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

End to End ECV Uncertainty Budget (E3UB) - Part 1: CDR-1 & CDR-2

Ref: D2.3

Date: 03 February 2021

Issue: 2.1

For: ESA / ECSAT

Ref: CCIWV.REP.010

Project :	**Water Vapour Climate Change Initiative (WV_cci) - CCI+ Phase 1**
Document Title:	**End to End ECV Uncertainty Budget (E3UB) - Part 1: CDR-1 & CDR-2**
Reference :	**D2.3**
Issued :	**03 February 2021**
Issue :	**2.1**
Client :	**ESA / ECSAT**

Author(s) : Juergen Fischer (SE), Rene Preusker (SE), Marc Schröder (DWD)

Document Change Log

Issue/ Revision	Date	Comment
1.0	12 June 2019	Initial issue
1.1	4 Sep 2019	Revised version based on RIDs of v1.0
2.0	05 November 2020	• Initial version for second review by ESA • Corrected typos in L3 uncertainty • Included measurement transformation in NIR • Links to HOAPS documentation included and with this changes to the structure of the document
2.1	03 February 2021	• Restructured • Added budget estimation section 2.1.2

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Purpose

The purpose of this document is to deliver an end-to-end uncertainty budget that is based on estimates of uncertainties that arise from each step within the merging process chain of the climate data record, here for the deliverables CDR-1 and CDR-2.

1.2 Scope

The end-to-end uncertainty budget aims at including the uncertainties from the two processing steps (1) L1 to L2 (retrieval step) and (2) L2 to L3 (merging step). The retrieval step consists of uncertainties based on each step in the retrieval process including instrument noise characteristics, geolocation and geophysical product retrieval. The merging step consists of uncertainties based on each step in the merging process chain including the uncertainty arising from geophysical sampling biases when moving from L2 to L3 for each satellite product and the uncertainty in the use of external ancillary data (i.e. model quality used as transfer function between each satellite product).

1.3 Terminology

The approach follows Merchant et al. (2017) and the recommendations therein, which are applied widely in the CCI programme and which also forms the main basis for definitions given in the PVP v2.0.

A *measurement* is a set of operations having the object of determining the value of quantity, the *measurand*. The measured value minus the true value of the measurand is the error, which is generally unknown. The measurand in CDR-1 and CDR-2 is the *total column water vapour*, the vertically integrated water vapour of the full atmosphere in units kg/m^2.

Uncertainty is the lack of *certainty* of knowledge about a state. With regards to a measurement of a measurand, uncertainty is quantified by the statistical dispersion of the measurement values, commonly the second moment (standard deviation) of the corresponding frequency distribution.

The uncertainty of measurement can be estimated by *propagating* uncertainties of assumptions, auxiliary data, instrumental limitations and specificities, model simplifications, etc., through a mathematical description of the measurement and retrieval process. For the total column water vapour retrieval we perform a *linear* *propagation of uncertainty*, which basically assumes, that the effect of small perturbations of the retrieval process can be described by a first-order (linear) Taylor expansion.

Alternatively and complementary uncertainty can be characterised and quantified by comparing measurements to reference measurements, which are trusted (ideally with a low uncertainty). Eventually, these comparisons can be further used to validate the uncertainty propagation.

The *end-to-end uncertainty budget* is the result (where possible) of the uncertainty propagation to the final product, ideally an uncertainty per datum and validated using comparisons to reference measurements (see Merchant et al., 2017).

2. TOTAL COLUMN WATER VAPOUR CONTENT CLIMATE DATA RECORD (CDR-1 AND CDR-2)

CDR-1, a vertically integrated (total column) water vapour ECV, in units of kg/m^2, is a gridded L3 data product over land based on L2 retrievals applied to MERIS, MODIS and OLCI measurements. The same product will be combined with the EUMETSAT CM SAF HOAPS product over ocean, with sea-ice and coastal areas filled with NIR observations. This global TCWV product comprises 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.

2.1 TCWV MERIS, OLCI, and MODIS L2 data uncertainties (CDR-1 and CDR-2)

Total column water vapour, in units of kg/m^2, is retrieved from MERIS, MODIS and OLCI observations. In this section, the estimation of associated L2 uncertainties is described. The uncertainty estimation is applied over land, coasts and sea-ice.

2.1.1 Quantification of linear uncertainty based on uncertainty propagation of the inversion from L1 to L2 [TCWV-NIR]

The algorithm to estimate water vapour from MERIS, OLCI and MODIS measurements is based on an optimal estimation (OE) inversion of a forward model *F*(*x*). OE comprehends a linear uncertainty estimation, which propagates uncertainties of parameterisations, measurements and ancillary data to product uncertainties.

