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

Product Specification Document (PSD)

Ref: D1.2

Date: 27 July 2021

Issue: 3.2

For: ESA / ECSAT

Ref: CCIWV.REP.002

This Page is Intentionally Blank

Project	**:**	**Water Vapour Climate Change Initiative (WV_cci) - CCI+ Phase 1**
Document Title:		**Product Specification Document (PSD)**
Reference	**:**	**D1.2**
Issued	**:**	**27 July 2021**
Issue	**:**	**3.2**
Client	**:**	**ESA / ECSAT**

Authors : Marc Schröder (DWD), Michaela Hegglin (U. Reading), Hao Ye (U. Reading), Ulrike Falk (DWD), Olaf Danne (Brockmann Consult), Jürgen Fischer (Spectral Earth), Alexandra Laeng (KIT), Richard Siddans (RAL), Christopher Sioris (EC), Gabriele Stiller (KIT), Kaley Walker (U. Toronto)

Document Change Log

Issue/ Revision	Date	Comment
1.0, draft	11 Dec 2018	Draft submitted to ESA for review
1.1	31 Jan 2019	RIDs from ESA implemented
2.0, draft	01 Nov 2019	Update based on discussions triggered at PM4, draft submitted to ESA for review
2.1	27 Nov 2019	RID from ESA implemented, updates following updated URD v2
3.0	19 Feb 2021	Third version, submitted to ESA for review
3.1	22 March 2021	Updates following discussions at PM9
3.2	27 July 2021	Final updates prior to final meeting

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

1. INTRODUCTION

1.1 Purpose of the document

The purpose of this Product Specification Document (PSD) is to provide detailed specifications on the different CDR products delivered for the ECV Water Vapour within the WV_cci project. The specifications aim at compliance with the user requirements from GCOS-200, other CCI projects as defined within the CCI Data Standards (v2.1) and established in consultation with the WV_cci user group and CRG that are summarised in the URD (v3.0).

1.2 Definitions

The terminology is given in the PVP (PVP, 2021).

2. WATER DATA PRODUCT OVERVIEW

This section provides an overview of the different water vapour data products that will be available from the WV_cci at the end of the project. Detailed specifications of these data products and their input data are provided in Section 3.

The primary datasets (or CDRs) that WV_cci produced are illustrated in Figure 2-1. These include gridded L3 total column water vapour over land (CDR-1) and merged over land and ocean (CDR-2), and also L3 vertically resolved water vapour with focus on the stratosphere (CDR-3) and with focus on the UTLS (CDR-4). In addition to the primary datasets, WV_cci also makes the input data (L2 and L3) that are used to generate CDR-3 and CDR-4 available.

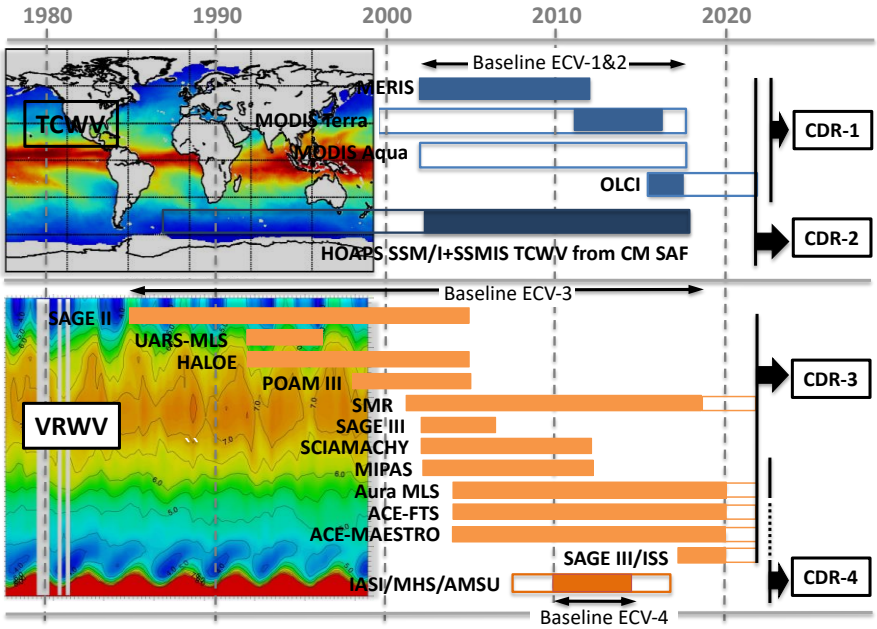

Figure 2-1: Instruments from which input observations (L2 data products) are obtained, and resulting climate data records (CDRs; L3 data products) for total column (TCWV; CDR-1 and CDR-2) and vertically resolved (VRes WV; CDR-3 and CDR-4) water vapour.

