Water Vapour Climate Change Initiative (WV_cci) - CCI+ Phase 1

ATBD Part 1 - MERIS-MODIS-OLCI L2 Products

Ref: D2.2

Date: 21 January 2021

Issue: 2.1

For: ESA / ECSAT

Ref: CCIWV.REP.004

This Page is Intentionally Blank

Project   :   Water Vapour Climate Change Initiative (WV_cci) - CCI+ Phase 1

Document Title: ATBD Part 1 - MERIS-MODIS-OLCI L2 Products

Reference   :   D2.2

Issued   :   21 January 2021

Issue :   2.1

Client: ESA / ECSAT

Author(s)   :   Jürgen Fischer and Rene Preusker (Spectral Earth), Marc Schröder (DWD), and Olaf Danne (Brockmann Consult)

Copyright   :   Water_Vapour_cci Consortium and ESA

Document Change Log

Issue/ Revision	Date	Comment
1.0	28 March 2019	Initial version; Spectral Earth
1.1	03 April 2019	Updated version for submission to ESA after Team Review and format update
2.0	10 November 2020	Transformation of measurements, references to HOAPS included, description of aggregation and combination included
2.1	21 January 2021	RID responses implemented

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

1.   INTRODUCTION

1.1 Purpose

This document provides information about the physical background, technical structure and the functional principle of the Total Column Water Vapour (TCWV) retrieval as defined within ESA's Water Vapour CCI project, The project aims to the development of a sensor-comprehensive retrieval of TCWV above land surfaces using the $\rho\sigma\tau$ H2O absorption band, measured by the different satellite instruments MERIS, MODIS and OLCI. The results from this retrieval will be gridded and comprise CDR-1. The same product is combined with the EUMETSAT CM SAF HOAPS product over ocean, with sea-ice and coastal areas filled with NIR observations, resulting in CDR-2. The HOAPS product is based on L2 retrievals applied to microwave imager observations. The final data records (CDR-1 and CDR-2) cover the period July 2002 to December 2017.

1.2 Structure of the document

This document addresses the Algorithm Theoretical Baseline (ATBD) and the Input/Output Data Definition (IODD). It is structured as following:

- Overview of the retrieval algorithm scheme for MERIS, MODIS and OLCI (Section 2) with subsections on:

    a. Description of the forward operator (Section 2.3)

    b. Input / output structure (Section 2.4)

- Brief section on the retrieval algorithm scheme for microwave imager observations (Section 3).

- Aggregation and combination (Section 4)

- Assumptions and limitations (Section 5)

1.3 Relevance of water vapour measurements and their status

Water vapour is a key climate variable. In the lower troposphere, condensation of water vapour into precipitation provides latent heating, which dominates the structure of tropospheric diabatic heating (Trenberth and Stepaniak, 2003a,b). Water vapour is also the most important gaseous source of infrared opacity in the atmosphere, accounting for about 60% of the natural greenhouse effect for clear skies (Kiehl and Trenberth, 1997), and provides the largest positive feedback in model projections of climate change (Held and Soden, 2000).

In the stratosphere, there are potentially important radiative impacts due to anthropogenic sources of water vapour, such as from methane oxidation. In the troposphere, the radiative forcing due to direct anthropogenic sources of water vapour (mainly from irrigation) is negligible. Rather, it is the response of tropospheric water vapour to warming itself – the water vapour feedback – that matters for climate change. In GCMs, water vapour alone provides the largest positive radiative feedback, it roughly doubles the warming in response to forcing (such as from greenhouse gas increases). There are also possible stratospheric water vapour feedback effects due to tropical tropopause temperature changes and/or changes in deep convection.

The radiative effect of absorption by water vapour is roughly proportional to the logarithm of its concentration, so it is the fractional change in water vapour concentration, not the absolute change that governs its strength as a feedback mechanism. Randall et al. (2007) stated in the IPCC Fourth Assessment Report that calculations with GCMs suggest that water vapour remains at an approximately constant fraction of its saturated value (close to unchanged relative humidity (RH)) under global-scale warming. Under such a response, for uniform warming, the largest fractional change in water vapour, and thus the largest contribution to the feedback, occurs in the upper troposphere. In addition, GCMs find enhanced warming in the tropical upper troposphere, due to changes in the lapse rate. This further enhances moisture changes in this region, but also introduces a partially offsetting radiative response from the temperature increase, and the net effect of the combined water vapour/lapse rate feedback is to amplify the warming in response to forcing by around 50%. The close link between these processes means that water vapour and lapse rate feedbacks are commonly considered together. The strength of the combined feedback is found to be robust across GCMs, despite significant inter-model differences, for example, in the mean climatology of water vapour.

For NWP models the water vapour information is assimilated directly from the satellite radiances (infrared and microwave) in channels affected by water vapour absorption. This is achieved by 4D variational analysis utilising a fast-radiative transfer model as the observation operator. Currently infrared radiances from IASI, AIRS, HIRS, SEVIRI, GOES-Imager and MTSAT-Imager and microwave radiances from AMSU-A, AMSU-B, MHS, SSM/I, SSMI(S) and AMSR-E are all used in global NWP centres to complement the radiosonde measurements network. GPS total zenith delay is also used to define the total column water vapour at GPS sites and the OLCI total column water vapour is beginning to be used over land during the day. The impact of the water vapour measurements improves the water vapour analyses, but it has been harder to demonstrate improvements to the forecasts by improving the water vapour field especially for low level water vapour analysis changes. This is because the models are

often tuned to their own water vapour climatology and modifying the low-level water vapour can often lead to changes in the model spin up which have detrimental effects on the forecasts.

The global network of radiosonde measurements provides the longest record of water vapour measurements in the atmosphere, dating back to the mid-1940s. However, early radiosonde sensors suffered from significant measurement biases, particularly for the upper troposphere, and changes in instrumentation with time often lead to artificial discontinuities in the data record (e.g., see Elliott et al., 2002). Consequently, most of the analysis of radiosonde humidity has focused on trends for altitudes below 500 hPa and is restricted to those stations and periods for which stable instrumentation and reliable moisture soundings are available.

Satellite observations provide near-global coverage and thus represent an important source of information over the oceans, and over land where radiosonde observations are scarce, and in the upper troposphere and stratosphere, where radiosonde sensors are unable to measure or at best unreliable.

An area of concern for infrared based products is that they cannot infer the tropospheric water vapour in cloudy areas whereas the microwave products can. Various studies (e.g. Mieruch et al., 2010 and Xavier et al., 2010) have shown this can lead to a dry bias for the infrared products. All satellite measurements of lower tropospheric water vapour over land are challenging but developments using products based on visible and near infrared radiance are showing new inside of the heterogeneous spatial water vapour distribution above land surface.

2. WATER VAPOUR ALGORITHM FOR MERIS, MODIS AND OLCI

2.1 Overview

The retrieval estimates total column water vapour (TCWV) over cloud-free land surfaces, coasts and sea-ice. The algorithm is based on differential absorption using satellite measurements within the $\rho\sigma\tau$-H_2O absorption band between 890 and 1000 nm (Fischer, 1988; Gao et al., 1990; Bartsch et al., 1996; Albert et al., 2001; Albert et al., 2005). The procedure can be adapted easily for different sensors, which measure reflected sunlight in this spectral region. It requires measurements in at least one absorption band and one nearby band with only little or no absorption (reference bands). It has been initially designed for the Medium Resolution Imaging Spectrometer MERIS (Fischer, 1988; Lindstrot et al., 2012) and has been recently adapted for the Moderate Resolution Imaging Spectroradiometer (MODIS) (Diedrich et al., 2015) and the Ocean- and Land Colour Instrument (OLCI).

The backbone of the retrieval is the forward operator that simulates satellite measurements. A 1D-Var scheme optimizes the difference between simulated and measured radiances band by iteratively varying the TCWV value and additional state parameter, following the scheme after Rodgers (2000) (see section 2.3.4.1).

For the forward operator, radiative transfer simulations were performed for various atmospheric conditions and stored in look up tables (LUTs). The needed gaseous absorption coefficients, based on HITRAN 2016 (Gordon et al., 2017) data plus the continuum absorption parametrisation by Mlawer et al. (2012), were calculated using an advanced k-distribution routine for different standard temperature profiles (Bennartz and Fischer, 2000; Doppler et al., 2013, Preusker and Fischer, pers. 2018). Interactions with aerosol particles and molecules are simulated with the radiative transfer model MOMO (Hollstein and Fischer, 2012). The land surface is assumed to be Lambertian.

The standard forward operator is an n-dimensional linear interpolation in the pre-calculated look up tables. However, the forward model is exchangeable, a full online Radiative Transfer Solver (RTS) as well a different LUT parametrisations can be used.

