated algorithms applied in a consistent way to a homogeneous (reprocessed) FCDR Level-1 data-set, which qualifies them as TCDRs.

1.3 Product Description

The algorithm used for the generation of the L3 monthly gridded products takes as input the Level 2 reprocessed GOME-2 NO_2 and H_2O column data generated by DLR using the GOME Data Processor (GDP) version 4.8 [ATBD, PUM]. GOME-2/MetOp-A data are available over the time period from January 2007 onwards, and GOME-2/MetOp-B data from January 2013 onwards.

All L3 GOME-2 atmospheric products data are organized into user-friendly and self-describing NetCDF-4 (Network Common Data Form) files, based upon the platform (MetOp-A or MetOp-B) and the temporal period of collection (Monthly data set). The NO_2 and H_2O L3 climate data products are stored in separate files.

The logical file name convention is:

SENSOR_GAS_LV_MM_YYYY_RV.TYPE

The meaning of the different subset of character string is given below:

- SENSOR denotes the instrument, i.e. GOME-2A / GOME-2B
- GAS is the name of the trace gas included in the product, i.e. NO_2 or H_2O
- LV is the product level, i.e. L3
- MM and YYYY are the month and year of the processed products
- RV is the two digit product review
- TYPE denotes the product format used, i.e. NetCDF-4 (.nc)

Global and group attributes for the files common to all products are listed in the following table:

Table 1-1: List of Global and Group attributes with typical values and a short description.

Global Attributes	typical values	Description
Description	Level 3 NO_2 data Level 3 Water Vapour data	short description of the file content
Conventions	CF-1.6	Climate and Forecast convention for variable names, units and dates
Filename	GOME_NO2_Global_YYYY MM_METOPA_DLR_RV.nc GOME_H2O_Global_YYYY MM_METOPB_DLR_RV.nc	The file name starts with the sensor name, followed by the content (NO2 or H2O L3), the coverage, the year and the month, the platform and ends with the version number and type

Group Attributes	typicale values	Description
composite_type	1_month	Gridded monthly data set
institution	DLR Deutsches Zentrum für Luft und Raumfahrt	Name of the institution responsible for the data
reference	https://acsaf.org/datarecord_access.html	
creator_name	Pieter Valks	Name of the person responsible for the data
creator_email	Pieter.Valks@dlr.de	contact address
processingTime	YYYY-MM-DD HH:mm:ss	date and time (UTC) when this file was created
baseProduct	Level 2 GDP	Basis of this data product
baseProductVersion	v 4.8	version of the underlying data
productAlgorithmVersion	v 1.0	Version of this data product
productFormatType	netCDF	File type
productFormatVersion	4	type version
project	EUMETSAT AC SAF	
geospatial_lat_min	-90	minimum latitude (degree)
geospatial_lat_max	90	maximum latitude (degree)
geospatial_lat_resolution	0.25 / 0.5	Lat. resolution (degree)
geospatial_lat_units	degrees North	
geospatial_long_min	-180	minimum longitude (degree)
geospatial_long_max	180	maximum longitude (degree)
geospatial_long_resolution	0.25 / 0.5	Lon. resolution (degree)
geospatial_long_units	degrees East	
sensor	GOME_2	sensor name
platform	MetOp-A/B	Satellite
time_coverage_start	YYYYMMdd	the first day of measurement period
time_coverage_end	YYYYMMdd	the last day of measurement period
product_ID	O3M-87 / -88	
DOI	10.15770/EUM_SAF_O3M_0020 / _0021	http://dx.doi.org/10.15770/EUM_SAF_O3M_0020

The GOME-2 L3 gridded data fields are grouped in the NetCDF file as follows (see also Chap. 3):

PRODUCT group:

- Total and tropospheric NO_2 or Water Vapour content
- Errors associated to NO_2 or Water Vapour column content
- Statistical information
- DETAILED_RESULTS/SUPPORT DATA sub-group:
    - CLOUD_PARAMETERS:
        - Cloud Fraction
        - Cloud Height
        - Cloud Albedo

- SURFACE_PROPERTIES:
  - Surface Albedo
  - Surface Height
  - Surface Flag

Each L3 product is produced at a spatial resolution of 0.25°×0.25° for NO_2 and 0.5°×0.5° for H_2O, from the complete set of L2 orbit files that span a particular month. The latitude/longitude grid defines the dimensions of the data set (see Table 1.2). A number of statistical summaries are computed for each of these L3 products, depending on the parameters being considered (e.g. standard deviations, fraction of pixels which satisfy some condition for the surface flag).

Table 1-2: Overview of the dimension in the data file.

Dimension			
Dimension name	**Unit**	**Size**	**Description**
Latitude	Degree north (-90, 90)	720 / 360	0.25°/0.5° latitudinal resolution
Longitude	Degree east (-180, 180	1440 / 720	0.25°/0.5° longitudinal resolution

1.4 Structure of the document

Section 2 gives an overview of the GOME-2/MetOp-A and GOME-2/MetOp-B instruments together with a brief description of the level 1-to-2 algorithm used for the retrieval of the total and tropospheric NO_2 and H_2O columns. We also introduce the GOME-2 cloud parameters (which are included in the support data provided in the L3 NetCDF-4 files), and comment on the validation of the L2 NO_2 and H_2O columns with external data sets.

In Section 3, the sampling and gridding approaches and the L3 NO_2 and H_2O column products are described. Topics covered in this section include the computational approaches and the derived statistics associated to each of the L3 monthly mean products.

2. INSTRUMENT AND LEVEL-2 RETRIEVAL OVERVIEW

2.1 The GOME-2 instruments

The GOME-2 sensor (Callies et al., 2000) is the follow up of the Global Monitoring Experiment (GOME), launched in 1995 on ERS-2 (Burrows et al., 1999), and the SCIAMACHY sensor, launched in 2002 on ENVISAT (Bovensmann et al., 1999). GOME-2 is a nadir viewing scanning spectrometer which covers the same spectral range as GOME, i.e. from 240 to 790 nm, with a spectral resolution of about 0.54 nm in the visible spectral region. Additionally, two polarization components are measured with polarization measurement devices (PMDs) using 30 broadband channels covering the full spectral range at higher spatial resolution. The German Aerospace Centre (DLR) plays a major role in the design, implementation and operation of the GOME-2 ground segment for trace gas products, including TCWV, as well as cloud properties in the framework of the EUMETSAT's AC SAF project.

The first GOME-2 instrument was mounted on the MetOp-A satellite (GOME-2A), which follows a sun-synchronous orbit with a mean altitude of 817 km. The overpass local time at the equator is 09:30 Local Time (LT) with a repeat cycle of 29 days. MetOp-A was launched on the 19 October 2006 and GOME-2 TCWV products are available from January 2007 onwards. A second GOME-2 type sensor on board of the MetOp-B satellite (GOME-2B) was launched on the 17 September 2012 and has been fully operational since December 2012. GOME-2 tandem operations started on 15 July 2013. In the tandem mode, GOME-2A operates on a reduced swath width of 960 km, thereby increasing its spatial resolution (40 by 40 km), while GOME-2B continues to operate on a nominal wide swath of 1920 km. This configuration allows the use of the higher spatial resolution data to further study the consistency of the two products in the overlap regions of the GOME-2A and GOME-2B orbits. Finally, the third satellite of the EUMETSAT Polar System series, GOME-2/MetOp-C, is planned to be launched in 2018, guaranteeing the continuous delivery of high-quality NO_2 and H_2O data until 2023.

We can identify important differences between the GOME-type instruments (summarized in Table 2.1), in particular with respect to spatial resolution, default swath width and equator crossing time.

Table 2-1: Summary of the GOME-type instrument characteristics, illustrating the main improvement of GOME-2 compared to its predecessors.