Table 2-1: Overview of final data products delivered by WV_cci

#	Product Name	Level	Notes	Time period
1	CDR-1	L3	TCWV gridded monthly and daily data (land)	2002–2017
2	CRD-2	L3	TCWV gridded monthly and daily data (land and ocean)	2002–2017
3	CDR-3	L3	Stratospheric vertically resolved zonal monthly mean WV CDR (2D-fields)	1985–2019
4	CDR-4	L2 & L3	UTLS vertically resolved monthly mean WV CDR (3D-fields)	2010–2014

2.1 CDR-1: TCWV (land)

This vertically integrated (total column) water vapour ECV, in units of kg/m^2, is a gridded L3 data product over land based on ESA (MERIS, OLCI) and NASA instruments (MODIS). The final dataset covers the period July 2002 to December 2017.

2.2 CDR-2: TCWV (land and ocean)

The CDR-2 dataset is a global, i.e. land and ocean coverage, TCWV product in units of kg/m^2. It contains the CDR-1 over land, coasts and sea-ice and the HOAPS microwave imager based TCWV data over ocean. The HOAPS data are generated by EUMETSAT CM SAF. CDR-2 is released by the EUMETSAT CM SAF. The final dataset covers the period July 2002 to December 2017.

2.3 CDR-3: Vertically resolved stratospheric water vapour climatologies

This vertically resolved water vapour ECV is a merged data product based on a range of ESA, third-party and NASA instruments including SAGE II, HALOE, UARS-MLS, POAM III, SAGE III, SMR, SCIAMACHY, MIPAS, Aura-MLS, ACE-FTS, ACE-MAESTRO, and SAGE III/ISS and spans the time period 1985 to the end of 2019. The product is provided as zonal means and in units of ppmv (mixing ratios).

2.4 CDR-4: Vertically resolved water vapour profiles in the UTLS

This CDR consists of three-dimensional vertically resolved water vapour data, covering the troposphere and lower stratosphere from 2010 to 2014. The water vapour is based on IMS, MIPAS, and Aura-MLS. The global product includes mixing ratios in units of ppmv.

3. SPECIFICATIONS OF PRIMARY WATER VAPOUR DATA PRODUCTS

3.1 CDR-1

The NIR spectral range provides daytime, cloud-free retrieval of TCWV over land. Since most land surfaces are bright in the NIR, this spectral range is well suited for this purpose. 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 only few absorption features.

The specification of the TCWV product and the required input for the retrieval are described in the following.

3.1.1 Used input data

The input datasets for the generation of the CDR-1 water vapour data products are:

- MERIS L1b reduced resolution (2002–2012), 3rd reprocessing
- MODIS TERRA MOD021KM (2011–2017), collection 6.1
- OLCI L1b reduced resolution (2016–2017), 1st reprocessing.

The MERIS L1b data originates from the latest available data reprocessing version. The data from the 3rd reprocessing of L1b reduced resolution was used for the final production run, as results from the 4th reprocessing were not available in time.The L1b dataset from the OLCI 1st reprocessing is used as input. Both datasets are available at Brockmann Consult.

Data from ERA5 are used as *a priori* input to the retrieval.

3.1.2 Resolution and coverage in space and time

3.1.2.1 TCWV L2 Products

For each L1b input product, a corresponding TCWV L2 product is generated on the same grid. The TCWV retrieval is applied on pixels classified as **land and cloud-free**. The collection of the generated L2 products serves in return as input for the TCWV L3 product generation.

Table 3-1 gives an overview of the product layers in these products. In addition, the TCWV L2 product contains a copy of the tie point grids of the L1b input product.

Table 3-1: Product layers in TCWV L2 products

Name in product	Unit	Type	Description
tcwv	kg/m^2	int16	Total Column Water Vapour
tcwv_uncertainty	kg/m^2	int16	Uncertainty of Total Column Water Vapour
tcwv_quality_flag	dl	int8	Quality flag of Total Column Water Vapour
surface_type_flag	dl	uint16	Pixel classification flag

3.1.2.2 TCWV L3 Products

The final TCWV L3 land products delivered as the CDR-1 dataset are generated from the L2 products by:

- Temporal aggregation (daily averages and monthly averages[1])
- Spatial aggregation (global, WGS84, 0.5 and 0.05 degree resolution).

Table 3-2 provides an overview of the technical specifications.

Table 3-2: CDR-1 - Resolution and coverage in space and time

Dimension	Specification
Temporal resolution	Daily averages and monthly averages
Temporal coverage	July 2002 – December 2017
Spatial coverage	Global
Spatial resolution	0.5 or 0.05 deg

[1] The monthly L3 products were not part of the initially agreed products to be delivered, but were identified as a useful add-on.

3.1.3 Uncertainty estimates, quality indicators, and expected accuracy

The expected quality of CDR-1 is summarised in Table 3-3.