Depending on the surface type, the algorithm uses different sets of state parameters and fixed, additionally assumed parameter. The retrieved state consists of the surface albedo in the window bands and the TCWV. The auxiliary surface temperature and pressure constrain the choice of the assumed temperature profile. The aerosol optical thickness constrains the atmospheric state further.

Currently ERA-5 is used to provide the priors for MODIS and MERIS. OLCI L1b data already comes with ECMWF analysis (non-time critical) or forecast (near real-time).

Uncertainty estimates are provided by considering all relevant sources of error such as sensor noise, and errors of forward modelling parameters such as aerosol optical thickness, aerosol vertical distribution, surface elevation and temperature.

Eventually, the algorithm relies on a strict prior cloud detection. Undetected clouds can lead to underestimation of TCWV, due to shielding effects as well as to overestimation of TCWV due to multiple scattering effects.

2.2   Satellite instruments

The measurements of the satellite sensors MERIS, MODIS and OLCI used in this study are briefly described in the following.

2.2.1   MERIS and OLCI

The key mission objective for the Sentinel-3 OLCI instrument is the continuity of the ENVISAT MERIS instrument capability. The primary mission of OLCI is the observation of the spectral distribution of upwelling radiance just above the sea surface (the water-leaving radiance) that is then used to estimate several geophysical parameters through the application of specific bio-optical algorithms. Atmospheric correction for ocean colour data is challenging (International Ocean Colour Coordinating Group - IOCCG, 2010) as only about 4% of the radiation measured by a satellite instrument originates from the water surface and sensors require high signal to noise ratio (SNR), particularly for the 'blue' bands (Donlon et al, 2012). This requires an accurate retrieval and description of the atmospheric state with respect to scattering and absorption processes. This points to the secondary objective, the detection of atmospheric properties, which include cloud detection (pixel classification) and aerosol detection, which is important not only for atmospheric correction but also for the monitoring of air-pollution.

The S-3 OLCI instrument is based on the opto-mechanical and imaging design of ENVISAT MERIS (see Figure 2-1). The instrument is a quasi-autonomous, self-contained, visible push-broom imaging spectrometer and incorporates the following significant improvements when compared to MERIS:

- An increase in the number of spectral bands (from 15 to 21)

- Improved SNR

- Improved long-term radiometric stability

11

- Mitigation of sun-glint contamination by tilting cameras in westerly direction by 12.6°

- Complete coverage over both land and ocean at 300 m Full-Resolution (FR)

- Improved instrument characterization including stray-light, camera overlap, and calibration diffusers.

The cameras are arranged to slightly overlap with each other to cover a wide 68.5° across-track field of view as shown. The OLCI swath is not centred at nadir (as in the MERIS design) but is tilted 12.6° westwards to mitigate the negative impact of sun-glint contamination that affects almost half of the MERIS ocean colour observations at sub-tropical latitudes.
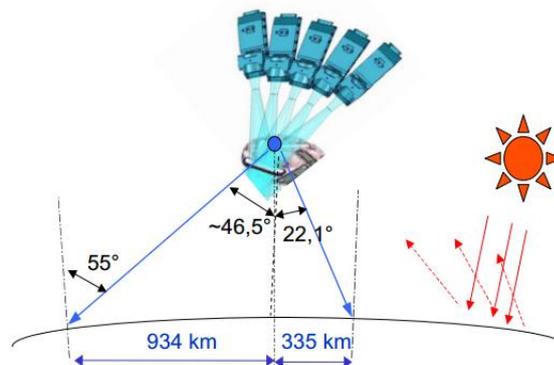
Figure 2-1: OLCI features a tilted field of view to avoid sun-glint.

OLCI bands are optimised to measure ocean colour over the open ocean and coastal zones. A new channel at 1.02 µm has been included to improve atmospheric and aerosol correction capabilities, additional channels in the O_2 A-band spectral region are included for improved cloud top pressure (height) and water vapour retrieval, and a channel at 673 nm has been added for improved chlorophyll fluorescence measurements. In principle, the OLCI programmable acquisition design allows spectral bands to be redefined in both location and width during commissioning of the instrument after which time they will be fixed for the mission duration.

For the water vapour retrieval, the MERIS channels Me14 and Me15 are used as reference and absorption band, respectively. For OLCI there are three reference channels, Oa17, Oa18 and Oa21, and two absorption channels Oa19 and Oa20 (see Table 2-1).

12

Table 2-1: The spectral bands of OLCI and MERIS

OLCI Band	Centre (nm)	Width (nm)	MERIS Band	Centre (nm)	Width (nm)
Oa1	400	15			
Oa2	412,5	10	Me1	412,5	10
Oa3	442,5	10	Me2	442,5	10
Oa4	490	10	Me3	490	10
Oa5	510	10	Me4	510	10
Oa6	560	10	Me5	560	10
Oa7	620	10	Me6	620	50
Oa8	665	10	Me7	670	15
Oa9	673,75	7,5			
Oa10	681,25	7,5	Me8	681,25	10
Oa11	708,75	10	Me9	708,75	10
Oa12	753,75	7,5	Me10	753,75	10
Oa13	761,25	2,5	Me11	761,875	3,75
Oa14	764,375	3,75			
Oa15	767,5	2,5			
Oa16	778,75	15	Me12	778,75	15
Oa17	865	20	Me13	865	10
Oa18	885	10	Me14	885	10
Oa19	900	10	Me15	900	10
Oa20	940	20			
Oa21	1020	40			

2.2.2 MODIS

The MODIS instruments provide high radiometric sensitivity (12 bit) in 36 spectral bands ranging in wavelength from 0.4 µm to 14.4 µm. The responses are custom tailored to the individual needs of the user community and provide exceptionally low out-of-band response. Two bands are imaged at a nominal resolution of 250 m at nadir, with five bands at 500 m, and the remaining 29 bands at 1 km. A ±55-degree scanning pattern at a sun-synchronous polar orbit of 705 km achieves a 2,330-km swath and provides global coverage every one to two days. The MODIS instrument is flying on board the Terra (10:30 equator crossing time) and on board Aqua (13:30 equator crossing time). The latter is also part of the A-train.

The wide swath of the MODIS instrument stretches over 2330 km with a spatial resolution of 250 m for the first two bands, Md1 and Md2. Md3-Md7 features 500 m, while the remaining ones feature 1000 m (see Table 2-2).

For the water vapour retrieval, the MODIS channels Md2 and Md5 (averaged over 4 b4 pixels) are used as reference and the three channels Md17 to Md19 for the water vapour absorption.

Table 2-2: The spectral bands of MODIS within the solar spectral domain, the spatial resolution is 250 m for Md1-Md2 (dark grey), 500m for Md3-Md7 (light grey), 1000m for Md8-Md19, and MODIS water vapour sensitive bands Md16-Md19 (dotted gray)

MODIS Band	Centre (nm)	Width (nm)
Md1	645	50
Md2	858,5	35
Md3	469	20
Md4	555	20
Md5	1240	20
Md6	1640	24
Md7	2130	50
Md8	412,5	15
Md9	443	10
Md10	488	10
Md11	531	10
Md12	551	10
Md13	667	10
Md14	678	10
Md15	748	10
Md16	869,5	15
Md17	905	30
Md18	936	10
Md19	940	50

2.3   Algorithm description

2.3.1   Theoretical description

Water vapour has various absorption features in the solar and terrestrial spectrum, which is due to a combination of rotational and the three fundamental vibration modes of the water molecule. Measurements of reflected sunlight in this absorption bands enable a determination of TCWV if the following conditions are given:

1. Solar radiation is available, limiting the retrieval to daytime measurements.

2. The absorption bands are in the sensitive part of the spectrum but not saturated (like 1.3 µm MODIS band).

3. The surface brightness in the absorption band can be estimated.

4. The lower troposphere, holding the main part of the TCWV, is not masked by clouds or optical thick aerosol layers.

The NIR spectral range is perfectly suited for daytime, cloud-free retrieval of TCWV over land. Here, almost all land surfaces provide a good background. The retrieval is based on the differential absorption technique (Fischer, 1988; Gao et al., 1993; Lindstrot et al., 2012). The basic principle of the method is the comparison of the measured radiance in an absorption band to a close by band with no or few absorption features.