Sensor Satellite	GOME ERS-2	SCIAMACHY ENVISAT	GOME-2 MetOp-A	GOME-2 MetOp-B
Data period	06/1995 - 07/2011	08/2002 - 04/2012	01/2007 - present	12/2012 - present
Spectral coverage	240 – 790 nm	240 – 2380 nm	240 – 790 nm	240 – 790 nm
Ground pixel size	320×40 km^2	60×30 km^2	80×40 km^2 - 40×40 km^2 (*)	80×40 km^2
Swath width	960 km	960 km	1920 km - 960 km (*)	1920 km
Equator crossing time	10 : 30 a.m. LT	10 : 00 a.m. LT	9 : 30 a.m. LT	9 : 30 a.m. LT
Global coverage	3 days (**)	6 days	1.5 days	1.5 days

(*) GOME-2A tandem operation since 15 July 2013. (**) GOME global coverage was lost in June 2003.

2.2 Level 2 retrieval algorithms

As input for the creation of the L3 gridded products, the GOME-2 L2 NO_2 and water vapour columns from the GDP 4.8 are used [ATBD]. A summary of the level 2 retrieval algorithms for NO_2 and water vapour is given below.

2.2.1 Total and tropospheric NO_2 column

The Differential Optical Absorption Spectroscopy (DOAS) method is used to determine NO_2 slant column densities from calibrated GOME-2 (ir)radiance data in the wavelength interval 425–450 nm (Valks et al., 2011). NO_2 absorption features are prominent in this wavelength range, and GOME-2 measurements have high signal-to-noise and manageable interference effects. A single NO_2 cross-section reference spectrum at 240 K (Vandaele et al. 2002) is used, and the interfering species ozone, O_2-O_2 and H_2O are included in the DOAS fit, as well as an additive Fraunhofer Ring spectrum. The NO_2 absorption cross-section has a marked temperature dependence, which has to be taken into account to improve the accuracy of the retrieved columns. In the GDP, an a posteriori correction for the difference between the atmospheric temperature and the 240 K cross-sections temperature is performed on the air mass factor level.

The initial total VCD is computed under the assumption that the troposphere is not polluted. Therefore, the air mass factor is based on stratospheric NO_2 profiles only, whereas the tropospheric NO_2 amount is assumed to be negligible. This approach is valid over large parts of the Earth, but in areas with significant tropospheric NO_2, the total column densities are underestimated and need to be corrected, as described below. The air mass factors are calculated with the LIDORT radiative transfer model for the window mid-point (437.5 nm), since NO_2 is an optically thin absorber in this wavelength region. To incorporate the seasonal and latitudinal variation in stratospheric NO_2 in the air mass factor calculations, a harmonic climatology of stratospheric NO_2 profiles is used (Lambert and Granville, 2004).

Tropospheric NO_2 columns are obtained from the initial total columns by estimating the stratospheric content and removing it from the total amount. Several methods exist for the stratosphere estimation, e.g. see Boersma et al. (2007) and Beirle et al. (2010). In the GDP 4.8, a spatial filtering approach (Wenig et al., 2004) is used by masking potentially polluted areas and then applying a low-pass filter in the zonal direction. This method has been shown to be an improvement on the Pacific reference sector method, which rests on the assumption of a longitudinally homogeneous stratospheric NO_2 layer.

After the stratosphere-troposphere separation, the tropospheric VCD can be determined using a tropospheric air mass factor. For the tropospheric air mass factor computation, monthly average NO_2 profiles from the MOZART-2 CTM are used, determined for the satellite overpass time. GOME-2 derived cloud properties, determined with the OCRA and ROCINN algorithms are used to calculate the air mass factors for scenarios in the presence of clouds (see also Sect. 2.3). The calculation of the tropospheric VCD is complicated in case of (partly) cloudy conditions. For many measurements over cloudy scenes, the cloud-top is well above the NO_2 pollution in the boundary layer, and when the clouds are optical thick, the enhanced tropospheric NO_2 concentrations cannot be detected by GOME-2. Therefore, the tropospheric VCD calculated for observations with a cloud fraction > 20% are flagged in the GOME-2 L2 product.

2.2.2 Total column Water Vapour

Various retrieval methods of the TCWV from space-born spectrometers operating in the visible region have been developed (AMC-DOAS: Noël et al., 1999, Lichtenberg et al., 2010; ERA: Casadio et al., 2000; OCM: Maurellis et al., 2000; IGAM: Lang et al. 2003, 2007; Classical DOAS: Wagner et al. 2003). In contrast to most other methods, the GDP algorithm for the retrieval of water vapour is directly based on a classical DOAS approach, performed in the wavelength interval 614-683 nm, and does not include explicit numerical modelling of the atmospheric radiative transfer. One specific advantage of the DOAS method is that it is only sensitive to differential absorptions, which makes the retrievals less sensitive to instrument changes or instrument degradation. The algorithm consists of three basic steps (described in detail by Wagner et al., 2003, 2006): (1) DOAS fitting, (2) non-linearity absorption correction and (3) Vertical Column Density (VCD) calculation.

In the first step, the spectral DOAS fitting is carried out, taking into account the cross sections of O_2 and O_4, in addition to that of water vapour. To improve the broadband filtering, 3 types of vegetation spectra are included in the fit, together with a synthetic Ring spectrum and, finally, an inverse solar spectrum to correct for possible offsets, e.g. caused by instrumental stray light. In the second step, the water vapour slant column density (SCD) is corrected for the non-linearities arising from the fact that the fine structure water vapour absorption lines are not spectrally resolved by the GOME instrument. In the last step, the water vapour SCD is divided by a "measured" Air Mass Factor (AMF) which is derived from the simultaneously retrieved O_2 and it is defined as the ratio between the measured SCD of O_2 and the known VCD of O_2 for a standard atmosphere. This simple approach has the advantage that it corrects in first order for the effect of varying albedo, aerosol load and cloud cover without the use of additional independent information. It is also important to remark that, in contrast to most other algorithms, the water vapour product from GOME-2 does not rely on additional information, except for the use of an albedo database for the AMF correction. The surface albedo used for the correction is taken from monthly varying albedo maps, which are composite of albedo derived from GOME-1 observations (Koelemeijer et al., 2003) for high latitude (>50°), and SCIAMACHY observations (Grzegorski, 2009) at mid and low latitudes (<40°). For the transition between 40° and 50°, both products are interpolated linearly. This serves the aim to derive a climatologically relevant time series of Total Column Water Vapour (TCWV) measurements (Wagner et al., 2006; Lang et al., 2007; Noël et al., 2008). Currently, the effect of elevated surface terrain is not taken into account in the AMF correction, i.e. over high mountain areas (> 1000 m), the retrieval error in H_2O column is significantly higher.

Compared to GOME/ERS-2 and SCIAMACHY, the observations of GOME-2 have a much wider swath (1920 km scan width). While this broader swath results in a largely improved coverage, also some modification to the H_2O retrieval becomes necessary. In particular, we observe that water vapour total column present a significant Scan Angle Dependency (SAD). There is a bias up to 10 kg/m^2 between the H_2O product for the west and east part of the swath and the central ground pixels. The effect is particularly strong over ocean areas, while the land surface is less affected.

In GDP 4.8 we use the empirical correction for the scan angle dependency introduced in the previous version of the algorithm (Grossi et al., 2015). The correction is based on the GOME-2A full time series and is computed separately over land and ocean surfaces, to take into account the diverse reflectivity properties of the surface. It is computed as follows. Multi-annual monthly mean H_2O total columns are created and employed to select the latitudinal binned regions which contain a sufficient large number of measurements to avoid that the correction is affected by natural variability in the H_2O total columns. Scan angle read-outs toward the nadir scan angle (scan pixel numbers 9-10-11) are then used as reference values to normalize the H_2O total column for every forward angle position and derive a self-consistent correction. Finally, a polynomial is fitted to the normalized measurements in order to remove outliers and obtain a smooth correction function. With this procedure, residuals are of the order of few percent and the bias between the east and west part of the scan is reduced to negligible values.

The water vapour retrieval algorithm uses two cloud indicators to identify and flag cloudy pixels. This is necessary to remove potential systematic effects due to the different altitude profiles of H_2O and O_2 which might still appear in the water vapour product.