Table 3-3: Expected quality of CDR-1

Quantity	Specification	Notes
Accuracy: Systematic component	1 kg/m^2 (initial estimate)	Bias or systematic difference to other data records, monthly global means, at lower end of results from ESA DUE GlobVapour
Accuracy: Random component	3 kg/m^2 (initial estimate)	cRMSD to other data records, monthly global means, at lower end of results from ESA DUE GlobVapour
Stability	<0.02 (kg/m^2)/year (initial estimate)	This is the observed trend in MERIS time series and a decent estimate of an expected trend can hardly be given (see also Section 4.1). The evaluation of stability on a global scale is challenging

3.1.4 Product layers, auxiliary output, format and metadata

Table 3-4 gives an overview of the product layers in CDR-1. For clarity, the definition of tcwv_err (i.e. propagated uncertainty) is recalled here (see E3UB v2.1 for details):

$$tcwv_{ran} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2}$$

Unphysical values are declared as an undefined or missing value. The uncertainty describes the reliability of the TCWV data. The quality flag can take four values: 0 – no known issues; 1 – cost function above a value of 1; 2 – cost function above 2; 3 – invalid (see Table 3-4).

Table 3-4: Product layers in TCWV L3 product (CDR-1).
NIR: near infrared, MW: microwave

Name in product	Type	Description
tcwv	float32	Total Column Water Vapour
stdv	float32	Standard deviation of Total Column Water Vapour
tcwv_err	float32	Average retrieval uncertainty
tcwv_ran	float32	Propagated retrieval uncertainty

Name in product	Type	Description
tcwv_quality_flag*,**	int8	Quality flag of Total Column Water Vapour 0: no known issues, 1: cost function above 1, 2: cost function above 2, 3: invalid
surface_type_flag	int8	Surface mask 0: land, 1: ocean, 2: clouds_NIR, 3: heavy_precipitation_MW, 4: sea_ice, 5: coast, 6: partly_cloudy over land, 7: partly_sea_ice
num_obs	int16	Number of TCWV L2 retrievals contributing to L3 grid cell
num_days_tcwv***	int16	Number of valid days in monthly CDR (over land without clouds)

* Monthly TCWV data files do not contain the “tcwv_quality_flag”.
**Implementation for CDR-2: 0 = no known issues, 1 = cost function above 2, 2 = "invalid",

***Daily TCWV data files do not contain the “num_days_tcwv”.

An estimate of the file sizes is given in Table 3-5.

Table 3-5: Estimate of final product output data volume (CDR-1)

	month	year	Full period
Size	1.2 GB	14 GB	225 GB

The file format used for storing the data is NetCDF-4 classic. All (NetCDF) files follow the NetCDF Climate and Forecast (CF) Metadata Conventions version 1.7. CF standard names are used for the main variables and global attributes required to ensure compliance with CCI Data Standards have been added. Compliance with the CCI Data Standards was cross-checked and confirmed as the products were being generated.

3.2 CDR-2

The CDR-2 TCWV product combines the land-based CDR-1 product and the ocean-based microwave imager HOAPS product from EUMETSAT CM SAF. Coastal areas and sea-ice regions are filled with NIR-based products. This way, sensor specific advantages have been utilised to produce a global product without changing the characteristics of the individual sensor products. The approach was first proposed for

the ESA GlobVapour project and the resulting product is described in Lindstrot et al. (2014).

3.2.1 Used input data

CDR-1 is used to cover land areas. Additionally, NIR based TCWV data using the CDR-1 retrieval as baseline is used to cover sea-ice and coastal areas. Input data specifications for CDR-1 are described in Section 3.1.

Over ocean TCWV from the CM SAF HOAPS product is used only. The HOAPS product suite is a purely microwave imager-based satellite product (except SST from microwave imagers and a static profile data base from ERA-Interim). HOAPS was originally developed at the MPI-M and the University of Hamburg and has been successfully transferred into the operational environment of EUMETSAT CM SAF. Since version 3.1 CM SAF is generating all HOAPS products which includes among others a TCWV product. This product is defined over the global ice-free ocean with a 50 km distance to nearest coasts. TCWV is retrieved with a 1D-Var scheme under clear-sky and cloudy-sky conditions and can not be reliably applied in presence of strong scattering events such as in strong precipitation cases. CM SAF generated the TCWV data by applying the HOAPS version 4 software using SSM/I, SSMIS, AMSR-E, and TMI data as input. CM SAF generated input to CDR-2 with the following specifications: 0.5° / 0.05° (oversampled) spatial resolution, daily and monthly resolution, global ice-free coverage over the period July 2002 – December 2017.

3.2.2 Resolution and coverage in space and time

The technical specifications for CDR-2 are the same as described in Section 3.1.2. However, in addition to the application of the NIR retrieval over clear-sky land areas, the retrieval is also applied to sea-ice and coastal areas, both in clear-sky conditions.

3.2.3 Uncertainty estimates, quality indicators, and expected accuracy

The expected quality of CDR-2 is summarised in Table 3-6.