For monochromatic radiation, neglecting scattering processes along the photon path, the transmittance T through the atmosphere can be related to its optical depth τ and the air mass along the geometrical path (amf=1/$cos(\theta_s)$ + 1/$cos(\theta_v)$), following the Beer-Lambert law:

$$-ln(T_{noscat}) = \tau * amf \qquad\qquad (1)$$

The optical depth of a medium is a measure of the amount of the absorption of atmospheric gases, here the water vapour. But the absorption measured by a medium resolution satellite instrument cannot be quantified by the simple Lambert Beer law. Instead, the superposition of many individual absorption lines must be considered, whereby their depths and widths result from pressure- and temperature dependent broadening processes. Consequently, the knowledge of the actual temperature-pressure profile is necessary in order to simulate the correct atmospheric transmittance. However, Lindstrot et al. (2012) and Lindstrot and Preusker (2012) showed, that it is sufficient to use the surface pressure and the 2m-temperature to approximate the transmittance corresponding to the actual temperature profile by adequately mixing the pre-calculated transmittance values corresponding to two close standard profiles.

In a previous investigation, we found a systematic overestimation of TCWV, estimated from MODIS NIR measurements, compared to ground-based microwave radiometer measurements (Diedrich et al., 2015). We identified several possible reasons for this wet bias:

- deficits in spectral and radiometric characterisation of the MODIS bands,

- deficits in NIR forward simulations and
- a systematic underestimation of the microwave radiometer-based retrievals.

The latter was excluded, since the wet bias was consistently found when compared to ground-based GPS measurement and radiosonde observations. Diedrich et al. (2015) further excluded spectral calibration reasons by analysing synthetic sensitivity tests.

These findings are consistent with TCWV retrievals from MERIS and OLCI. First, there is a systematic overestimation of about 10 percent. Second, OLCI's standard TCWV retrieval does not show distinct camera dependent features, although the central wavelength of the bands jumps by more than 1.2 nm between some of the 5 cameras.

The origin of the deficiency of the retrieval might be caused by the estimation of the water vapor absorption, in particular in the insufficient description of the continuum absorption and broadening effects.

In the following we describe essential aspects of the generation of the look-up tables (LUT), core of the TCWV retrieval scheme. The LUTs for all used satellite sensors have been built by the radiative transfer model MOMO (Hollstein and Fischer, 2012; Doppler et al., 2013).

Although there have been slight modifications of the used HITRAN line parameters during the last 2 decades, we always found a wet bias when the different HITRAN database versions are used for the development of the retrieval scheme. But the introduction of the water vapour foreign- and self-continuum absorption in the $\rho\sigma\tau$-absorption band leads to significant enhanced absorption and reduced transmission for a given TCWV value. These continua have two origins, the H_2O dimer interaction and their collisions with other gases (Shine et al., 2016). The self-continuum does not exist in all parts of the water vapour spectrum but is very important in the thermal infrared and later found also in the $\rho\sigma\tau$-H_2O absorption band (Ptashnik et al, 2011). The foreign- and self-continua do not have the same dependence to water vapour amount, pressure and temperature. The used values of the continua are computed with the help of tabulated coefficients, which can be found in the MT-CKD database (Clough et at., 2005; Mlawer et al., 2012) and in CAVIAR (Continuum Absorption at Visible and Infrared wavelengths and its Atmospheric Relevance, Ptashnik et al., 2011).

The computation of the appropriate line profiles and the implementation of the continua in the computation of the atmospheric absorption spectra is challenging (Doppler et al., 2014). Spectroscopy codes (e.g. CGASA - Coefficient of Gas

Absorption, Doppler et al., 2013; LBLRTM - Line by Line Radiative Transfer Model, Clough et al., 2005) provide state of the art estimations of gas absorption coefficients at high spectral resolution. But still, it has to be recognized, that the continua coefficients are derived by fitting line-by-line absorption calculations and observations – a solid theoretical model to calculate foreign- and self-continuum absorption coefficients is still missing (Mlawer et al., 2012). However, CGASA as well as LBLRTM assume a cut-off wavelength at 64 times the half-width of an absorption line, which justify the use of the published foreign- and self-continuum absorption coefficients.

Preusker and Fischer (2019) calculated the transmission of the different absorption processes due to pure line-by-line absorption T_{lbl} and due to line-by-line plus foreign- and self-continuum absorption $T_{lbl+cont}$ for the spectral region of OLCI's bands Oa19 and Oa20 (see Figure 2-2). The difference $1-(T_{lbl}-T_{lbl+cont})$ (green lines) quantifies the effect of the continuum absorption on the spectral transmission. Since, line-by-line and continuum absorption have to be considered as coherent processes, the continuum absorption does not contribute to additional absorption in saturated line centers but is most effective along the unsaturated line wings.

The TCWV algorithm, as applied in the WV_cci, is based on line-by-line calculations using the HITRAN 2016 database plus the continuum absorption parametrisation by Mlawer et al. (2012). Simulated total atmospheric transmittance due to water vapour in the NIR and the relative response function of the five relevant MODIS bands (blue curves), MERIS bands (green) and OLCI bands (green + red) are plotted in Figure 2-3.
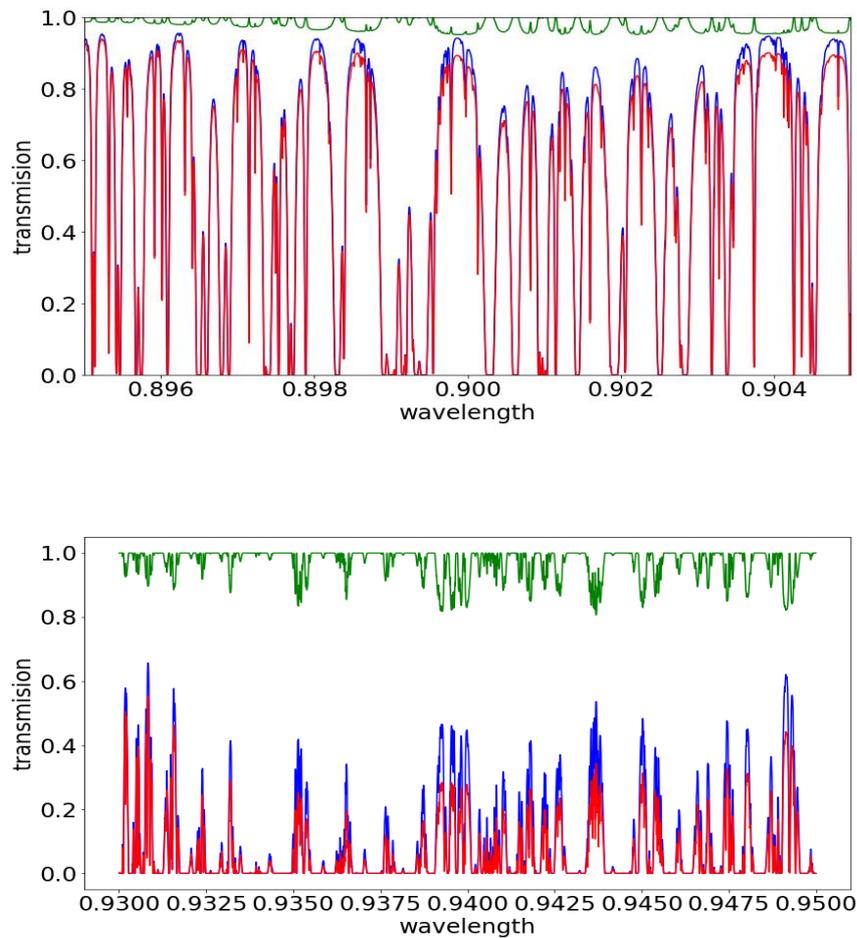
**Figure 2-2: Transmission versus wavelength [in μm] for a mid-latitude summer atmosphere (TCWV=30Kg/m^2) and an air-mass factor of 4, pure line–by-line absorption T$_{lbl}$ (blue), line-by-line plus foreign- and self-continuum absorption T$_{lbl+cont}$ (red), and 1-(T$_{lbl}$-T$_{lbl+cont}$) (green) for the spectral range of Oa19 (upper) and of Oa20 (lower) [Preusker and Fischer, 2019].**
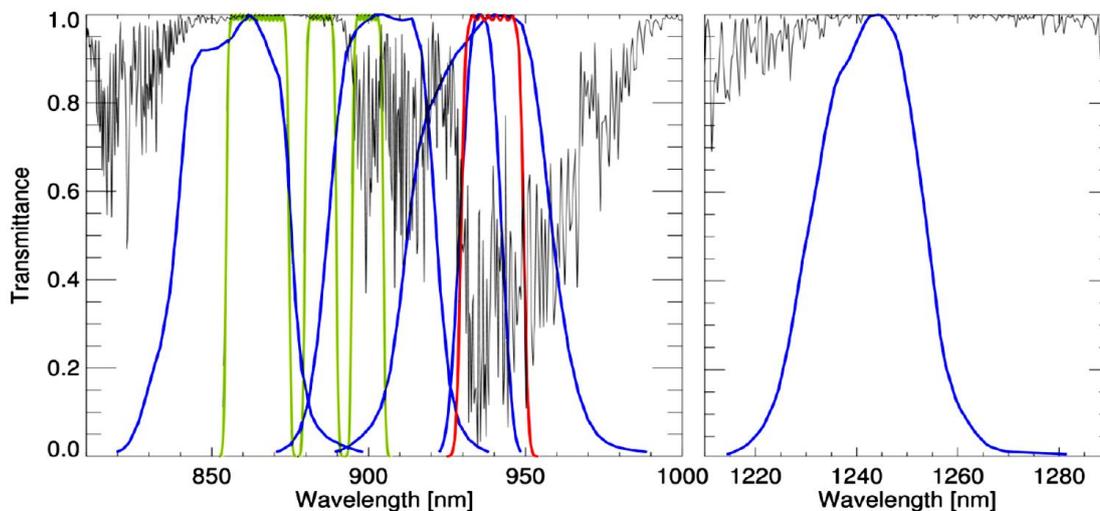
Figure 2-3: Simulated total atmospheric transmittance due to water vapour in the NIR (absorption coefficients from HITRAN database, Rothman et al., 2013, black curve) and relative response function of the relevant MODIS bands (blue curves), of MERIS bands (green) and OLCI bands (green + red).