The first H_2O cloud flag is set if the retrieved O_2 slant column is below 80% of the maximum O_2 SCD for the respective solar zenith angle (roughly when about 20% from the column to ground is missing). Especially for low and medium high clouds, the relative fraction of the VCD from the ground which is shielded by clouds for O_2 and H_2O can be quite different. Therefore, we require that the main part of the O_2 column is present.

The second cloud flag is set if the product of the GOME-2 cloud fraction and cloud top albedo (see below) exceeds 0.6 (anomalously high cloud top reflection). In this case, the H_2O total column is also set to ”invalid” as the pixel might be considered fully clouded.

2.3 Cloud parameters

The presence of clouds significantly affects the retrieval of tropospheric trace gases in the visible spectral range and it is very important to derive information on cloud properties from the GOME-2 observations to study the indirect effects of residual cloud contamination in the Level 2 product. In most cases the predominant effect of clouds is to shield the absorption below the cloud. However, there can also be an enhancement in the absorption due to multiple scattering events within the cloud. In the NO_2 retrieval the computation of the VCD assumes the independent pixel approximation (IPA) for cloud treatment. For water vapour, especially in the case of low clouds and large cloud fractions, very large errors can occur because of the differential shielding of the O_2 and H_2O profiles and cloudy cases are flagged.

In GDP 4.8, the OCRA and ROCINN algorithms (Loyola et al., 2007) are used for obtaining GOME-2 cloud information. Clouds are regarded as reflecting Lambertian surfaces and cloud information is reduced to the specification of three parameters: cloud fraction, cloud-top albedo and cloud-top pressure. While OCRA provides the radiometric cloud fraction using the broad-band polarization UVN measurements, ROCINN provides effective cloud pressure and cloud albedo from measurements in and adjacent to the oxygen A-band around 760 nm.

The basic idea in OCRA (Optical Cloud Recognition Algorithm (Loyola and Ruppert, 1998; Lutz et al., 2016) is to break down each optical sensor measurement into two components: a cloud-free background and a residual contribution expressing the influence of clouds. The key to the algorithm is the construction of a cloud-free composite that is invariant with respect to atmosphere, to topography and to solar and viewing angles. The effective cloud fraction is determined by examining the separation between the reflectance measured by the PMDs of GOME-2 and their cloud-free composite values. Note that being sensitive to light scattered by clouds, OCRA is also sensitive to scattering by aerosols present in a given GOME-2 scene, so that both effects are subsumed in the retrieved cloud fraction. Moreover, an important upgrade for GOME-2 is the ability to distinguish clouds in measurements affected by ocean surface sun-glint, a phenomenon that is common at the edges of the GOME-2 swath. OCRA discriminates clouds in the region affected by sun-glint by analysing the broad-band polarization measurements (Loyola et al., 2011; Lutz et al., 2016).

The cloud fraction derived from the OCRA algorithm is taken as fixed input to the ROCINN algorithm (Loyola, 2004), which delivers cloud-top height and cloud albedo using measurements inside the O_2 A absorption band. In the simulations, attenuation through oxygen absorption of the direct solar beam and its reflection from ground or cloud-top is considered and surfaces are assumed to be lambertian reflectors. The surface albedo is an important input parameter for the simulations and in ROCINN version 3, the MERIS black-sky albedo climatology is used. In ROCINN v3, the radiative transfer simulations include also Rayleigh scattering and polarization. High-resolution reflectances computed with VLIDORT (Spurr, 2006) are used to create a complete data set of simulated reflectance for all viewing geometries and geophysical scenarios, and for various combinations of cloud fraction, cloud-top height and cloud-top albedo. The inversion is performed using neural network techniques. Finally, the cloud-top pressure for GOME-2 scenes is derived from the cloud-top height provided by ROCINN and an appropriated pressure profile (U.S. Standard Atmosphere).

2.4 Error Estimates and validation

2.4.1 Total and tropospheric NO_2

An estimation of the error budget for the GOME-2 total and tropospheric NO_2 column is provided in Table 2.2. This includes typical errors on NO_2 slant columns and the AMF for the total NO_2 column for unpolluted conditions, and the tropospheric NO_2 column (for polluted conditions). A detailed error analysis and description of the different error sources in the GOME-2 NO_2 product can be found in Valks et al. (2011).

Table 2.2 Estimation of error sources for the total and tropospheric NO_2 column.

Error source	Percent error	
	Total column (unpolluted)	Tropospheric column (polluted)
NO_2 slant column		
NO_2 absorption cross-sections	2-5	2-5
Instrument signal-to-noise	5	5
Instrument spectral stability (wavelength registration)	0.5	0.5
Ring and molecular Ring effect	<2	<2
Stratospheric NO_2 column	n.a.	10-20
NO_2 Air Mass Factor	2-5	15-50
NO_2 vertical column (accuracy)	5-15	40-80

The total and tropospheric NO_2 products (level 2 product retrieved with GDP 4.8) are regularly validated and monitored by comparing to data sets acquired by other satellites (GOME, SCIAMACHY, OMI) and GOME-2 retrievals performed with other processors, as well as ground-based reference measurements. These have been acquired by UV-visible DOAS zenith sky looking spectrometers and zenith-sky looking spectrometers for the NO_2 stratospheric column, and multi axis (MAX-DOAS) spectrometers for the NO_2 tropospheric column (Pinardi et al., 2015).

The main findings from the different comparison studies are summarised below.

The step by-step verification of the GDP 4.8 GOME- 2B product against GOME-2A, and GOME-2A and B alternative retrievals has highlighted a global underestimation by GOME-2B slant columns of 8×10^{14} molec/cm^2 with respect to GOME-2A, which translates into a bias of about 1×10^{14}-3×10^{14} molec/cm^2 (~5-15%) in total and stratospheric columns. GOME-2B tropospheric column data underestimate GOME-2A by less than 5×10^{14} molec/cm^2 (~20%) in moderately polluted conditions, while larger differences (up to 8×10^{14} molec/cm^2) occur in polluted regions. The latter can be explained by the local time difference between GOME-2A and GOME-2B overpasses and the associated impact of the variability in NO_2 content and cloud cover on the comparison results.

The comparison for total NO_2 between GOME-2A/B total NO_2 and correlative ground-based measurements acquired by 25 zenith-sky DOAS instruments show a good agreement, usually within $1–5\times10^{14}$ molec/cm^2 (~10-20%) a value close to the combined uncertainty of the comparison method and within the target accuracies for total NO_2 of $3–5\times10^{14}$ molec/cm^2 in unpolluted conditions [PRD] (GOME-2/NDACC co-located data are filtered to avoid GOME-2 data contaminated by tropospheric pollution, and corrected for photochemical diurnal effects arising from solar local time differences between NDACC and GOME-2 observations).

The GOME-2 tropospheric NO_2 columns have been compared with MAX DOAS measurements at four different BIRA-IASB sites. Very different conditions of tropospheric NO_2 are sampled at the four stations, from clean/remote region (OHP, south of France), city (Uccle, Belgium) and heavily polluted region in Beijing and Xianghe in China (just outside the city, at 60 km south-east of Beijing). Good correlations between GOME-2A and the ground-based MAXDOAS data are obtained, both in terms of correlation coefficients (ranging from 0.65 in Uccle to 0.92 in Beijing) and slopes of the regression analysis, with values ~0.8 at Xianghe, slightly smaller in Uccle and OHP (~0.6) and with larger differences (~0.5) in Beijing. The impact of the location of the ground-based instrument can be seen when comparing the results between Beijing and Xianghe, that are only 60 km apart: in the first case the MAXDOAS is situated in the centre of the Beijing megacity and then it has been moved to the Xianghe site outside the city, in a zone more representative of what is seen by a satellite pixel. This large differences in Beijing compared to results in Xianghe are due to a representation error for stations in urban locations, affected by local pollution episodes, not seen in the averaged GOME-2 pixel. Time series of GOME- 2A and B above OHP in southern France and its comparisons to ground-based MAXDOAS tropospheric NO_2 data, show that the pollution episodes at this site are captured well by both GOME-2 instruments, and the comparisons of monthly averaged columns show consistent seasonal variations, with high NO_2 in winter and low NO_2 in summer.