Table 3-6: Expected quality of CDR-2

Quantity	Specification	Notes
Accuracy: Systematic component	1 kg/m^2 (initial estimate)	Bias or systematic difference to other data records, monthly global means, at lower end of results from ESA DUE GlobVapour

Quantity	Specification	Notes
Accuracy: Random component	3 kg/m^2 (initial estimate)	cRMSD to other data records, monthly global means, at lower end of results from ESA DUE GlobVapour
Stability	<0.02 (kg/m^2)/year (initial estimate)	This is the observed trend in MERIS time series and a decent estimate of an expected trend can hardly be given (see also section 4.1). The evaluation of stability on a global scale is challenging.

3.2.4 Product layers, auxiliary output, format and metadata

The product layers are the same as for CDR-1 and are recalled for convenience in Table 3-7.

Table 3-7: Product layers in TCWV L3 product (CDR-2).
NIR: near infrared, MW: microwave

Name in product	Type	Description
tcwv	float32	Total Column Water Vapour
stdv	float32	Standard deviation of Total Column Water Vapour
tcwv_err	float32	Average retrieval uncertainty
tcwv_ran	float32	Propagated retrieval uncertainty
tcwv_quality_flag*,**	int8	Quality flag of Total Column Water Vapour 0: no known issues, 1: cost function above 1, 2: cost function above 2, 3: invalid
surface_type_flag	int8	Surface mask 0: land, 1: ocean, 2: clouds over land_NIR, 3: heavy_precipitation_MW, 34: sea_ice, 45: coast, 56: partly_cloudy over land, 67: partly_sea_ice
num_obs	int16	Number of TCWV L2 retrievals contributing to L3 grid cell
num_days_tcwv***	int16	Number of valid days in monthly CDR (over land without clouds)
num_hours****	int16	Number of valid hours in daily CDR (over land without clouds)

* Monthly TCWV estimates do not include the “tcwv_quality_flag”.
**Implementation for CDR-2: 0 = no known issues, 1 = cost function above 2, 3 = "invalid",

***Daily TCWV data files do not contain the “num_days_tcwv”.

Name in product	Type	Description

**** Not available over land, sea-ice and coasts.

An estimate of the file sizes is given in Table 3-8.

Table 3-8: Estimate of final product output data volume (CDR-2)

	month	year	Full period
Size	1.2 GB	14 GB	225 GB

The file format used for storing the data is NetCDF-4 classic. All (NetCDF) files follow the NetCDF Climate and Forecast (CF) Metadata Conventions version 1.7. CF standard names used for the main variables and global attributes required to ensure compliance with CCI Data Standards were added. Compliance with the CCI Data Standards was cross-checked and confirmed as the products were being generated.

3.3 CDR-3

3.3.1 Used input data

CDR-3 used as input data observations from international satellite limb sounders as obtained directly from the different data providers as both single profile data (see Section 3.3.1.1) and zonal monthly mean climatologies processed for the SPARC Data Initiative (see Section 3.3.1.2).

3.3.1.1 L2 (HARMOZ-like) satellite limb sounder profiles

The original profile data have been assembled by the SPARC WAVAS activity for the following satellite limb sounders: ACE-FTS, ACE-MAESTRO, Aura-MLS, HALOE, MIPAS, POAM III, SAGE II, SCIAMACHY, and SMR.

These observations have been converted into a harmonised format, following the conventions of the HARMonized dataset of OZone profiles, HARMOZ (Sofieva et al., 2013). The harmonised dataset consists of original retrieved water vapour profiles (Level 2) from each instrument, which are screened for invalid data according to the recommendations of the instrument teams. The harmonised dataset was processed

into two versions: on a common fixed pressure grid and on a common fixed altitude grid, both in NetCDF-4 format. The pressure grid corresponds to vertical sampling of ~1 km below 20 km and 2–3 km above 20 km. The fixed altitude grid has 1-km steps between 10 and 70 km. The vertical range of the water vapour profiles is specific for each instrument, thus all information contained in the original data is preserved. Geolocation, uncertainty estimates, and vertical resolution are provided for each profile. For each instrument, optional parameters, which are related to the data quality and the profile position, are also included. These water vapour datasets are published by KIT under the name WAVAS_SAHAR (Laeng, 2019).

3.3.1.2 L3 satellite limb sounder zonal mean climatologies

Intermediate instrument time series of L3 zonal monthly mean water vapour fields, which are based on L2 observations that are used for generating the harmonised L2 limb vertical profile WAVAS_SAHAR dataset and which are described in Section 3.3.1.1 were used to produce the final CDR-3. These were provided to WV_cci via the SPARC Data Initiative (SPARC, 2017; Hegglin et al, 2013) and updated to include the latest data versions, now also including climatologies from ACE-MAESTRO and SAGE III/ISS (Hegglin et al., 2020).