Main inputs for the propagation are S_E and S_A, being the measurement and *a priori* error co-variance matrices, respectively. To incorporate forward model uncertainties, S_E is composed from the error co-variance matrix of the measurement S_M (from e.g. instrument signal to 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*:

$$S_E = S_M + K_B^T S_B K_B$$

This approach assumes Gaussian probability density functions and bias-free measurements, priors and models. Using this quantity and the Jacobian of the forward model *F* with respect to the state *K*, OE automatically provides linear uncertainty measures:

- The retrieval error co-variance. The diagonal element that belongs to the TCWV entry of the state vector is the uncertainty σ^2_{TCWV} of the water vapour, resulting from the linear propagation of measurement and parameter error, expressed by S_a:

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

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

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

using the gain *G*:

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

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

$$dof \;= tr\; A$$

- The information content *H*, as the logarithm of the distance of the prior and retrieval (posterior) error co-variance:

$$H = \frac{1}{2} \ln \left| S_a \cdot \hat{S}^{-1} \right|$$

- The retrieval noise:

$$S_n = G S_e G^T$$

- The smoothing error:

$$S_S = (I - A)S_a(I - A)^T$$

where *I* is the identity matrix.

The residual difference between the optimal simulated measurement and the measurement:

$$\eta = F(\hat{x}) - Y$$

may give indications to deficits in the measurement (calibration or band characterisation) or the radiative transfer (band setting, parameterisations, interpolations) or other anomalies (deficits in pixel identification, in particular cloud detection).

A common and useful treatment of an inverse problem to increase convergence speed is to transform it to a ‘more linear’ form. This has also been done for the NIR TCWV retrieval. The measurement in the absorption bands is transformed to quantities, which almost linearly depend on the total amount of water vapour by:

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

ρ_i is the top of atmosphere reflectance in the respective absorption bands. amf is the air mass factor, $\tilde{\rho}_i$ is top of atmosphere reflectance of the window bands (ρ_0, ρ_1) inter- or extrapolated to the wavelength λ_i of the absorption band:

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

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

Since the measurement is transformed, the corresponding measurement error co-variance must be transformed as well. Assuming that the SNR is approximately the same for all bands and a simple uncertainty propagation of upper equation, the uncertainty in the transformed absorption bands is:

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

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

Here we must assume an uncertainty σ_{inter}^2 (of e.g. 10% interpolation error which results in 0.01 for σ_{inter}^2) due to the extra- or interpolation of the window bands to the absorption bands. Eventually $\sigma_{\tau\,i}^2$ are used to populate the diagonal elements of the measurement error covariance S_M .

2.1.2 Sources and effects of uncertainties for the NIR TCWV product

The following bullet points list the known sources of uncertainty for the NIR TCWV product and give an approximation of their impact.

- **The signal-to-noise ratio**. Currently OLCI and MERIS do not provide a pixel by pixel uncertainty of the L1b measurement. Instead, we refer to the ESA requirement of approximately 300 for the NIR bands (Donlon et al., 2012; Drinkwater et al., 2007). Validation activities of the mission performance centre indicate a higher (better) signal-to-noise ratio (Bourg et al., 2019). However, the crucial point is, that signal-to-noise ratios are always given at a *typical signal level* and they cannot simply be transferred to a specific measurement. Thus, we decided to assume conservatively 300 for all MERIS and OLCI measurements. MODIS provides signal-to-noise ratios between 60 and 250, but for very low *typical radiances* (factor 10 smaller than what is expected for land surfaces; see MODIS specifications). Nevertheless, for MODIS we assume 200. Eventually, the effect of the pure signal-to-noise ratio on the uncertainty of the retrieved water vapour is of the same order of magnitude, certainly below 1%.