2.4.2 Total column Water Vapour

The error budget in the H_2O product can be separated in two parts: errors affecting the retrieval of the slant columns (DOAS-related errors), and errors affecting the conversion of the SCD into VCD (AMF-related errors). A preliminary estimation of the error budget for the H_2O slant column (without saturation correction), and for the H_2O total column, is provided in Table 2.3. The slant column error-estimates are mainly based on the DOAS analyses using GOME data (Wagner et al., 2003).

Table 2.3 Estimation of error sources for the total H_2O column.

Error source	Percent error
H_2O slant column	
H_2O absorption cross-sections	< 5
Atmospheric (effective) temperature	3
Other (Instrument signal-to-noise, interference, Ring effect)	< 3
H_2O Air Mass Factor	
Clear Sky	10 – 25
Cloudy*	20 – 100
H_2O vertical column (accuracy)	15 – 100

* For the creation of the L3 product only cloud screened GOME-2 measurements are used, see also Sect. 3.1

AMF-related errors are difficult to quantify, because the water vapour AMF is not based on explicit RT calculations, and there may be compensating effects. For example, in the case of snow surfaces, the high surface reflectivity would lead to a relatively high sensitivity for H_2O in the lower troposphere, and hence a lower AMF-ratio of O_2 to H_2O, but above cold surfaces the tropospheric H_2O column is reduced, causing the opposite effect.

In the GDP algorithm, the following potential error sources are taken into account: relative DOAS fit error of H_2O (Δ_{H2O}) and O_2 (Δ_{O2}), uncertainties in the spectroscopic data (about 10%) and uncertainties due to clouds (Δ_{RTM}). The total relative error can then be derived by the following formula (Wagner et al., 2011):

$$\Delta_{total} = \sqrt{\Delta^2_{H_2O} + \Delta^2_{O_2} + (0.1)^2 + \Delta^2_{RTM}}.$$

The DOAS fit error of the H_2O column typically varies between ~ 0.5 to 5%, mainly depending on humidity, surface albedo and viewing geometry. The uncertainty due to clouds is included empirically in the RTM error, which increases with decreasing O_2 slant column (indicating strong cloud shielding), indicating strong cloud shielding. Clouds may shield a major part of the total H_2O column from the GOME-2 view. This effect is partly compensated for by using the “measured” AMF of O_2, and assuming that the observed O_2 and water vapour absorptions are affected by clouds in a similar way. Nevertheless, since the vertical profile of H_2O is much more peaked in the troposphere with respect to that of O_2 (the H_2O scale height is only about 2 km compared to 8 km for O_2), the measured AMF derived from the O_2, absorption is in general larger than the AMF for water vapour. In particular, in the case of low lying clouds, the shielding effect can differ because the partial columns of water vapour and O_2 below the clouds differ substantially.

Previous validation studies indicated the necessity to reduce the strong cloud influence in the measurements by flagging for cloudy conditions. As discussed in Section 2.2.2, both a criteria based on the measured surface reflection and a threshold based on the of O_2 absorption are used to reduce the errors in the final products (Wagner et al., 2005; Kalakoski et al. 2011, 2014). In particular, the derived water vapour total column depends significantly on the chosen threshold value for O_2 absorption. In most cases, the true H_2O total columns decreases significantly with a lower absorption threshold, as more cloudy measurements are considered. However, during winter and

over the continents the selection of mostly cloud free scenes affects the sampling statistics in the opposite way and can lead to a reverse dependence. It was found that a threshold of 80% for the maximum O_2 SCD is a good compromise in order to exclude the strongest cloudy conditions still preserving many observations.
GOME-2/MetOp-A and MetOp-B water vapour total columns (level 2 product retrieved with GDP 4.8) have been evaluated using comparisons with (1) ground-based observations obtained from the IGRA radiosonde data set and GPS observations from the COSMIC/SuomiNet network (Kalakoski et al., 2014), (2) model data from the ECMWF ERA-Interim reanalysis (Dee et al., 2011a, 2011b) and (3) independent SSMIS F16 satellite measurements from the Remote Sensing System (REMSS, Wentz et al., 1997, 2003). Moreover, an inter-comparison between GOME-2A and GOME-2B data in the period January 2013 to February 2015 has been performed.

The main findings from the different comparison studies are summarised below. All results are derived for cloud free conditions.

From the inter-comparison between GOME-2A and GOME-2B data, overall a very good consistency is found. The GOME-2A water vapour total columns are only slightly drier than the GOME-2B measurements and present a small, negative bias of about -0.37 kg/m^2 (generally less than 1-2%), when averaging all the results for the January 2013 - February 2015 period. In the large majority of cases (> 90% of the time), the relative differences between GOME-2A and GOME-2B H_2O product are within 5 kg/m^2 in absolute value. The results do not change substantially whether we take into account co-located observations or all measurements.

Comparison with soundings and GPS observations show that both GOME-2A and GOME-2B observations are in good agreement with ground-based observations for water vapour amounts below 5 kg/m^2. For very large water vapour columns, both GOME-2 instruments underestimate the ground-based observations. Long-term comparisons show that the product is very stable over validation period (2007-2015). Some seasonal and latitudinal variation was also observed.

Both the GOME-2A and GOME-2B data are in good agreement with the SSMIS measurements. A bias between -0.8 kg/m^2 and 1.3 kg/m^2 for the GOME-2 instruments in the full period January 2007 – April 2013 was found. The correlation coefficient between GOME-2 and SSMIS data sets is close to one (between 0.9 and 0.95). The comparison between GOME-2 and SSMIS data also revealed a seasonal cycle in the geographical distribution of the bias, with a positive bias in June-August and a larger negative bias in December-January. The dominating cause for this pronounced seasonal component is a large positive bias in regions at high latitude, in particular the northern areas of the Atlantic and Pacific Ocean (bias values typically range between 1 and 6.2 kg/m^2). These variations can mainly be related to the impact of clouds on the accuracy of the GOME-2 observations and to the different sampling statistics of the instruments.

Finally, the ECMWF ERA-Interim data set are typically slightly drier than the GOME-2 retrievals. On average GOME-2 data overestimate the ERA-Interim reanalysis by only 3.4 kg/m^2 (GOME-2A) and 0.67 kg/m^2 (GOME-2B). The seasonal behaviour is not as evident when comparing GOME-2 TCWV to the ECMWF ERA-Interim data sets, since the different biases over land and ocean surfaces partly compensate each other. It was identified a very good agreement between GOME-2 total columns and ECMWF model data in February and August 2014, although some discrepancies (bias larger than 5 kg/m^2) over ocean and over land areas with high humidity or a relatively large surface albedo are observed.

Based on these inter-comparisons it was possible to demonstrate that the current GOME-2 H_2O product fulfils the 10% target accuracy in TCWV [PRD] for most conditions, especially for climatologically relevant data.

3. LEVEL 3 DATA

3.1 Theoretical description: sampling and gridding

GOME Level 3 data products are developed with the aim of providing the scientific community with easily translatable data, to both facilitate scientific progress (e.g. on climate trend analysis and low-frequency climate variability) and satisfy public interest.

The creation of GOME-2 L3 data involves the selection of a global, 2- dimensional spatial grid for binning the equivalent L2 data into the grid cell. An "equal angle" squared grid of 0.25º × 0.25º for NO_2 and 0.5º × 0.5º for H_2O was used for translating the spatial information expressed in the GOME-2 ground pixels into a gridded system. Analyses using various different resolutions (i.e. 0.1º × 0.1º, 0.25º × 0.25º and 0.5º × 0.5º) showed that the spatial variations in the water vapour monthly means are properly captured by the larger cell size. For NO_2, a higher 0.25º × 0.25º resolution is needed to capture the larger spatial variations in the tropospheric NO_2 columns.