The zonal monthly mean time series of water vapour (in volume mixing ratio, VMR) have been calculated for each instrument on the SPARC Data Initiative climatology grid, using 5 degree latitude bins (with midpoints at 87.5°S, 82.5°S, 77.5°S, . . . , 87.5°N) and 28 pressure levels (300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa). To this end, profile data have been carefully screened before binning and a hybrid log–linear interpolation in the vertical has been performed. For instruments that provide data on an altitude grid, a conversion from altitude to pressure levels was performed using retrieved temperature/pressure profiles or meteorological analyses (ECMWF, GEOS-5, or NCEP). Similarly, this information was used to convert retrieved number densities into VMR, where needed. Along with the monthly zonal mean value, the standard deviation and the number of averaged data values are given for each grid point.

Note, considered were a sub-selection of the available instrument retrieval versions in WAVAS_SAHAR that have achieved specified quality targets in the L2 limb vertical profile data Round Robin evaluation and some that are additionally assessed within the SPARC Data Initiative (Hegglin et al., 2013). The final instruments chosen included: SAGE II, SAGE III, HALOE, UARS-MLS, POAM III, SMR, MIPAS (IMK), SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS and SAGE III/ISS.

3.3.2 Resolution and coverage in space and time

The spatial, vertical and temporal resolution and coverage of CDR-3 follows the SPARC Data Initiative convention using a latitude-pressure grid. Table 3-9 lists the associated technical specifications.

Table 3-9: CDR-3 (L3) - Resolution and coverage in space and time

Dimension	Specification	Notes
Temporal resolution	Monthly mean	
Temporal coverage	1985-2019	
Spatial coverage	Zonal mean	Latitude-pressure coordinates
Latitudinal resolution	SPARC Data Initative latitude grid	The 5° latitude bins are centered around: 87.5°S, 82.5°S, 77.5°S, … , 87.5°N
Vertical resolution	SPARC Data Initative pressure grid	Levels: 300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa

3.3.3 Uncertainty estimates, quality indicators, and expected accuracy

The expected quality of CDR-3 is summarised in Table 3-10.

Table 3-10: Expected quality of CDR-3 (L3)

Quantity	Specification	Notes
Accuracy: Systematic component	10–20% (initial estimate)	Bias or systematic difference to other data records such as SWOOSH and GOZCARDs and also a multi-model mean from chemistry climate models
Accuracy: Random component	5–10% (initial estimate)	The cRMSD in the monthly mean values to above data records
Stability	2% per decade	Stability cannot be quanitifed satisfactorily over the full 1985–2019 time period due to a lack of high-quality reference observations. However, this estimate is given under the assumption that the merged product will be produced out of datasets having small drifts (estimated to be 0.3% dec^{-1}) with respect to each other.

3.3.4 Product layers, auxiliary output, format and metadata

Table 3-11 and Table 3-12 compile the product layers, shown here separately for main and additional variables.

Table 3-11: CDR-3 (L3) – Main variable list and description

Variable shortname	unit	Dimensions	Longname/description/comment
lat	degrees_north	N_{lat}	Latitude / predefined latitude bands
plev	hPa	$N_{pressure}$	Pressure / predefined atmospheric pressure grid
time	months since 15 January 1980-01-01	N_{time}	Time
zmh2o	mole mole-1	N_{time} x N_{lat} x $N_{pressure}$	Zonal Mean Water Vapour Volume Mixing Ratio

Table 3-12: CDR-3 (L3) – Additional variable list and description

Variable shortname	unit	Dimensions	Longname/description/comment
zmh2o_uncertainty	%	N_{lat} x $N_{pressure}$ x N_{time}	Uncertainty of Zonal Mean Water Vapour Volume Mixing Ratio
zmh2o_stdv	mole mole-1	N_{lat} x $N_{pressure}$ x N_{time}	Standard Deviation of Zonal Mean Water Vapour Volume Mixing Ratio
quality_flag	N/A	N_{lat} x $N_{pressure}$ x N_{time}	Quality Flag of Zonal Mean Water Vapour Volume Mixing Ratio

An estimate of the file sizes is given in Table 3-13.

Table 3-13: Estimate of final product output data volume (CDR-3)

	month	year	Full period
Size	40 KB	500 KB	17.5 MB

The file format used for storing the data is NetCDF-4 classic. All (NetCDF) files follow the NetCDF Climate and Forecast (CF) Metadata Conventions version 1.7. CF standard names used for the main variables and global attributes required to ensure compliance with CCI Data Standards were added. Compliance with the CCI Data Standards was cross-checked and confirmed as the products were being generated.

3.4 CDR-4

CDR-4 input data consist of L2 vertical profile data from RAL IMS satellite profiles (see product specification in DARD v3.1) and limb sounders (MIPAS and Aura-MLS, see product specification in Section 3.3.1.1).