The calculated transmission functions are further used to estimate band integrated pseudo spectral absorption coefficients by an uncorrelated k-binning approach (Doppler et al., 2013). These k-bins are used in the radiative transfer code MOMO (Hollstein and Fischer, 2012) to simulate spectral radiances of the different satellite instruments.

Furthermore, instrument effects, such as over- or under-corrected spectral stray light, cannot be ruled out either. Although a state-of-the-art line by line transmission calculation including H_2O continua absorption has been performed to generate the retrieval LUTs, a transmission correction has to be applied in order to adjust the atmospheric transmittance.

$$\tau_{corr}^{abs} = a + b * \tau^{abs}$$

(2)

τ^{abs} is the optical thickness of water vapor absorption, based on HITRAN calculations, τ_{corr}^{abs} is the corrected optical thickness. The coefficients *a* and *b* were obtained by minimizing the difference between the simulated and measured MERIS, MODIS and OLCI observations using the ARM-MWR data set as a reference. It is important to note, that this was done for each band 19 and 20 individually. The estimated and applied correction coefficients for MERIS, MODIS and OLCI are given in Table 2-3.

There is no specific procedure to harmonize the retrieved TCWV products of the different satellite, however, there is an implicit harmonisation due to the introduction of

a correction of the calculated optical thickness by the ARM-MWR TCWV retrievals for MERIS, MODIS and OLCI.

The impact of aerosols is considered by a continental aerosol with variable optical depth between τ^{cont} =0 and τ^{cont}=0.7.

Table 2-3: Correction coefficients (see equation 2) for the adjustment of the absorption optical thickness of water vapour for OLCI band 19 and 20, MERIS band 15 as well as MODIS bands 17,18 and 19

Instrument/Band	a	b
OLCI ba19	00054	1.061
OLCI ba20	0.023	1.147
MERIS bn15	0.0006	1.045
MODIS_bn17	-0.003	0.98
MODIS_bn18	-0.0005	1.066
MODIS_bn19	0.0116	1.058

The ocean surface is treated as a black ocean with a sea surface roughness which is parametrized after Cox and Munk (1954), considering a wind speed taken from ECWMF. No ocean sediment load is considered, which might cause problems above bright ocean surfaces. Lakes are treated as oceans, so far they are large enough and recognized by the land/sea/lake mask. For all surfaces, land and water, the same maximum of a solar zenith distance of 73.4° is considered, which is equivalent to an air-mass factor of 3.5.

2.3.2 Cloud screening

As cloud contamination has a significant impact on TCWV retrievals from the method described above, a strict filtering of clouds is required for the measurements from all NIR sensors considered (MERIS, OLCI, MODIS).

2.3.2.1   MERIS/OLCI

The applied cloud screening schemes for both MERIS and OLCI are quite similar. They are both based on a neural network approach as implemented in the IdePix Pixel Classification processor developed by Brockmann Consult as plugin for SNAP (Sentinels Application Platform). As input for the neural nets, the square roots of the TOA reflectances (obtained from an internal radiance-to-reflectance conversion) in all available bands are used. The final pixel classification provides land, water, cloud,

snow/ice and (over water) sun glint flags. More details are given in the integrated help documentation of the SNAP toolbox. See http://step.esa.int/main/toolboxes/snap/ for details on SNAP.

The distinction of clouds and snow/ice is known as a challenging issue. Therefore, an improved scheme available for OLCI for the detection of snow/ice has been added. This scheme was recently developed in the frame of the ESA SEOM S3 Snow Project. It is described in detail in Brockmann et al. (in preparation).

In practice, the IdePix processor is run on the MERIS/OLCI L1b TOA radiance input products to obtain the corresponding cloud mask for the TCWV retrieval.

In the TCWV retrieval, snow/ice pixels, including sea ice, were initially treated in the same way as a 'water' pixel. However, during the analysis and the validation of the first version of the full TCWV dataset it turned out that the treatment of sea ice as 'land' pixel (using the 'land' algorithm with corresponding lookup tables) would better reflect the physical conditions. Therefore, the land algorithm was applied for OLCI in the generation of the final TCWV dataset. For MERIS, a full reprocessing starting from L2 would be out of scope for the remaining project frame, thus the data obtained with the water algorithm are kept here. In case of MERIS, the differences for sea ice pixels if treated as land or water do not seem to be very large anyway. However, in a future project phase a reprocessing with the land algorithm should be foreseen also for MERIS, ideally in combination with an improved scheme for distinction of clouds and snow/ice, similar to the current one for OLCI.

2.3.2.2   MODIS/Terra

For a given MODIS (Terra) MOD021KM L1b radiances input product, the corresponding MODIS MOD35_L2 Cloud product, which is available on the same grid, is used for cloud masking.

The MODIS Cloud Mask product is a Level-2 product generated at 1-km and 250-m (at nadir) spatial resolutions. The cloud detection algorithm employs a series of visible and infrared threshold and consistency tests to specify confidence that an unobstructed view of the Earth's surface is observed. The 250-m cloud-mask flags are based on the visible channel data only. Radiometrically accurate radiances are required, so holes in the Cloud Mask will appear wherever the input radiances are incomplete or of poor quality.

In addition to the cloud flag, the MODIS Cloud Mask product contains other processing path flags inherited from upstream processing steps. Among these are a day/night flag and a surface type flag (land, water, snow/ice, desert etc.), All these flags are used in the WV_cci processing chain for MODIS TCWV L2 for pixel classification (clouds, land, water, snow/ice), and for the exclusion of MODIS L1b night mode products. To apply a cloud filtering as strict as possible, only pixels finally identified as 'certainly clear' are considered for TCWV computation. Again, snow/ice pixels, including sea ice, are treated in the same was as a 'land' pixel in the retrieval.

The underlying MODIS cloud detection algorithm together with the 48-bit encoding of the flags with a large number of specific cloud tests is described in great detail in the MODIS Cloud Mask ATBD (MODIS Cloud Mask Team, 2010).

2.3.3 Universal forward operators

The operator has been implemented as an n-dimensional linear interpolation in pre-calculated sensor specific look up tables, having the state vector elements as dimensions. However, the full algorithm is modular, and the forward operator could be replaced by any RTM like module. The look up tables are organized as follows:

- ocean:
  - dimension names and sequence: *wvc, aot, wsp, tmp, wvl, azi, vie, suz, bands*
  - *wvc*, water vapor content: [0.1, 0.5, 5, 20, 40, 75] kg/m^2
  - *aot*, aerosol optical depth at 900 nm: [0., 0.04, 0.08, 0.16, 0.63, 0.95]
  - *wsp*, wind speed at 10 m: [2, 3.3, 4.6, 5.9, 7.2, 8.5, 9.8, 11.1, 12.4, 13.7, 15] m/s
  - *wvl*, central wavelength of band 19: [898.3, 899.27, 900.24] nm. It is assumed, that for a specific pixel, all bands have the same offset with respect to their mean position. The mean positions of each band are given in the /cha/[18,19,20]/cwvl attribute in the same file. Having a central wavelength of 899.27 nm for band 19, the central wavelength dimension belongs to a shift of -0.97nm, 0.0 nm and +0.97nm.
  - *tmp*, air temperature at 2 m: [258.6, 272.2, 287.2, 288.2, 294.2, 299.7] K
  - *suz* and *vie*, sun and viewing zenith angle: [0, 9.8, 18.9, 28.0, 37.1, 46.1, 55.2, 64.3, 73.4] °
  - *azi*, azimuth difference angle: [0, 18, 36, 54, 72, 90, 108, 126, 144, 162, 180]°