The binning process for the data set included taking the arithmetic mean and standard deviation of all L2 data points falling into the grid cell in a given month, with possible trimming of low quality measurements due to cloud contamination. Only forward-scan pixels were used for the gridding. Of the forward-scan TCWV measurements available, about 20% have the second cloud flag (product of cloud fraction and cloud-top albedo > 0.6) set and 50% have the first cloud flag (< 80% O_2 minimum SCD) set. This leaves about 40% of the GOME-2 L2 data set for the creation of the L3 product. The cloud screening of the tropospheric NO_2 columns (cloud fraction < 20%) also leaves about 40% of the L2 data.

Because of the relatively large GOME-2 pixel size, a significant grid effect would be induced by assigning each GOME-2 measurement to a single grid point based on the center coordinates of the GOME-2 ground pixel, without taking into account the pixel geometry and extension. Therefore, an area weighted tessellation has been used by sub-gridding each (0.25º×0.25º / 0.5º×0.5º) cell into sub-longitude and sub-latitude cells and then computing a weighted mean and standard deviation for each of the L3 variables. The area weight is assigned to each grid cell based on the number of it's sub-cells that fall into the satellite ground pixel measurement and it depends on the corner coordinates of the satellite ground pixel only (not on the viewing geometry). Because of the large dimension of the data set, the weighted mean and variance are calculated in a single pass, using a stable online algorithm (Knuth, 1998). To compute the variance we update the sum of squares of differences from the current mean (below denoted as $M_{2,n}$), as:

$$s_n^2 = \frac{M_{2,n}}{n-1}, \quad M_{2,n} = M_{2,n-1} + (x_n - \bar{x}_{n-1}) * (x_n - \bar{x}_n) \qquad (1)$$

where:

$$\bar{x}_n = \bar{x}_{n-1} + \frac{(x_n - \bar{x}_{n-1})}{n} \qquad (2)$$

is the sample mean of the first n measurements. This algorithm is much less prone to catastrophic cancellation and can be extended to handle the area weights by replacing the simple counter "n" in Eq. 1 and 2 with the sum of weights in the grid cell (West, 1979).

3.2 Level 3 mapped NO_2 parameters

The monthly averaged GOME-2 total and tropospheric NO_2 column for each grid cell is given in the PRODUCT group. The temporal resolution of the product is monthly (calendar). Grid maps coordinates range from -180.0° to +180.0° in longitude and from -90.0° to 90.0° in latitude. Both the mean errors associated to the L2 retrieval and the standard deviations associated to the variation of the individual (gridded) measurements in each grid-cell are computed. An area weighted tessellation technique is used to compute the NO_2 columns and the associated standard deviations, as described in Sect. 3.1.

Table 3-1: List of Level 3 mapped NO_2 products.

Variable name	**Unit**	**Size**	**Description**
Latitude	Degree north	720 / 0.25°	Centre of the gridded data Latitudinal resolution
Longitude	Degree east	1440 / 0.25°	Centre of the gridded data Longitudinal resolution
group: PRODUCT			
NO2total	molec cm^{-2}	720 x 1440	averaged total NO_2 column
NO2total_err	molec cm^{-2}	720 x 1440	averaged error associated to the total NO_2 column
NO2total_stddev	molec cm^{-2}	720 x 1440	total NO_2 column standard deviation
NO2trop	molec cm^{-2}	720 x 1440	averaged tropospheric NO_2 column
NO2trop_err	molec cm^{-2}	720 x 1440	averaged error associated to the tropospheric NO_2 column retrieval
NO2trop_stddev	molec cm^{-2}	720 x 1440	standard deviation associated to the tropospheric NO_2 column grid cells
nobs	-	720 x 1440	number of individual observations in the grid cell
group: SUPPORT_DATA			
group: DETAILED_RESULTS			
CLOUD_PARAMETERS		Table 3.3	
SURFACE_PROPERTIES		Table 3.4	

As an example, the monthly gridded tropospheric NO_2 column (parameter *NO2trop* in PRODUCT group) derived from GOME-2/MetOp-A measurements using area weighted tessellation is shown in Fig. 3.1 for February 2008. Clearly visible are the high tropospheric NO_2 concentrations above large urban and industrial areas in Asia, Europe and America.

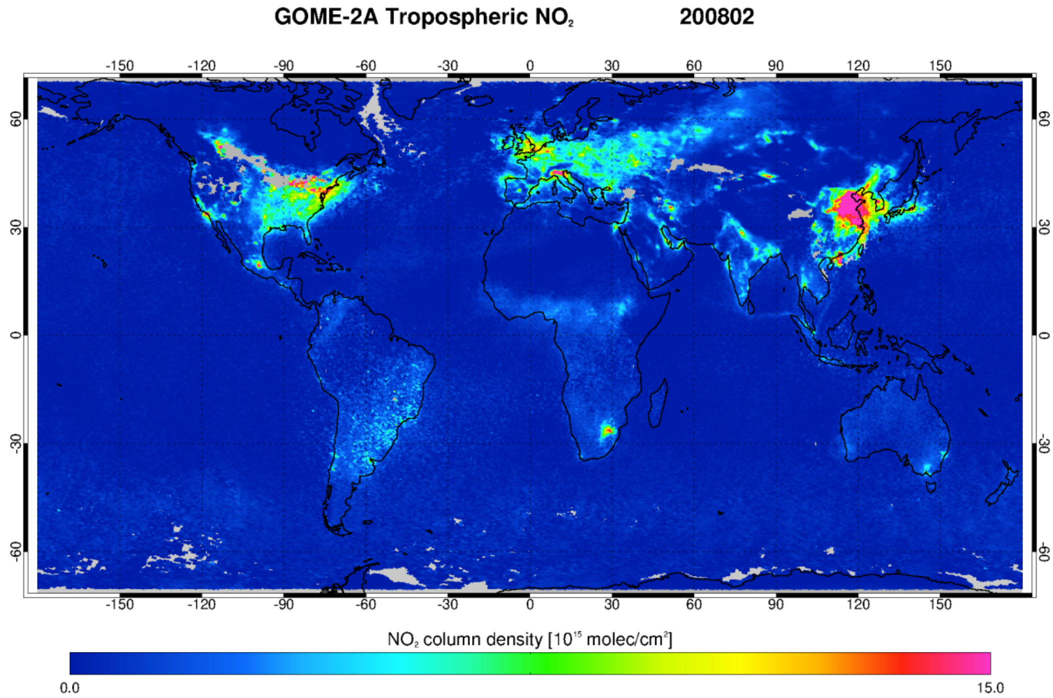

Figure 3.1: Geographical distribution of the averaged tropospheric NO_2 column derived from GOME-2/MetOp-A measurements in February 2008. The tropospheric NO_2 column measurements have been aggregated by area weighted tessellation. Note that tropospheric NO_2 columns are only retrieved for the latitude range 70°N-70ºS, see [ATBD].