3.4.1 Used input data

CDR-4 used as input data for the stratosphere observations from international satellite limb sounders as obtained directly from the different data providers (see information in Section 3.3.1.1). However, for this prototype version we focussed only on MIPAS and Aura-MLS. For the troposphere below 300 hPa, we used water vapour retrievals as derived using the RAL IMS scheme, which provides profiles of water vapour from combining measurements of Metop IASI, AMSU and MHS. In the UTLS region between 100 hPa and 300 hPa, a bias-correction, which is based on the climatological VRWV field from balloon-borne hygrometer profiles, is applied to the L2 VRWV profiles from Aura-MLS, MIPAS, and IMS. The final CDR-4 product comprises merged VRWV profiles from 1000 hPa up to 10 hPa (1000, 950, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 225, 200, 175, 150, 125, 100, 70, 50, 30, 10 hPa) with a horizational resolution of 5 degrees and 5 degrees in latitude and longitude.

3.4.2 Resolution and coverage in space and time

The spatial, vertical and temporal resolution and coverage of CDR-4 is similar to the SPARC Data Initiative convention (see Section 3.3.2) but using a latitude–longitude–pressure grid. Table 3-14 lists the associated technical specifications.

Table 3-14: CDR-4 - Resolution and coverage in space and time

Dimension	Specification	Notes
Temporal resolution	Monthly mean	
Temporal extent	2010-2014	Only short time period is provided since this is a prototype CDR
Spatial domain	Global	Latitude–longitude–pressure coordinates
Vertical resolution	Pressure grid	26 levels: 1000, 950, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 225, 200, 175, 150, 125, 100, 70, 50, 30, 10 hPa

Dimension	Specification	Notes
Vertical extent	1000-10 hPa	From the troposphere to the stratosphere , with special focus on the UTLS
Latitudinal resolution	5 degrees	90°S to 90°N
Longitudinal resolution	5 degrees	180°W to 180°E

3.4.3 Uncertainty estimates, quality indicators, and expected accuracy

The expected quality of CDR-4 is summarised in Table 3-15.

Table 3-15: Expected quality of CDR-4

Quantity	Specification	Notes
Accuracy: Systematic component	20–40% (initial estimate)	Bias against other datasets such as reanalyses, FPH sondes. Value is highly uncertain, since information is currently not available. Initial estimate is based on Hegglin et al. (2013)
Accuracy: Random component	10–20% (initial estimate)	cRMSD based on monthly means against other datasets. Value highly uncertain, since information is currently not available. Initial estimate is based on Hegglin et al. (2013)
Stability	2% dec^{-1} (initial estimate)	Value highly uncertain, since information is currently not available. Same assumption used as for CDR-3, see Table 3-10

3.4.4 Product layers, auxiliary output, format and metadata

Table 3-16 and Table 3-17 compile the product layers, shown here separately for main and additional variables.

Table 3-16: CDR-4 (L3) – Main variable list and description

Variable shortname	unit	Dimensions	Longname/description/comment
lat	degrees_north	N_{lat}	Latitude / predefined latitude bands
lon	degrees_east	N_{lon}	Longitude / predefined longitude bands

Variable shortname	unit	Dimensions	Longname/description/comment
plev	hPa	$N_{pressure}$	Pressure / predefined atmospheric pressure grid
time	months since 15 January 1980-01-01	N_{time}	time
vmrh2o	ppmv	N_{lon} x N_{lat} x $N_{pressure}$ x N_{time}	Water Vapour Volume Mixing Ratio

Table 3-17: CDR-4 (L3) – Additional variable list and description

Variable shortname	unit	Dimensions	Longname/description/comment
vmrh2o_stdv	ppmv	N_{lon} x N_{lat} x $N_{pressure}$ x N_{time}	Standard Deviation of Water Vapour Volume Mixing Ratio
vmrh2o_uncertainty	%	N_{lon} x N_{lat} x $N_{pressure}$ x N_{time}	Uncertainty of Water Vapor Volume Mixing Ratio
quality_flag	N/A	N_{lon} x N_{lat} x $N_{pressure}$ x N_{time}	Quality Flag of Water Vapor Volume Mixing Ratio

An estimate of the file sizes is given in Table 3-18.

Table 3-18: Estimate of final product output data volume (CDR-4)

	month	year	Full period
Size	1 MB	10 MB	50 MB

The file format used for storing the data is NetCDF-4 classic. All (NetCDF) files follow the NetCDF Climate and Forecast (CF) Metadata Conventions version 1.7. CF standard names used for the main variables and global attributes required to ensure compliance with CCI Data Standards were added. Compliance with the CCI Data Standards was cross-checked and confirmed as the products were being generated.

4. COMPLIANCE WITH USER REQUIREMENTS

We here discuss the compliance of the different envisaged WV_cci CDRs with the user requirements as derived in the URD (v3.0).

4.1 Level 3 CDR-1 and CDR-2

For the Level 3 CDR-1 and CDR-2 TCWV product, compliance with the user requirements as defined in the URD are judged to be as follows.

Resolution:

These products (with a resolution of 0.05°/0.5° and daily and monthly estimates) aims at fulfilling breakthrough requirements on spatial and temporal resolution. Note that the over ocean TCWV data relies on microwave observations and thus is oversampled for the 0.05° spatial grid. It is further noted that the daily estimates over land rely on one estimate per day at maximum at the equator. This might be uncritical in view of the small diurnal cycle of TCWV which was analysed on basis of GNSS data.