- land:
  - dimension names and sequences: *wvc, al0, al1, aot, prs, tmp, azi, vie, suz, bands*:
  - *wvc*, water vapor content: [0.1, 0.5, 5, 20, 40, 75] kg/m^2
  - *al0* and *al1*, surface albedo in window band 0 and 1: [0.001, 0.01, 0.1, 0.3, 1]
  - *aot*, aerosol optical depth at 900 nm: [0, 0.05, 0.1, 0.2, 0.7]
  - *prs*, surface pressure: [1030, 780, 530] hPa
  - *tmp,* air temperature at 2 m: [263.13, 288.13, 313.13] K
  - *suz* and *vie*, sun and viewing zenith angle: [0, 9.8, 18.9, 28.0, 37.1, 46.1, 55.2, 64.3, 73.4] °
  - *azi*, azimuth difference angle: [0, 18, 36, 54, 72, 90, 108, 126, 144, 162, 180]°

Note, that the surface pressure and the water vapor column are non-linearly, scaled and transformed (with natural logarithm and square root, respectively) in order to increase the linearity of the problem and decrease the necessary sampling size of the look up table (see also section 2.3.4.2).

2.3.4 Retrieval scheme

This retrieval estimates TCWV values with the help of an inverse modelling scheme. Deviations between modelled and measured radiances as well as between the estimated state and the prior knowledge about the state are iteratively reduced, using a variational method. There is a quality-filtering, defined by the cost-function < 1 (see equation 3) and a maximum of six iteration. The uncertainties are estimated by taking into account all error influences.

2.3.4.1 Inversion technique

The estimation of a state vector X based on a measurement Y and a priori knowledge X_a is considered as optimal, if the following cost function $J(X)$ is minimized:

$$J(X) = \frac{1}{2}(Y - F(X))^T S_E^{-1}(Y - F(X)) + \frac{1}{2}(X_a - X)^T S_a^{-1}(X_a - X), \tag{3}$$

with S_E and S_a being the measurement and a priory error co-variance matrices and $F(X)$ a forward model calculating a measurement Y from a state X. To incorporate forward model uncertainties into S_E, it is composed from the error co-variance matrix of the measurement S_M (from e.g. instrument signal noise ratio at signal level) and from the forward model parameter error co-variance S_B. K_B is the *Jacobian* of the forward model F with respect to its parameterisations B:

$$S_E = S_M + K_B^T S_B K_B \tag{4}$$

This approach assumes Gaussian probability density functions of the priors and measurements, and further bias free measurements, priors and forward operators. The state X, that minimizes equation (3) is referred to as the maximum a posteriori (MAP) solution. It is the same as the expected solution. For non-Gaussian statistics the expected value and the maximum a posteriori solution will differ, and thus the found solution won't be optimal. However, it is reasonable to assume, that the measurement and prior uncertainties follow a normal distribution. We do not know if the forward model

errors are Gaussian, but they are small, as long as the clear-sky pre-condition holds. In any case, eventually the retrieved state as well as its uncertainty will be validated. If they agree, the assumptions made are very likely correct.

There are several methods to find the optimal solution of equation (3). If the dependencies between state vector *X* and measurement *Y* are not high-frequency (ideally the cost is convex) or if a first guess estimation X_{fg} in the convex vicinity of the optimal solution exists (thus a solver does not trap into local minima), gradient based methods can be used, such as: *Steepest Descent*, *Gauss-Newton*, *Levenberg-Marquardt* and *Newton* method. The commonality among them is, that starting from a first guess they iteratively approximate the optimal solution using gradient (the *Jacobian*) or (approximations of) the second order derivatives (the *Hessian*) for the update step (*us*).

$$X_{i+1} = X_i - us_i. \tag{5}$$

A *Gauss-Newton* update step is used:

$$us_i = (S_a^{-1} + K_i^T \cdot S_E^{-1} \cdot K_i)^{-1} \cdot [K_i^T \cdot S_E^{-1} \cdot (F(X_i) - Y) - S_a^{-1} \cdot (X_a - X_i)] \tag{6}$$

Additionally, to numerical criteria (maximum number of iterations is reached or the update step is smaller than machine accuracy) to stop the iteration, the following criteria as is used (Rodgers (2000)), which is based on the step width relative to the retrieval error co-variance $\widehat{S}_i$:

$$(X_i - X_{i+1})^T \cdot \hat{S}_i^{-1} \cdot (X_i - X_{i+1}) \le n \cdot \epsilon. \tag{7}$$

n is the number parameter in the state vector and ϵ is adjustable, e.g. to 0.01.

In addition to the cost function, 1D-Var automatically provides a number of linear uncertainty measures:

- The retrieval error co-variance:

$$\hat{S} = (S_a^{-1} + K_i^T \cdot S_E^{-1} \cdot K_i)^{-1} \tag{8}$$

- The averaging kernel (the sensitivity of the retrieved state $\hat{x}$ to the truth):

$$A = \frac{\delta \hat{x}}{\delta x} = \frac{\delta \hat{x}}{\delta y} \cdot \frac{\delta y}{\delta x} = G \cdot K \tag{9}$$

using the gain *G*:

$$G = \hat{S} K^T S_e^{-1} \tag{10}$$

- The trace of A, giving the degrees of freedom *dof* :

$$dof = tr\, A \qquad (11)$$

- The retrieval noise:

$$S_n = GS_eG^T$$

- The smoothing error:

$$S_S = (I - A)S_a(I - A)^T \qquad (12)$$

The first guess of TCWV f_g is taken from ECMWF analysis (for OLCI) or re-analysis (MERIS and MODIS).

2.3.4.2   Transformation of measurement

A common and useful treatment of an inverse problem to increase convergence speed is to transform it to a 'more linear' form. For the universal forward operator, we use a transformed measurement τ_i^r in the absorption bands. The transformed measurement almost linearly depend on the total amount of water vapour and thus allows an inversion with few iterations.

$$\tau_i^r = -\ln\left(\rho_i/\tilde{\rho}_i\right) \cdot \frac{1}{\sqrt{amf}} = (\ln(\tilde{\rho}_i) - \ln(\rho_i)) \cdot \frac{1}{\sqrt{amf}} \qquad (13)$$

ρ_i is the top of atmosphere normalized radiance in the absorption band (OLCI: 19,20; MODIS: 18,19; MERIS: 15), amf is the air mass factor, $\tilde{\rho}_i$ is top of atmosphere normalized radiance   of the window bands (ρ_0, ρ_1) inter- or extrapolated to the wavelength λ_i of the absorption band:

$$\tilde{\rho}_i = \rho_0 + \frac{\rho_1 - \rho_0}{\lambda_1 - \lambda_0} \cdot (\lambda_i - \lambda_0) \qquad (14)$$

With that in mind, τ_i^r is a kind of rectified optical thickness, which is roughly proportional to the amount of water vapour.

2.3.4.3 Uncertainty estimates

After the iteration procedure, the retrieval uncertainty is calculated using equation (8), considering the following sources of uncertainty:

- measurement uncertainty quantified by signal noise ratio (SNR)
- uncertainty of the aerosol optical depth (assumed to be 0.1)
- uncertainty of the surface pressure and –temperature (assumed to be 5 hPa and 5 K, respectively).
- uncertainty of the wind speed (5 m/s)

Since the measurement (top of atmosphere reflectance) is transformed before injected into the inversion, the corresponding measurement error co-variance must be transformed too. Assuming, that the signal noise ratio (SNR) is approximately the same for all bands and that a simple uncertainty propagation of equation (13) is valid, the uncertainty in the transformed absorption bands (the rectified optical thicknesses) $\sigma^2_{\tau i}$ is:

$$\sigma^2_{\tau i} = \left[\frac{\sigma^2_i}{\rho^2_i} + \frac{\tilde{\sigma}^2_i}{\tilde{\rho}^2_i} \right] \cdot \frac{1}{amf}$$

$$= \left[\frac{1}{SNR^2} + \left(\frac{1}{SNR^2} + \sigma^2_{inter} \right) \right] \cdot \frac{1}{amf} \tag{15}$$

Here we assume an uncertainty σ^2_{inter} (of e.g. 0.01) due to the extra/ interpolation of the window bands to the absorption bands (equation 14). $\sigma^2_{\tau i}$ are the diagonal elements of the measurement error co-variance S_M.