3.3 Level 3 mapped H_2O parameters

The monthly averaged GOME-2 TCWV for each grid cell is given in the PRODUCT group. The temporal resolution of the product is monthly (calendar). Grid maps coordinates range from -180.0° to +180.0° in longitude and from -90.0° to 90.0° in latitude. Both the mean errors associated to the L2 measurements and the standard deviations associated to the observed values are computed. An area weighted tessellation technique is used to compute the TCWV and the associated standard deviation, as described in Sect. 3.1

Table 3-2: List of Level 3 mapped H_2O products

Variable name	Unit	Size	Description
Latitude	Degree north	360 / 0.5°	Centre of the gridded data Latitudinal resolution
Longitude	Degree east	720 / 0.5°	Centre of the gridded data Longitudinal resolution
group: PRODUCT			
tcwv	Kg m^{-2}	360 x 720	averaged total column water vapour
tcwv_err	Kg m^{-2}	360 x 720	averaged error associated to the total column water vapour retrieval
tcwv_stddev	Kg m^{-2}	360 x 720	standard deviation associated to the total column water vapour grid cells
nobs	-	360 x 720	number of individual observations in the grid cell
group: SUPPORT_DATA			
group: DETAILED_RESULTS			
CLOUD_PARAMETERS		Table 3.3	
SURFACE_PROPERTIES		Table 3.4	

Figure 3.2 shows an example of the monthly gridded total column water vapour (parameter *tcwv* in PRODUCT group) derived from GOME-2/MetOp-A measurements in April 2013. In contrast to other satellite H_2O data sets, the GOME-2 product has the advantage that it covers the entire Earth, including both ocean and continents, leading to a more consistent picture of the global distribution of the atmospheric humidity. Moreover, the Level-2 data retrieval is performed in the visible/near-infrared spectral range and it is very sensitive to water vapour in the lower troposphere, which contributes the major fraction of the total atmospheric column. The spatial patterns in the humidity distribution show high values in the tropics and low humidity at higher latitudes. The global averaged total column water vapour (i.e. averaged over all latitude and longitude cells) is 19.03 kg/m^2.

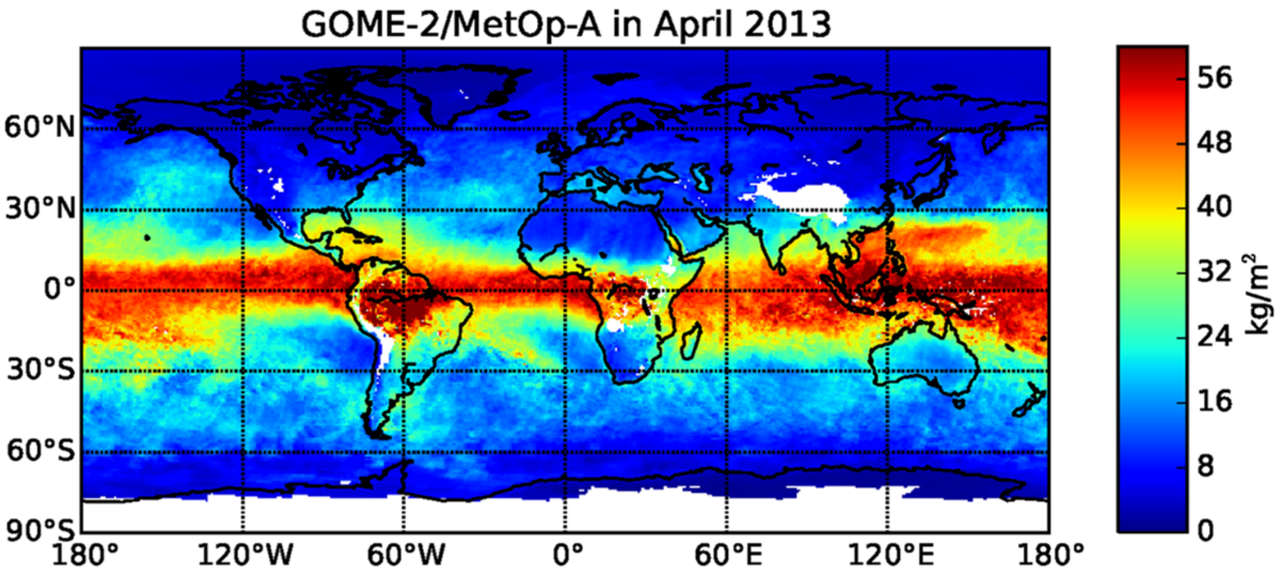

Figure 3.2: Geographical distribution of the averaged H_2O vertical columns derived from GOME-2/MetOp-A measurements in April 2013. The total column water vapour measurements have been aggregated by area weighted tessellation.

In Fig. 3.3, the geographical distribution of the number of observations retrieved from GOME-2A in April 2013 is shown. The average number of measurements in each grid cell is 10.6, while the median number is 4.9. Due to the polar orbit of MetOp, the number of observations is generally largest in the polar region. In addition, the number of observations depends strongly on the cloud filtering, and there are some regions without any measurements due to persistent clouds or high surface elevation. At mid latitudes, there are about 8 observations per grid cell. At higher latitudes the number of available observations increases drastically, and we observe a maximum of ~90 observations per grid cell.

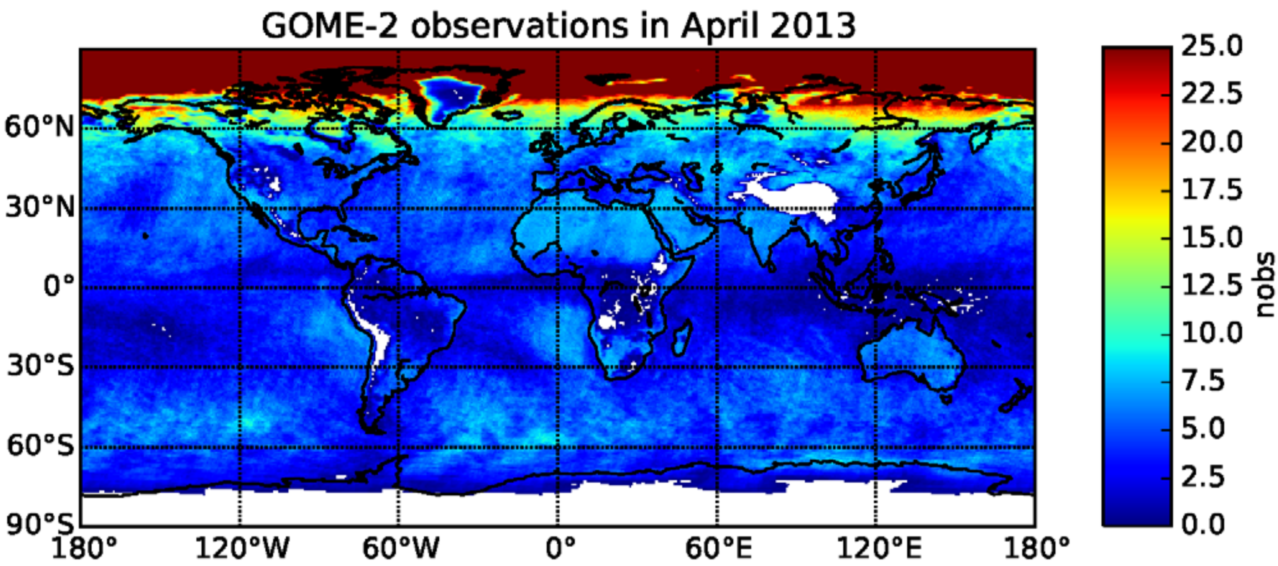

Figure 3.3: Geographical distribution of the number of H_2O vertical columns observations (for each grid cell) from GOME-2/MetOp-A in April 2013.

Fig. 3.4 and 3.5 illustrate the geographical distribution of the averaged error (*tcwv_err*) and standard deviation (*tcwv_stddev*) of the total column water vapour retrieved from GOME-2A in April 2013. Larger errors in the water vapour column are found over regions with high surface albedo and residual cloud contamination. The two figures clearly illustrate that for the low- and mid-latitudes, the (natural) variation in the water vapour columns is usually much larger than the estimated error.