Accuracy:

The systematic (random) component of the uncertainty is estimated to be 1 kg/m^2 (3 kg/m^2) and based on the precursor, MERIS and SSM/I only product. This transforms into 4.5% (14%) when using an average global TCWV value of 22 kg/m^2 as observed by the ESA GlobVapour product in July 2006. Thus, it is expected that the threshold requirement for the random component is within reach and that the breakthrough requirement on the systematic component will likely be achieved. Note that the provided initial estimates are based on intercomparisons using monthly means. Thus, the cRMSD contains differences in sampling and not necessarily the clear-sky sampling bias which is however contributing to the total uncertainty. Further note that the cRMSD of daily data will likely be larger due to larger natural variability.

Stability:

For the precursor version of the ESA GlobVapour project the stability has not been analysed. Thus, the provided value reflects the observed trend of the MERIS based TCWV product. When achieved this would be slightly below target requirement. It is also noted that partly the output from the utilised intruments are merged and combined for the first time and stability issues might still be present.

4.2 Level 2 input data to CDR-3

For Level 2 stratospheric profiles, compliance with the user requirements as defined in the URD (v3.0) are judged to be as follows.

Resolution:

The typical horizontal and vertical resolution of current stratospheric limb sounders (400–500 km and 3–4 km, respectively) are approximately compliant with the threshold requirements (500 km and 3 km, respectively), but do not achieve breakthrough requirements as defined in the URD.

Accuarcy:

A random error of 5% can be achieved for individual instruments, but the instruments are deviating from each other up to ±10%, though in some cases it can go down to ±3%.

Stability:

Results of the WAVAS-II assessment show that the datasets can drift with respect to each other (not with respect to the truth) as small as 0.3% per decade. Hence, if the merged product will be produced out of these datasets only, then the drift of 0.3% would be a realistic expectation. The absolute drift of these datasets is not known yet.

4.3 Level 3 CDR-3

For Level 3 zonal mean stratospheric profiles, compliance with the user requirements as defined in the URD (v3.0) are judged to be as follows.

Resolution:

As follows from the compliance assessement of the L2 data, the expected horizontal and vertical resolution of the L3 CDR-3 (since they are limited to the characteristics of the L2 input data) are compliant with the threshold requirements, but do not achieve breakthrough requirements as defined in the URD. Note that WV_cci will deliver CDR-3 as zonal monthly means and not as three-dimensional fields in order to optimise data quality and to avoid too many data gaps, which while not ideal, still 20% of the stratospheric data users find useful (Level 2 CDR-3 will also be available). The monthly temporal resolution also achieves the threshold requirements.

Accuracy:

The uncertainty in the systematic component is expected to be constrained to between 10 and 20% and hence lies halfway between threshold and breakthrough requirements.

The same is true for the uncertainty in the random component, with the expected 5–10%.

Stability:

The expected stability of 2% dec^{-1} is better than the threshold (2.5% dec^{-1}), but is not quite compliant with the breakthrough requirement (<1% dec^{-1}).

4.4 Level 2 input data to CDR-4

For the Level 3 CDR-4 vertically resolved UTLS product, compliance with the user requirements as defined in the URD are judged as follows.

Resolution:

The vertical resolution of the MLS and MIPAS product is 1–3 km in the UTLS, while IMS is about 3 km in the UTLS, which are compliant with the threshold requirement (0.5–3 km). The horizontal resolution of IMS however is relatively high (25 km), hence meets the breakthrough requirement.

Accuracy:

The uncertainty in the systematic component is estimated to be between 20 and 40%, which would not even meet the threshold requirement at its lower end of the range. However, UTLS specific validation may improve this assessement. For the uncertainty in the random component, an expected 15–30% (or 10–20% after accounting for averaging kernels) is meeting the threshold requirement at its lower end of the range.

Stability:

No measure of stability has yet been quantified for this product.

4.5 Level 3 CDR-4

For the Level 3 CDR-4 vertically resolved UTLS product, compliance with the user requirements as defined in the URD (v3.0) are judged to be as follows.

Resolution:

The L3 CDR-4 product (monthly averages with a resolution of 5°) aims at providing high vertical resolution within the UTLS regions. As follows the compliance assessment on the vertical resolution for l2 input data, the vertical resolution of the L3 CDR-4 (1–2 km) is compliant with the threshold requirement. The horizontal resolution of CDR-4 is set as 5 degrees in order to optimise data quality from limb sounders in stratosphere

and UTLS, which still fulfils threshold requirements. The monthly data also meets the threshold requirements for temporal resolution.

Accuracy:

The uncertainty in the systematic component is estimated to be between 20 and 40%, which would not meet the threshold requirement. For the uncertainty in the random component, an expected 10–20% may just about achieve the threshold requirement.

Stability:

The expected stability of 2% dec^{-1} is expected to be larger than the threshold (1% dec^{-1}).