There is a number of potentially significant sources of uncertainties that are not included in the current TCWV retrieval, these are due to

- undetected clouds, which depend on the applied cloud mask
- high aerosol load, such as high load of desert dust
- non-black sea-surface, such as above ocean regions with high sediment load
- uncertainties in the estimation of the spectral albedo slope
- uncertainties in the spectral characterisation of all cameras, which is variable in time and over the field of view.

2.4   Input Output data

This section defines the input data, required for the processing of MERIS, OLCI and MODIS data as well as the output of the water vapour processor.
The TCWV processor is using normalised radiances. The observation geometry is expressed in viewing zenith angle, Sun zenith angle and azimuth difference angle, which are all given in the L1b product-files (see Table 2-4).

Table 2-4: Satellite measurements taken from the Level 1b instrument data files

Quantity	Unit	Valid range	Source	Comment
Normalized radiance	1/sr	0 -1	L1b	Depending on instrument different bands serve as absorption and as window bands (see Section 1.3). Normalized radiance is defined as top of atmosphere radiance divided by the in-band solar irradiance.
Viewing zenith angle	deg	0-60	L1b	
Sun zenith angle	deg	0-75	L1b	
Azimuth difference angle	deg	0-180	L1b	

The auxiliary data used in the TCWV processor are taken for MERIS and OLCI from ESA's meteorology data-files provided by ECWMF, and additional accessible data on aerosols, i.e. from ESA/NASA archive. For MODIS the ERA-interim data are used (see Table 2-5).

Table 2-5: Auxiliary data used in the TCWV processor and its valid range

Quantity	Unit	Valid range	Source	Comment
Surface pressure	hPa	1050-200	QNH from ECMWF, adapted to actual surface height	GFS and other model outputs are also possible. The accuracy requirements for the surface pressure are low (15hPa), compared to the accuracy of models
2-m temperature	K	260-330	ECMWF	
Total column water vapour	kg/m2	0 75	ECMWF	*A priori* and first guess
Aerosol optical thickness	1	0-1.0	Taken from climatology	If not available, a value of 0.1 is used
Aerosol optical thickness uncertainty	1	0.1	Taken from climatology	If not available, a value of 0.2 is used

The output of the TCWV processor is the total column water vapour and its uncertainty in kg/m^2. For further analysis of the convergence of the inversion process, the estimated transmission and the number of required iterations are given (see Table 2-6).

Table 2-6: Output of the TCWV algorithm after applying to the Level 1b instrument data files; TCWV and its estimated uncertainty, convergence as a logical expression, intermediate results to check the inversion process

Quantity	Unit	Valid range	Source	Comment
TCWV	kg/m2	0-75		Values larger than 75 kg/m^2 are clipped and flagged accordingly
TCWW uncertainty	kg/m2	0-60		
Convergence	logical			
Intermediate results, i.e. first guess, transmission, number of iterations				Can be used for debugging purpose

2.4.1 Input structure

The following lines give an example for the JSON input for the processor configured for MODIS (the comment behind the hash signs are not part of the JSON)

```
{'tmp':303.                 # surface temperature (or 2m) in [K]

,'prs':1003.                # surface pressure  [hPa]

,'suz':9.7968997955322270   # sun zenith [deg]

,'vie':46.12860107421875    # view zenith [deg]

,'azi':18.                  # azimuth difference = view_azi-sun_azi
(0,...180) [deg]

,'aot': {'2':0.1            # aot at wavelength of band 2

    , '5':0.08             # aot at wavelength of band 5

    , '17':0.1             # aot at wavelength of band 17

    , '18':0.1             # aot at wavelength of band 18

    , '19':0.1             # aot at wavelength of band 19

    }

,'sig_aot': {'2':0.1       # uncertainty of aot at wavelength of band 2

    , '5':0.1             # uncertainty of aot at wavelength of band 5

    , '17':0.1            # uncertainty of aot at wavelength of band 17
```

```
                 , '18':0.1                 # uncertainty of aot at wavelength of band 18

                 , '19':0.1                 # uncertainty of aot at wavelength of band 19

                 }

          ,'rtoa':{'2':  0.0636619783227144      # normalized radiance in band 2

                 , '5':  0.06525352778078226     # normalized radiance in band 5

                 , '17': 0.050292962874944384    # normalized radiance in band 17

                 , '18': 0.019098593496814323     # normalized radiance in band 18

                 , '19': 0.029921129811675773    # normalized radiance in band 19

                 }

          }
```

2.4.2 Output structure

The following lines show a corresponding example output. Most lines are only useful
for debugging. Relevant are the TCWV fields.

```
{'alb': {'17': 0.20744058708043236,

         '18': 0.20803469858273346,

         '19': 0.2080450532480437,

         '2': 0.20653781139830463,

         '5': 0.21382843735086557},

 'amf': 2.4577125799685628,

 'aot': {'17': 0.1, '18': 0.1, '19': 0.1, '2': 0.1, '5': 0.08},

 'azi': 18.0,

 'convergence': true,

 'fgu': 13.325100325715178,    #first guess TCWV in kg/m2

 'niter': 3,

 'prs': 1003.0,
```

```
 'rtoa': {'17': 0.050292962874944384,

          '18': 0.019098593496814323,

          '19': 0.029921129811675773,

          '2': 0.0636619783227144,

          '5': 0.06525352778078226},

 'rtoa_0': {'17': 0.06492165866278476,

            '18': 0.06501949555510961,

            '19': 0.06502120073716902,

            '2': 0.06477299167542006,

            '5': 0.06597359490147904},

 'sig_aot': {'17': 0.1, '18': 0.1, '19': 0.1, '2': 0.1, '5': 0.1},

 'sig_tcwv': 0.7624883523001241, #uncertainty TCWV in kg/m2

 'suz': 9.796899795532227,

 'tcwv': 15.249767046002482,  #retrieved TCWV in kg/m2

 'tmp': 303.0,

 'trans_fg': {'17': 0.8018521635828549, # gaseous transmission for retrieved
TCWV & geometry

              '18': 0.31104987302486253, # gaseous transmission for retrieved
TCWV & geometry

              '19': 0.47906689882847997, # gaseous transmission for retrieved
TCWV & geometry

              '2': 0.9828475831674877, # gaseous transmission for retrieved TCWV
& geometry

              '5': 0.9890855254776993}, # gaseous transmission for retrieved
TCWV & geometry

 'vie': 46.12860107421875}
```

3. WATER VAPOUR ALGORITHM FOR MICROWAVE IMAGER OBSERVATIONS

CDR-2 combines the EUMETSAT CM SAF HOAPS data over ocean with MERIS, MODIS and OLCI based observations over land, sea-ice and coasts.

The HOAPS product relies on inter-calibrated microwave imager observations. In particular, it utilises the CM SAF FCDR (Fennig et al., 2017) and currently CM SAF is working on including data from additional microwave imagers into HOAPS after application of the inter-calibration approach described in Fennig et al. (2020).

For CDR-2, the following input is used: the CM SAF FCDR with temporal coverage until 2015, SSM/I and SSMIS observations as beta-version of the CM SAF FCDR (2015-2017) and AMSR-E (2002-2011) and TMI (2002-2014) observations, inter-calibrated to the CM SAF FCDR. Further details are given in the HOAPS ATBD (2017).

The HOAPS product is developed, generated and documented by EUMETSAT CM SAF. The retrieval of L2 TCWV and associated uncertainties from microwave imager observations is based on a 1D-Var retrieval scheme which was provided by NWP SAF, a similar approach as described in section 2.3. The HOAPS ATBD (2017) describes the physical baseline for the retrieval of TCWV from microwave imager observations. Further reading on the 1D-Var retrieval is available at https://nwp-saf.eumetsat.int/site/software/1d-var/documentation/.

4. AGGREGATION AND COMBINATION

Data from NIR observations are at first aggregated and gridded onto a plat carrée longitude/latitude grid of 0.5° and 0.05° on a daily basis and per satellite. Every grid cell contains the average TCWV, the average uncertainty, the standard deviation within the grid cell and number of valid observation. Details of the aggregation are described in the SSD (2020) and PUG (2020). For the overlapping periods of MERIS-MODIS and MODIS-OLCI the results of both instruments are merged by a weighted sum. The weights are the respective counts of valid L2 retrievals. Additionally monthly averages are provided, created by a temporal aggregation of the daily products.

L2 data from CM SAF HOAPS is available from SSM/I, SSMIS, AMSR-E and TMI observations. In view of the differences in sampling, L2 TCWV are first gridded onto a plat carrée longitude/latitude grid of 0.5° on hourly basis and per satellite. Then, for each hourly bin averages over all available satellites are carried out. Finally, the daily average is computed based on the hourly bins. This way it is ensured that all hourly bins and satellites have the same weight during averaging. In each case the averaging is carried out over all valid observations. In order to generate a HOAPS based L3 product at 0.05° the 0.5° product is oversampled.