- σ^2_{inter}, **the uncertainty of interpolation and inter-calibration** of the window reference bands to the spectral position of the water vapour band. Lindstrot et al. (2012) have shown that the uncertainty of a corresponding interpolation for MERIS is in most cases 0.001–0.004 (in albedo units). Typical surfaces have spectral albedos in the NIR in the order of 0.1–0.8, so the corresponding interpolation error is somewhere between 0.001% and 0.04%. This is in the same order of magnitude as found in uncertainties of the inter-band radiometric calibration of OLCI (Lamquin et al., 2020). They did not investigate the water vapour bands, but all investigated bands did show similar features. The relative calibration accuracy for MODIS is in the order of 2% (Xiong and Butler, 2020), but the spectral distance between window reference and the absorption bands is larger. Hence, an assumed interpolation uncertainty of 8% is reasonable. This translates into a similar uncertainty for the total column of water vapour.
- **Surface pressure and surface temperature**. These two quantities are used to constrain the vertical profile of temperature. Lindstrot et al. (2012) have investigated the effect of this simplified assumption. They found that the resulting error is smaller than 5% (maximum) and the uncertainty is smaller than 2% (standard deviation).
- **Uncertainty of aerosol optical thickness and height**. Depending on the amount and height of the aerosols and on the brightness of the surface, aerosols alter the mean geometrical path of photons. This alters the measured transmission, when, in particular, humid low-level layers of the atmosphere are shielded. We use the MODIS L3 collection 6.1 data as source with an uncertainty of 20% + 0.1 (0.2) (Wei et al., 2019). In the work of Lindstrot et al. (2012), the influence of aerosol thickness and height is quantified by a scattering factor. As long as the surface albedo is above 0.2 and aerosol optical thickness at 900 nm is below 0.4, the influence of a change in aerosol thickness of ±0.1 on the signal in the simulated effective transmission is below 2%. The impact of aerosol uncertainty can increase dramatically, when the surface is dark, such as water surfaces. Here the perturbation can easily reach 100% (see Lindstrot et al., 2012; and Diedrich et al., 2013).
- **Uncertainties of spectral characterisation**. Due to the optical design of OLCI and MERIS, all spectral bands show a slightly varying spectral response over their field of view, and further small discontinuities (Sentinel 3 CalVal Team, 2016). The central wavelengths exhibit variations of up to 1 nm and jumps of up to 1.6 nm between the cameras. A visual inspection of the retrieved fields of total

column water vapour from OLCI and MERIS did not indicate that the shifts and jumps create visual discontinuities in the water vapour fields, and so the varying spectral specifics have not been considered in forward models. They assume the nominal spectral response functions. A first quantitative assessment of the impact by Preusker et al. (2021) shows an uncertainty of 0.5 kg/m^2 for an atmosphere of 16 kg/m^2 which is approximately 3%. MODIS does not show spectral variation over the field of view.

- **Uncertainty due to undetected clouds**. An undetected cloud would shield the moist atmosphere below the cloud from solar radiation. This leads to a systematic underestimation of the total column of water vapour. For a single incidence the shielding could reach 100%. If the cloud detection systematically misses clouds, the resulting total column L3 product would be dry biased.

From the above listed sources of uncertainty, the signal-to-noise ratio and uncertainty of the interpolation and intercalibration are considered in the linear uncertainty propagation. They are responsible for up to 10% uncertainty of the retrieved total column water vapour. *The interpolation uncertainty is the dominating factor*.

The uncertainty of the spectral characterisation of OLCI and MERIS contributes 3%, the forward model simplifications for the humidity and temperature profile 2%. The effects of aerosol uncertainty are small, if the surface is bright, which is the normal case for natural surfaces. Hence, we would expect a total uncertainty of around 15%. However, the comparisons with ground truth (GNSS and ARM microwave radiometer, see PVIR) show an agreement between the satellite products and the ground truth of approximately 12%. This is a hint that some sources of uncertainty are currently overestimated.

Dark surfaces (water) and clouds are improper conditions for the retrieval of total column water vapour in the NIR. Clouds must be avoided. For dark water surfaces, the uncertainty of aerosol scattering properties dominate the uncertainty budget.

2.2 TCWV CM SAF HOAPS L2 data uncertainties (CDR-2)

CDR-2 combines the EUMETSAT CM SAF HOAPS TCWV data over ocean with MERIS, MODIS and OLCI based TCWV data over land, sea-ice and coasts. 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 to CM SAF by NWP SAF, with similar approaches to estimate uncertainties as outlined in section 2.1. CM SAF is not providing a stand-alone E3UB document. Instead, the HOAPS ATBD (2017) is referenced, and the NWP SAF documentation on the 1D-Var retrieval is available at https://nwp-saf.eumetsat.int/site/software/1d-var/documentation/ .