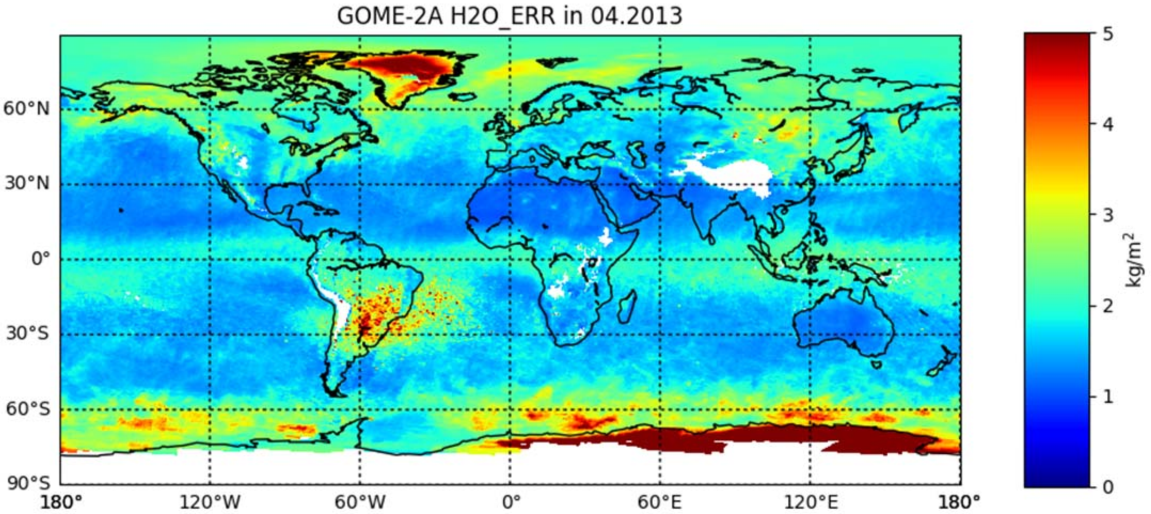

Figure 3.4: Geographical distribution of the average error in the H_2O vertical columns (for each grid cell) from GOME-2/MetOp-A in April 2013.

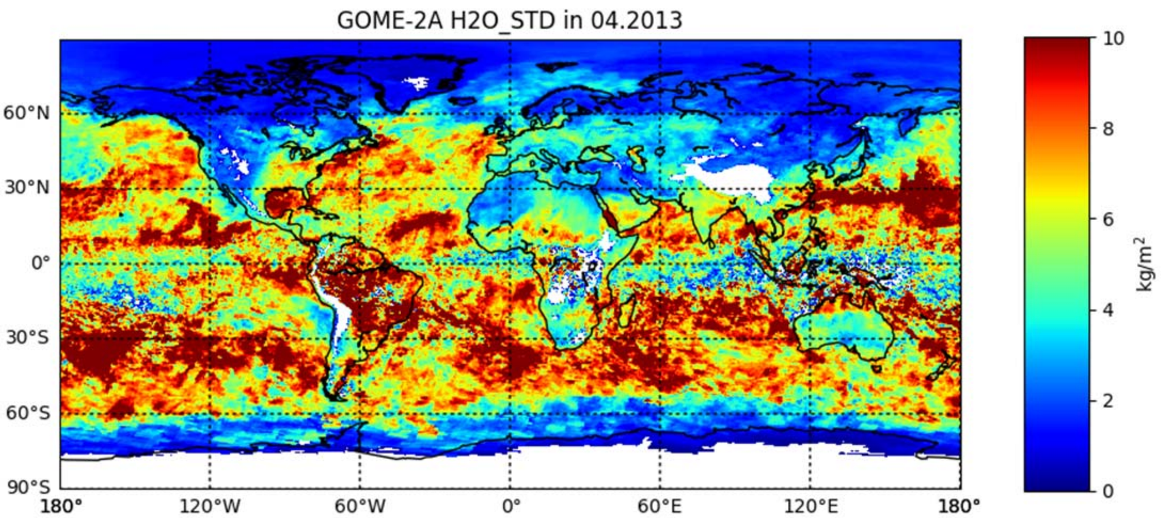

Figure 3.5: Geographical distribution of the standard deviation in the H_2O vertical columns (for each grid cell) from GOME-2/MetOp-A in April 2013.

3.4 Cloud properties

The L3 mapped products are computed from GOME-2 NO_2 and water vapour total columns which are not flagged as cloud contaminated in the L2 data. Only tropospheric NO_2 observations with cloud fraction < 20% are used (see also Sect. 2.2.1), while for water vapour we consider only the observations for which the product between cloud fraction and cloud-top albedo is smaller then 0.6 and the majority (80%) of the O_2 slant column is present (see also Sect. 2.2.2).

The effects of a residual cloud contamination on the L3 NO_2 and water vapour data can be investigated by looking at the averaged cloud parameters (cloud fraction, cloud height and cloud albedo) associated to the cloud screened NO_2 and water vapour measurements in a given grid cell (see Table 3.3). The standard deviations for each of these parameters are also given to appreciate how much variation there is from the average. The number of individual observation in each grid cell (*nobs*) is provided in the PRODUCT group.

Table 3-3: Averaged Level 3 cloud parameters.

Variable name	Unit	Size	Description
cloud_fraction	-	360 x 720 / 720 x 1440	average cloud fraction
cloud_fraction_std	-	360 x 720 / 720 x 1440	cloud fraction standard deviation
cloud_height	km	360 x 720 / 720 x 1440	average cloud height
cloud_height_std	km	360 x 720 / 720 x 1440	cloud height standard deviation
cloud_albedo	-	360 x 720 / 720 x 1440	average cloud top albedo
cloud_albedo_std	-	360 x 720 / 720 x 1440	cloud top albedo standard deviation

3.5 Surface properties

The surface height and albedo for each grid cell, and a surface condition flag is provided in the SURFACE_PROPERTIES group. The surface properties from the L2 data are averaged to the resolution on the L3 product. The number of individual observation in each grid cell (*nobs*) is provided in the PRODUCT group.

Table 3-4: Averaged Level 3 surface properties.

Variable name	Unit	Size	Description
surface_albedo	-	360 x 720 / 720 x 1440	average surface albedo
surface_height	km	360 x 720 / 720 x 1440	average surface height
surface_flag	-	360 x 720 / 720 x 1440	land sea flag (see Table 3.5)

3.5.1 Surface condition flag

The *surface_flag* gives additional information about the surface condition. The surface data are distinguished between land, sea and coast. The product is based on the Land/Sea flag of the different pixels of the GOME-2 Level-2 product. Also PMD subpixels affected by sun glints are considered to derive the surface flag. To distinguish smaller land masses (e.g. Hawaii or Indonesia) in the resolution of the L3 product (0.25º×0.25º / 0.5º×0.5º), a “coastal” value was introduced.

Table 3-5: Land / sea flag

Flag value	Description	Threshold
0	land	if less than 20 % of all GOME-2 level 2 observations were classified as sea
1	coast	between 20% and 80% were classified as sea
2	sea	more than 80% were classified as sea

REFERENCES

Applicable documents

ATBD Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, NO2 , BrO, HCHO, SO2, H2O and Cloud Properties, DLR/GOME-2/ATBD/01, Rev. 3/A, Valks, P. et al., 2016.

PUM Product User Manual for GOME-2 Total Columns of Ozone, NO2 , BrO, HCHO, SO2, H2O and Cloud Properties, DLR/GOME-2/PUM/01, Rev. 3/A, Valks, P., et. al., 2016.

PRD O3M SAF Product Requirements Document, SAF/O3M/FMI/RQ/PRD/001/Rev. 1.7, J. Hovila, et. al., 2015.

Reference documents

Beirle, S., Kühl, S., Pukite, J., and Wagner, T.: *Retrieval of tropospheric column densities of* NO_2 *from combined SCIAMACHY nadir/limb measurements*, Atmos. Meas. Tech., 3, 283–299, doi:10.5194/amt-3-283-2010, 2010.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: *Near-real time retrieval of tropospheric NO2 from OMI*, Atmos. Chem. Phys., 7, 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerik, J., No¨el, S., Rozanov, V. V., Chance, K. V., and Goede, A.: *SCIAMACHY – mission objectives and measurement modes*, J. Atmos. Sci., **56**(2), 127–150, 1999.

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., de Beek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., and Perner, D.: *The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results*, J. Atmos. Sci.,**56**, pp 151–175, 1999.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – MetOp's Second Generation Sensor for Operational Ozone Monitoring, ESA Bulletin, No. 102, 2000.

Casadio, S., Zehner, C., Piscane, G., and Putz, E.: Empirical Retrieval of Atmospheric Airmass factor (ERA) for the Measurement of Water Vapor Vertical Content using GOME Data, Geophys. Res. Lett. 27, 1483–1486, 2000.