The combination of the microwave imager based TCWV data over ocean and the NIR based observations over land, sea-ice and coasts utilises the spatial complementarity of both products. Within HOAPS a land/sea mask (static) and a sea-ice mask (per month) are generated. While the sea-ice mask is variable in time valid HOAPS values are 50 km or more away from the coast. The masks are described in the HOAPS ATBD (2017). These masks are used to identify grids that are filled with microwave imager (open ocean) or with NIR (land, coasts and sea-ice) data. Further details are given in the PUG (2020).

The uncertainty propagation from L2 to L3 is described in the E3UB (2020).

5.   ASSUMPTIONS AND LIMITATIONS

Generally, the quality of the TCWV retrieval algorithm strongly depends on the reliability of the cloud mask. For example, MERIS and OLCI do not provide measurements in the thermal infrared, the screening of optically thin cirrus clouds is difficult. A dry bias is to be expected, where the cloud detection fails to detect cirrus clouds or sub pixel clouds. For other satellite instruments, such as MODIS, cloud-mask-products are more reliable.

The TCWV retrieval is applicable only for day light and clear-sky sampling and above land surfaces. Additionally, the SNR increases in the NIR for low surface reflectances. More importantly, the physical basis of the NIR retrieval; that the solar light is essentially reflected at the surface and thus transmitted through the full atmosphere, does not hold for very dark surfaces. Water surfaces outside the glint is black in the NIR and hence the solar light is reflected at aerosols in the boundary layer and not transmitted through the full atmosphere. Depending on their height, aerosol could or could not shield the water vapor absorption and lead to high ambiguity and uncertainty.

Assumptions and limitations of the HOAPS part of the product are given in the HOAPS ATBD (2017).

The gridded products are potentially affected by sampling biases. While the uncertainty arising from temporal sampling (NIR, HOAPS) and from non-sampling under strong scattering conditions (HOAPS) are assumed to be small, the bias between all-sky observations and clear-sky observations (as is the case for NIR TCWV data)can be fairly large. Such a bias was analysed but is not included as uncertainty information in the product (see CAR, 2020; E3UB, 2020).

The sampling differs between observations over land (NIR) and ocean (microwave imagers): clear-sky vs almost all-sky, day vs day and night and one satellite per day vs up to six satellites per day, respectively. The sampling was not aligned between land and ocean observations and by this the sampling capacity of microwave imagers over ocean has not been substantially reduced. Among others, the microwave imagers allow observations under almost all-sky conditions and by this, are not affected by a dry clear-sky bias as this is the case for NIR observations (CAR, 2020).

6. CONCLUSIONS

A retrieval method has been defined and developed to estimate TCWV from satellite measurements, using the $\rho\sigma\tau$-H_2O absorption band, as available from MERIS, MODIS and OLCI. The 1D-Var algorithm is based on a fast forward operator, which includes all relevant radiation interactions, basically, (1) transmittance due to atmospheric water vapour and (2) modification of the photon path due to atmospheric and surface scattering. An uncertainty estimate is given on a pixel-by-pixel basis where the uncertainties of measurement and forward model are considered respectively. TCWV from MERIS, MODIS and OLCI comprises CDR-1. This product contains valid values over land, under clear-sky conditions and during day time.

The retrieval of TCWV over ocean is also based on a 1D-Var retrieval scheme. The scheme is applied to microwave imager observations. This retrieval was originally developed by the NWP SAF and is operated by CM SAF to generate the HOAPS data.

The spatial complementariy of TCWV from HOAPS over ocean and from MERIS, MODIS and OLCI over land, sea-ice and coasts is utilised to comprise CDR-2. The NIR TCWV algorithm discriminate between land and ocean surfaces, while the later retrieval is optimised for dark surfaces, which is not appropriate for TCWV retrievals above bright sea-ice.

APPENDIX 1: REFERENCES

Albert, P., R. Bennartz, and J. Fischer, 2001: Remote Sensing of Atmospheric Water Vapor from Backscattered Sunlight in Cloudy Atmospheres. Journal of Atmospheric and Oceanic Technology, 18:865.

Albert, P., R. Bennartz, R. Preusker, R. Leinweber, and J. Fischer, 2005: Remote Sensing of Atmospheric Water Vapor Using the Moderate Resolution Imaging Spectroradiometer. Journal of Atmospheric and Oceanic Technology, 22:309.

Bartsch, B., S. Bakan, and J. Fischer, 1996: Passive remote sensing of the atmospheric water vapour content above land surfaces. Advances in Space Research, 18:25–28.

Bennartz, R. and J. Fischer, 2000: A modified k-distribution approach applied to narrow band water vapour and oxygen absorption estimates in the near infrared. Journal of Quantitative Spectroscopy & Radiative Transfer, 66:539–553.

Brockmann, C., Preusker, R., Danne, O., Kern, M., van As, D., Mankoff, K.D., and J. Box: A new method for cloud screening over the cryosphere. In preparation.

CAR, 2020: Climate Assessment Report. ESA Climate Change Initiative on Water Vapour, version 1.1, November 2020.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes. Journal of Quantitative Spectroscopy & Radiative Transfer, 91:233–244.

Cox, C. and W. Munk, 1954: Measurement of the roughness of the sea surface from photographs of the sun's glitter. Journal of the Optical Society of America (1917- 1983), 44:838.

Diedrich, H., R. Preusker, R. Lindstrot, and J. Fischer, 2015: Retrieval of daytime total columnar water vapour from MODIS measurements over land surfaces. Atmospheric Measurement Techniques, 8:823-836.

C. Donlon, B. Berruti, A. Buongiorno, M.-H. Ferreira, P. Féménias, J. Frerick, P. Goryl, U. Klein, H. Laur, C. Mavrocordatos, J. Nieke, H. Rebhan, B. Seitz, J. Stroede, R. Sciarra, 2012: The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. Remote Sensing of Environment, 120, Pages 37-57.

Doppler, L., R. Preusker, R. Bennartz, J. Fischer, 2013: k-bin and k-IR: k-distribution methods without correlation approximation for non-fixed instrument response function and extension to the thermal infrared. Journal of Quantitative Spectroscopy and Radiative Transfer, 09 / 2013.

Doppler, L., C. Carbajal-Henken, J. Pelon, F. Ravetta, and J. Fischer, 2014: Extension of radiative transfer code MOMO, matrix-operator model to the thermal infrared - Clear air validation by comparison to RTTOV and application to CALIPSO IIR, 144:49–67.

E3UB, 2020: End to End ECV Uncertainty Budget. ESA Climate Change Initiative on Water Vapour, version 2.0, November 2020.

Elliott, W. P., R. J. Ross, and W. H. Blackmore, 2002: Recent changes in NWS upper-air observations with emphasis on changes from VIZ to Vaisala radiosondes. Bull. Amer. Meteor. Soc., 83, 1003-1017.

Fennig, Karsten; Schröder, Marc; Hollmann, Rainer (2017): Fundamental Climate Data Record of Microwave Imager Radiances, Edition 3, Satellite Application Facility on Climate Monitoring, DOI:10.5676/EUM_SAF_CM/FCDR_MWI/V003, https://doi.org/10.5676/EUM_SAF_CM/FCDR_MWI/V003.

Fennig, K., Schröder, M., Andersson, A., and Hollmann, R.: A Fundamental Climate Data Record of SMMR, SSM/I, and SSMIS brightness temperatures, Earth Syst. Sci. Data, 12, 647–681, https://doi.org/10.5194/essd-12-647-2020 , 2020.

Fischer, J. 1988: High Resolution Spectroscopy for Remote Sensing of Physical Cloud Properties and Water Vapour. In: Current Problems in Atmospheric Radiation, Ed. Lenoble and Geleyn, Deepak Publishing, 151-156.

Gao, B.-C. and A. F. H. Goetz, 1990: Column atmospheric water vapor and vegetationliquid water retrievals from Airborne Imaging Spectrometer data. Journal of Geophysic Research, 95: 3549–3564.

Gao, B.-C., A. F. H. Goetz, E. R. Westwater, J. E. Conel, and R. O. Green, 1993: Possible Near-IR Channels for Remote Sensing Precipitable Water Vapor from Geostationary Satellite Platforms. Journal of Applied Meteorology, 32: 1791–1801.

Gordon I.E., L.S. Rothman, C. Hill et al., "The HITRAN2016 Molecular Spectroscopic Database", J Quant Spectrosc Radiat Transfer 203, 3-69 (2017).

Held, I. and Soden, B., 2000: Water vapor feedback and global warming, Annu. Rev. Energy Environ., 25: 441-475.