2.3 L3 Data uncertainties (CDR-1 and CDR-2)

L3 products utilise the provided L2 uncertainty information as far as possible. We propose to be coherent to other CCI products (e.g. clouds). The following quantities

are provided per L3 bin, assuming that N samples of a quantity x are averaged

- Standard deviation of x:

$$\sigma^2_{\mathrm{x\,std}} = \frac{1}{\mathrm{N}} \sum_{1}^{\mathrm{N}} (\mathrm{x_i} - \langle \mathrm{x} \rangle)^2$$

- Mean uncertainty of x:

$$\langle \sigma_{\mathrm{x\,i}} \rangle = \frac{1}{\mathrm{N}} \sum_{1}^{\mathrm{N}} (\sigma_{\mathrm{x\,i}})$$

- Mean of the squared uncertainty (propagated uncertainty) of x:

$$\sigma^2{}_{\mathrm{x\,prop}} = \frac{1}{\mathrm{N}} \sum_{1}^{\mathrm{N}} \sigma^2_{\mathrm{x\,i}}$$

The above-mentioned L3 uncertainty terms can be used to approximate further properties: the natural variability of the variables observed and the uncertainty of averages (av), i.e. the propagated L2 to L3 uncertainty. Following Stengel et al. (2017) these two terms can be calculated by:

$$\sigma^2_{x\,natural} = \sigma^2_{x\,std} - (1 - c)\langle \sigma_{x\,i} \rangle^2$$

$$\sigma^2{}_{x\,av} = \frac{1}{N}\sigma^2_{x\,natural} + c\langle\sigma_{x\,i}\rangle^2 + (1-c)\frac{1}{N}\langle\sigma^2_{x\,i}\rangle$$

The latter equation allows the computation of the propagated uncertainty of L3 products. The most critical term however is the correlation *c* between the spatially and temporally averaged retrievals. This quantity is typically not known. Thus, the products include the standard deviation, the mean uncertainty and the mean of squared uncertainties. Then, when estimates of the correlation become available, e.g. over specific regions and time windows, the total uncertainty can be estimated using the provided information on uncertainty.

In order to provide an impression of the impact of the correlation on propagated uncertainty and consistency the extreme cases with c=0 and c=1 will be applied during validation. First analysis will be carried out to estimate the correlation by analysing time series of GPS and spatial variability of satellite imagery (Diedrich et al., 2016). Finally, the CDR-1 and CDR-2 products are potentially affected by sampling biases. In particular, the NIR retrievals are predominantly observed once per day and are applied under clear-sky conditions. Conditions in clouds are typically more humid than the surrounding clear-sky areas and are not taken into account by the satellite's clear-sky observations. This effect is called clear-sky bias (CSB, Sohn and Bennartz, 2008). In order to characterise the CSB for gridded and temporally averaged clear-sky products, ERA5 data was used to analyse differences between all-sky and clear-sky cases as a function of local time, month, longitude and latitude. Due to the local time dependency the CSB also includes uncertainties arising from temporal sampling. It is noted that in view of the fairly small diurnal cycle of TCWV (Diedrich et al., 2016; Schröder et al., 2017) this impact is assumed to be small. Full results are given and discussed in the CAR (2020). Because the estimated CSB is hardly significant it was concluded not to include the CSB into CDR-1 and CDR-2 (CAR, 2020). However, the CSB will be considered during consistency analysis.

TCWV from microwave imager observations over ocean is not retrieved in the presence of strong scattering. This might cause a sampling bias. Mears et al. (2018) analysed this bias relative to GNSS observations and observed a systematic, but small bias. They did not include this bias in their uncertainty estimate. Here, such a bias is not considered either.

APPENDIX 1: REFERENCES

Bourg, L.; Blanot, L.; Alhammoud, B.; Sterck, S.; Preusker, R. Sentinel-3 A and B OLCI instruments Calibration Status. In Proceedings of the Fifth Sentinel-3 Validation Team Meeting, Frascati, Italy, 7–9 May 2019.

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

Diedrich, H., R. Preusker, R. Lindstrot, and J. Fischer, 2013: Quantification of uncertainties of water vapour column retrievals using future instruments. Atmos. Meas. Tech., 6, 359–370, https://doi.org/10.5194/amt-6-359-2013.

Diedrich Hannes, Falco Wittchen, René Preusker, and Jürgen Fischer, 2016: Representativeness of total column water vapour retrievals from instruments on polar orbiting satellites. Atmos. Chem. Phys., 16, 8331–8339.

Donlon C, Berruti B, Buongiorno A, et al. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission Remote Sensing of environment.. 2012 May;120:37-57. DOI: 10.1016/j.rse.2011.07.024.

Drinkwater, M.R.; Rebhan, H. Sentinel-3 Mission Requirements Document. Ref: EOP-SMO/1151/MD-md. 2007. Available online: http://esamultimedia.esa.int/docs/GMES/GMES_Sentinel3_MRD_V2.0_update.pdf (accessed on 30 May 2020).

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.