Dee, D. and Coauthors: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. Soc., 137, 553-597, 2011a.

Dee, D. and Coauthors: The use of reanalysis data for monitoring the state of the climate [in "The State of the Climate in 2010"]. Bull. Amer. Meteor. Soc. 92(6), S34-S35, 2011b.

GCOS, *Systematic observation requirements for satellite-based products for climate*. Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC". GCOS-107 (WMO/TD No. 1338) September, 2006.

GCOS, *Systematic observation requirements for satellite-based data products for climate, 2011 Update*. Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update)". GCOS-154, December, 2011.

Grzegorski, M.: *Cloud retrieval from UV/VIS satellite instruments (SCIAMACHY and GOME)*, PhD thesis, University of Heidelberg, 2009.

Grossi, M., Valks, P., Loyola, D., Aberle, B., Slijkhuis, S., Wagner, T., Beirle, S., and Lang, R.: *Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B*, Atmos. Meas. Tech., 8, 1111-1133, doi:10.5194/amt-8-1111-2015, 2015.

Kalakoski, N., T. Wagner, K. Mies, S. Beirle, S.Slijkhuis, D. Loyola: *O3M SAF Validation Report, Offline Total Water Vapour*, SAF/O3M/FMI/VR/H2O/111, 2011.

Kalakoski , N., Kujanpää, J,. Sofieva, V,. Tamminen, J., Grossi, M., and Valks, P.: *Comparison of GOME-2/MetOp total column water vapour with ground-based and in-situ measurements*, amt-2014-292, 2014.

Knuth, D. E., The Art of Computer Programming, volume 2: *Seminumerical Algorithms*, 3rd edn., p. 232. Boston: Addison-Wesley, 1998.

Koelemeijer, R. B. A., Haan, J. F. D., and Stammes, P.: *A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations*, J. Geophys. Res.,108(D2), 4070 , doi:10.1029/2002JH002429, 1650, 1651, 2003.

Lambert, J.-C., and J. Granville, *Harmonic climatology of stratospheric* NO_2, BIRA-IASB, Brussels, 2004.

Lang, R., Williams, J. E., van der Zande, W. J., and Maurellis, A. N.: *Application of the Spectral Structure Parameterization technique: retrieval of total water vapour columns from GOME*, Atmos. Chem. Phys., 3, 145–160, doi:10.5194/acp-3-145-2003, 2003.

Lang, R., Casadio S., Maurellis, A.N., and Lawrence M.G.: *Evaluation of the GOME Water Vapor Climatology 1995-2002*, J. Geophys. Res. **112**, D12110, doi:10.1029/2006JD008246, 2007.

Lichtenberg, G., Bovensmann, H., Van Roozendael, M., Doicu, A., Eichmann, K.-U., Hess, M., Hrechanyy, S., Kokhanovsky, A., Lerot, C., Noel, S., Richter, A., Rozanov, A., Schreier, F. and Tilstra, L.G.: *SCIAMACHY Offline Level 1b-2 Processor ATBD* (ENV-ATB-QWG-SCIA-0085, issue 1A), 1–137, 2010.

Loyola, D., and T. Ruppert, *A new PMD cloud-recognition algorithm for GOM*E, ESA Earth Observation Quarterly, 58, 45-47, 1998.

Loyola, D., *Automatic Cloud Analysis from Polar-Orbiting Satellites using Neural Network and Data Fusion Techniques*, IEEE International Geo science and Remote Sensing Symposium, 4, 2530-2534, Alaska, 2004.

Loyola D., Thomas W., Livschitz Y., Ruppert T., Albert P., Hollmann. R.: *Cloud properties derived from GOME/ERS-2 backscatter data for trace gas retrieval*, IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 9, pp. 2747-2758, 2007.

Loyola D., Koukouli M. E., Valks P., Balis D. S., Hao N., Van Roozendael M., Spurr R. J. D., Zimmer W., Kiemle S., Lerot C., Lambert J.-C: *The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation*, J. Geophys. Res., 116, D07302, 2011.

Lutz, R., Loyola, D., Gimeno García, S., and Romahn, F.: *OCRA radiometric cloud fractions for GOME-2 on MetOp-A/B*, Atmos. Meas. Tech., 9, 2357-2379, doi:10.5194/amt-9-2357-2016, 2016.

Maurellis, A. N., Lang, R., Van der Zande, W. J., Ubachs, W, and Aben, I: *Precipitable Water Column Retrieval from GOME*, Geophys. Res. Lett., 27, 903–906, 2000.

Noël, S., Buchwitz, M., Bovensmann, H., Hoogen, R., and Burrows, J. P.: *Atmospheric Water Vapor Amounts Retrieved from GOME Satellite Data*, Geophys. Res. Lett. 26, 1841 pp., 1999.

Noël, S., Mieruch S., Bovensmann, H., and Burrows, J. P.: *Preliminary results of GOME-2 water vapour retrieval and first applications in polar regions*. Atmos. Chem. Phys. **8**, pp 1519-1529, 2008.

Shindell, D., G. Faluvegi, D. Koch, G. Schmidt, N. Unger, S. Bauer, *Improved Attribution of Climate Forcing to Emissions*, Science, **326**, 716, DOI: 10.1126/science.1174760, 2009.

Pinardi, G., Lambert, J.-C., Y. Huan, I. De Smedt, J. Granville, M. Van Roozendael and P. Valks (2015), *O3M SAF Validation report of GOME-2 GDP 4.8 NO_2 column data*, SAF/O3M/IASB/VR/NO2, Issue 1/0, Oct. 2015.

Spurr, R. J. D., *VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media*, JQSRT, 102(2), 316-342, doi:10.1016/j/jqsrt.2006.05.005, 2006

Valks P., Pinardi G., Richter A., Lambert J.-C., Hao N., Loyola D., Van Roozendael M., Emmadi S. (2011), Operational total and tropospheric NO_2 column retrieval for GOME-2, Atmospheric Measurement Techniques, vol. 4, pp. 1491-1514.

Wagner, T., Heland, J., Zoeger, M., and Platt, U.: *A fast H_2O total column density product from GOME – Validation with in-situ aircraft measurements*, Atmos. Chem. Phys., **3**, 651–663, 2003.

Wagner, T., Beirle, S., Grzegorski, M., Sanghavi, S., and Platt, U.: *El-Nino induced anomalies in global data sets of total column precipitable water and cloud cover derived from GOME on ERS*-2, J. Geophys. Res., 110, doi:10.1029/2005JD005972,1642, 2005.

Wagner, T., Beirle, S., Grzegorski, M., and Platt, U.: *Global trends (1996–2003) of total column precipitable water observed by Global Ozone Monitoring Experiment (GOME) on ERS-2 and their relation to near-surface temperature*, J. Geophys. Res. **111**, D12102, doi:10.1029/2005JD006523, 2006.

Wagner, T., Beirle, S., and Mies, C.: *Description of the MPI-Mainz H2O retrieval (Version 5.0, March 2011), technical document*, http://www.sciamachy.org/products/H2O/H2Ovc_IUP_AD.pdf (last access: 3 March 2015), 2011.

Wenig, M., Kuhl, S., Beirle, S., Bucsela, E., Jahne, B., Platt, U., Gleason, J., and Wagner, T.: *Retrieval and analysis of stratospheric NO_2 from the Global Ozone Monitoring Experiment*, J. Geophys. Res., 109, D04315, doi:10.1029/2003JD003652, 2004.

Wentz, F.J.: *A well-calibrated ocean algorithm for SSM/I*, J.Geophys.Res. 102, No. C4, pp. 8703-8718, 1997.

Wentz, F. J.: *SSM/I Version-7 Calibration Report*, report number 011012, Remote Sensing Systems, Santa Rosa, CA, 46pp, 2013.

West, D. H. D.: *Updating Mean and Variance Estimates: An Improved Method*, Communications of the ACM, 22, 9, 532-5, 1979.