HOAPS ATBD, 2017: Algorithm Theoretical Baseline Document - HOAPS version 4.0. EUMETSAT Satellite Application Facility on Climate Monitoring, Issue 2.3, 31 January 2017, available at https://doi.org/10.5676/EUM_SAF_CM/HOAPS/V002.

Hollstein, A. and J. Fischer, 2012: Radiative transfer solutions for coupled atmosphere ocean systems using the matrix operator technique. Journal of Quantitative Spectroscopy & Radiative Transfer, 113: 536–548.Kiehl, J.T., and K.E. Trenberth, 1997: Earth's annual global mean energy budget. Bull. Am. Meteorol. Soc., 78, 197–208.Lindstrot, R. and R. Preusker, 2012: On the efficient treatment of temperature profiles for the estimation of atmospheric transmittance under scattering conditions. Atmospheric Measurement Techniques, 5: 2525–2535.

Lindstrot, R., R. Preusker, H. Diedrich, L. Doppler, R. Bennartz, and J. Fischer, 2012: 1D-Var retrieval of daytime total columnar water vapour from MERIS measurements. Atmospheric Measurement Techniques, 5: 631–646.

Mieruch, S., M. Schröder, S. Noël, and J. Schulz, 2010: Comparison of monthly means of global total column water vapor retrieved from independent satellite observations. Journal Of Geophysical Research, Vol. 115.

Mlawer, E. J., Payne V. H., Moncet, J., Delamere, J. S., Alvarado, M. J., and Tobin, D.C., 2012: Development and recent evaluation of the MT_CKD model of continuum absorption, Philos. T. R. Soc. A, 370, 2520–2556, https://doi.org/10.1098/rsta.2011.0295 , 2012.

MODIS Cloud Mask Team: Discriminating clear-sky from cloud with MODIS algorithm: theoretical basis document (MOD35). CIMSS, University of Wisconsin – Madison. Version 6.1, October 2010.

Ptashnik I.V., K.P. Shine, A.A. Vigasin, 2011: Water vapour self-continuum and water dimers: 1. Analysis of recent work. J. of Quant Spect. and Rad. Trans. 112(8); 1286-303.

Preusker and Fischer, 2018: Estimation of un-correlated k-binning coefficients from calculated line-by-line spectral absorption. Python and Fortran-code, private communication.

Preusker and Fischer, 2019: Improvement in Copernicus Sentinel-3 OLCI Water Vapour Product' (COWa). ATBD, EUMETSAT.

PUG, 2020: Product User Guide. ESA Climate Change Initiative on Water Vapour, version 1.1, October 2020.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor, 2007: Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H .L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Rodgers, C., 2000: Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific Pub Co.

Rothman, L.S, I.E. Gordon, Y. Babikov, A. Barbe, D.C. Benner, P.F. Bernath, et al., 2013: The HITRAN2012 molecular spectroscopic database. JQSRT, Volume 130, Pages 4-50.

Shine, K. P., A. Campargue, D. Mondelain, R. A. McPheat, I. V. Ptashnik, D. Weidmann, 2016: The water vapour continuum in near-infrared windows – Current understanding and prospects for its inclusion in spectroscopic databases. Journal of Molecular Spectroscopy 327 (2016) 193–208.

SSD: ESA CCI Water Vapour: System Specification Document. O. Danne and M. Hegglin, Issue 2.0, 15 May 2020

Trenberth, K.E., and D.P. Stepaniak, 2003a: Co-variability of components of poleward atmospheric energy transports on seasonal and interannual timescales. J. Clim., 16: 3690−3704.

Trenberth, K.E., and D.P. Stepaniak, 2003b: Seamless poleward atmospheric energy transports and implications for the Hadley circulation. J. Clim. 16: 3705–3721.

Xavier, P. K., V. O. John, S. A. Buehler, R. S. Ajayamohan, and S. Sijikumar, 2010: Variability of Indian summer monsoon in a new upper tropospheric humidity data set. Geophys. Res. Lett., 37, L05705.

APPENDIX 2: GLOSSARY

Term	Definition
ATBD	Algorithm Theoretical Basis Document
CM SAF	Satellite Application Facility on Climate Monitoring
Envisat	ESA satellite (see http://envisat.esa.int/)
EO	Earth Observation
ESA	European Space Agency (http://www.esa.int)
ESTEC	European Space Research and Technology Centre
ESRIN	European Space Research Institute (https://www.esa.int/About_Us/ESRIN)
GCOS	Global Climate Observing System
GEWEX	Global Energy and Water Exchanges Project
GNSS	Global navigation satellite system
GPS	Global Positioning System
GUAN	GCOS Upper Air Network
HITRAN	High-resolution transmission molecular absorption database
HOAPS	Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite
IODD	Input Output Data Definition Doc.
ISCCP	GEWEX International Satellite Cloud Climatology Project
IWV	Integrated water vapour (=TCWV)
L1/L2	Level 1 / Level 2
LBL	Line-by-line
LUT	Look-up table
MERIS	Medium Resolution Imaging Spectrometer Instrument (http://envisat.esa.int/)
MODIS	Moderate Resolution Imaging Spectroradiometer (on board the NASA EOS-Aqua satellite)
MOMO	Matrix Operator Modell
MSG	METEOSAT Second Generation
MWR	Microwave Radiometer
NASA	National Aeronautics and Space Administration
OE	Optimal estimation

Term	Definition
OLCI	Ocean and Land Colour Instrument on board Sentinel-3
POLDER	Polarization and Directionality of the Earth's Reflectances
RTC / RTM	Radiative Transfer Code / Model
TCWV	Total Column Water Vapour
TOA	Top of atmosphere
1DVar	1 Dimensional Variational

APPENDIX 3: N-DIMENSIONAL INTERPOLATION

Suppose we have an n-dimensional look up table, representing a function

$$y = f(x_1, \ldots, x_n). \tag{16}$$

The linear interpolation of f to the point $(x_{1i}, \ldots, x_{ni})$ is calculated by the weighted sum over the 2^n neighbours $\vec{x}_j$ creating the surrounding rectangular polytope:

$$f(\vec{x}_i) = \sum_{j=1}^{2^n} f(\vec{x}_j) \cdot w_j \tag{17}$$

The weights are the normalized volumes of the opposite n-dimensional rectangular polytopes created by the ambient points and the interior point $\vec{x}_i$:

$$w_{00\ldots0} = \frac{1}{N} \cdot (x_{1u} - x_{1i}) \cdot (x_{2u} - x_{2i}) \cdot \ldots \cdot (x_{nu} - x_{ni})$$

$$w_{00\ldots1} = \frac{1}{N} \cdot (x_{1u} - x_{1i}) \cdot (x_{2u} - x_{2i}) \cdot \ldots \cdot (x_{ni} - x_{nl}) \tag{18}$$

$$w_{11\ldots1} = \frac{1}{N} \cdot (x_{1i} - x_{1l}) \cdot (x_{2i} - x_{2l}) \cdot \ldots \cdot (x_{ni} - x_{nl})$$

The first index of x indicates the dimension, the second the neighbour point position: l is the lower point and u is the upper point of the polytope. The weight index sequence of zeros and ones is the binary representation of j. A zero means, that the polytope corner is the lower one in the respective dimension, a one means the upper corner. N is the total volume of the polytope:

$$N = (x_{1u} - x_{1l}) \cdot (x_{2u} - x_{2l}) \cdot \ldots \cdot (x_{nu} - x_{nl}) \tag{19}$$

APPENDIX 4: K-BINNING

The transmission is derived from pre-calculated absorption coefficients using an advanced k-distribution method (Bennartz and Fischer, 2000; Doppler et al, 2013). Basis are spectrally very high resolved extinction coefficients for a given profile of temperature, pressure and humidity. In order to obtain the transmission of the entire band, which contains a large number of spectral lines, the quasi-monochromatic intervals are sorted with respect to their optical thickness. Afterwards *similar* high resolution intervals are combined (binned) to a significantly lower number of pseudo-spectral interval *i* for every vertical layer *j* of the preconditioned TCWV-profile. A weight factor w_i is assigned to each bin, representing the number of included high resolution intervals and the relative weight with respect to the bands instrumental line shape. Subsequently, for every bin the transmittance is derived from the sum of the optical depth for all layers along the line of sight. The total transmittance is the sum over all pseudo spectral intervals multiplied with the associated weights w_i accordingly:

$$T = \sum_{i=1}^{\#bins} w_i * exp\left(- \sum_{j=1}^{\#layers} \tau_{ij} * amf * sf_j \right)$$

(20)

The scaling factor *sf* is the ratio of the actual amount of water vapor in the layer *j* to the amount used to calculate the high-resolution extinctions in layer *j*.

End of Document