Lamquin, N.; Clerc, S.; Bourg, L.; Donlon, C. OLCI A/B Tandem Phase Analysis, Part 1: Level 1 Homogenisation and Harmonisation. Remote Sens. 2020, 12, 1804.

Lindstrot, R., Preusker, R., Diedrich, H., Doppler, L., Bennartz, R.,and Fischer, J.: 1D-Var retrieval of daytime total columnar water vapour from MERIS measurements, Atmos. Meas. Tech., 5, 631–646, doi:10.5194/amt-5-631-2012, 2012.

Mears, C. A., Wang, J., Smith, D., and Wentz, F. J. (2015), Intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and ground-based GPS instruments. J. Geophys. Res. Atmos., 120, 2492– 2504. doi: 10.1002/2014JD022694.

Mears, C. A., Smith, D. K., Ricciardulli, L., Wang, J., Huelsing, H., & Wentz, F. J. (2018). Construction and Uncertainty Estimation of a Satellite-Derived Total Precipitable Water Data Record Over the World's Oceans. Earth and Space Science, 5. https://doi.org/10.1002/2018EA000363.

Merchant CJ, Paul F, Popp T, Ablain M, Bontemps S, Defourny P, et al. Uncertainty information in climate data records from Earth observation. Earth Syst Sci Data. 2017;9(2):511-27.

MODIS-PERF online: https://modis.gsfc.nasa.gov/about/specifications.php#3 (last time accessed on 29 Jan 2021).

Preusker, R., Carbajal Henken, Fischer, J. 2021: Retrieval of daytime total column water vapour fromOLCI measurements over land surfaces; Remote Sens. 2021 submitted

PVP, 2020: Product Validation Plan. ESA Climate Change Initiative on Water Vapour, version 2.0, November 2019.

Schröder, M., Lockhoff, M., Shi, L., August, T., Bennartz, R., Borbas, E., Brogniez, H., Calbet, X., Crewell, S., Eikenberg, S., Fell, F., Forsythe, J., Gambacorta, A., Graw, K., Ho, S.-P., Höschen, H., Kinzel, J., Kursinski, E.R., Reale, A., Roman, J., Scott, N., Steinke, S., Sun, B., Trent, T., Walther, A., Willen, U., Yang, Q., 2017: GEWEX water vapor assessment (G-VAP). WCRP Report 16/2017; World Climate Research Programme (WCRP): Geneva, Switzerland; 216 pp. Available at https://www.wcrp-climate.org/resources/wcrp-publications.

Sentinel 3 CalVal Team. Technical Note: Sentinel-3 OLCI-A spectral response functions. Technical report,684ESA, 2016. Reference: S3-TN-ESA-OL-660.

Sohn, B.-J., and R. Bennartz (2008), Contribution of water vapor to observational estimates of longwave cloud radiative forcing, J. Geophys. Res., 113, D20107, doi:10.1029/2008JD010053.

Stengel, M., Stapelberg, S., Sus, O., Schlundt, C., Poulsen, C., Thomas, G., Christensen,M., Carbajal Henken, C., Preusker, R., Fischer, J., Devasthale, A., Willén, U., Karlsson, K.-G., McGarragh, G. R., Proud, S., Povey, A. C., Grainger, R. G., Meirink, J. F., Feofilov, A., Bennartz, R., Bojanowski, J. S., and Hollmann, R.: Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud_cci project, Earth Syst. Sci. Data, 9, 881-904, https://doi.org/10.5194/essd-9-881-2017, 2017.

Wei, J., Peng Y. , Guo J., Sun L.: Performance of MODIS Collection 6.1 Level 3 aerosol products in spatial-temporal variations over land, Atmospheric Environment, Volume 206, 2019, Pages 30-44, ISSN 1352-2310,

Xiong, X.; Butler, J.J. MODIS and VIIRS Calibration History and Future Outlook. Remote Sens. 2020, 12, 2523. https://doi.org/10.3390/rs12162523

APPENDIX 2: GLOSSARY

Term	**Definition**
CDR	Climate Data Record
CM SAF	Satellite Application Facility on Climate Monitoring
E3UB	End to End ECV Uncertainty Budget
ECV	Essential Climate Variable
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
ESA	European Space Agency
HOAPS	Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data
MERIS	MEdium Resolution Imaging Spectrometer
MODIS	Moderate-resolution Imaging Spectroradiometer
NWP SAF	Satellite Application Facility on Numerical Weather Prediction
OE	Optimal Estimation
OLCI	Ocean and Land Colour Instrument
TCWV	Total Column Water Vapour

End of Document