5. SUMMARY AND CONCLUSIONS

The presented PSD lays out the product specifications for the WV_cci products. The compliance analysis with respect to accuracy was based on initial estimates of the expected product uncertainties. Demonstrating the stability of all WV_cci products has turned out to be a challenge. Overall, it is anticipated that the CDRs envisaged within WV_cci will be compliant with the new threshold and partly with the new target user requirements as defined by the URD.

APPENDIX 1: REFERENCES

DARD v3.1: Data Access Requirement Document (DARD). ESA Water_Vapour_cci, version 3.1, March 2021.

E3UB v2.1: End to End ECV Uncertainty Budget (E3UB) - Part 1: CDR-1 & CDR-2. ESA Water_Vapour_cci, version 2.1, February 2021.

Fischer, J.: High-resolution spectroscopy for remote sensing of physical cloud properties and water vapour, in: *Current problems in atmospheric radiation*, edited by: Lenoble, J. and Geleyn, J.-F., Deepak Publishing, Hampton, Virginia, USA, 151–156, 1988.

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 Applied Meteorology*, **32**(12), 1791-1801.

GCOS-200: GCOS Implementation Plan 2016. GCOS-200. Available at https://library.wmo.int/opac/doc_num.php?explnum_id=3417.

Hegglin, M. I., et al. (2013), SPARC Data Initiative: Comparison of water vapor climatologies from international satellite limb sounders, *J. Geophys. Res. Atmos*., **118**, 11,824–11,846, doi:10.1002/jgrd.50752.

Hegglin, M. I., Tegtmeier, S., Anderson, J., Bourassa, A. E., Brohede, S., Degenstein, D., Froidevaux, L., Funke, B., Gille, J., Kasai, Y., Kyrölä, E., Lumpe, J., Murtagh, D., Neu, J. L., Pérot, K., Remsberg, E., Rozanov, A., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., Wang, H.-J., Arosio, C., Damadeo, R., Fuller, R., Lingenfelser, G., McLinden, C., Pendlebury, D., Roth, C., Ryan, N. J., Sioris, C., Smith, L., and Weigel, K.: Overview and update of the SPARC Data Initiative: Comparison of stratospheric composition measurements from satellite limb sounders, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2020-342, in review, 2020.

Laeng, A., Water vapour profiles from WAVAS Satellite component in HARmonized format (WAVAS_SAHAR) – Version 2, KIT-Bibiliothek, doi:10.5445/IR/1000098908, 2019. (available at https://bwdatadiss.kit.edu/dataset/192)

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, https://doi.org/10.5194/amt-5-631-2012, 2012.

Lindstrot, R., M. Stengel, M. Schröder, J. Fischer, R. Preusker, N. Schneider, T. Steenbergen, and B. Bojkov, 2014: A global climatology of total columnar water vapour from SSM/I and MERIS. *Earth Syst. Sci. Data*, **6**, 221–233, 2014, www.earth-syst-sci-data.net/6/221/2014/, doi:10.5194/essd-6-221-2014.

PVP, 2021: Product Validation Plan. ESA Water_Vapour_cci, version 3.1, March 2021.

Sofieva, V. F., Rahpoe, N., Tamminen, J., et al., Harmonized dataset of ozone profiles from satellite limb and occultation measurements, *Earth Syst. Sci. Data*, **5**, 349-363, doi:10.5194/essd-5-349-2013, 2013.

SPARC, 2017: The SPARC Data Initiative: Assessment of stratospheric trace gas and aerosol climatologies from satellite limb sounders. By M. I. Hegglin and S. Tegtmeier (eds.), SPARC Report No. 8, WCRP-5/2017, https://doi.org/10.3929/ethza-010863911.

URD v3.0: User Requirements Document (URD). ESA Water_Vapour_cci, version 2.0, February 2021.

APPENDIX 2: GLOSSARY

This appendix explains all utilised abbreviations of this document.

Term	Definition
ABC(t)	Atmosphere Biosphere Climat (teledetection)
ACE-FTS	Atmospheric Chemistry Experiment Fourier Transform Spectrometer
ACE-MAESTRO	Atmospheric Chemistry Experiment Measurements of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
AMSU	Advanced Microwave Sounding Unit
ARA	Atmospheric Radiation Analysis
ARSA	Analyzed RadioSoundings Archive
AVHRR	Advanced Very High Resolution Radiometer
BC	Brockmann Consult
CARIBIC	Civil Aircraft for the Regular Investigation of the atmosphere Based on an Instrument Container
CCI	Climate Change Initiative
CDO	Climate Data Operators
CDR	Climate Data Record
CDS	Copernicus Climate Data Store
CEDA	Centre for Environmental Data Analysis
CF	Conventions for Climate and Forecast
CM SAF	EUMETSAT Satellite Application Facility on Climate Monitoring
CMAM	Canadian Middle Atmosphere Model
CMIP	Coupled Model intercomparison Project
CMUG	Climate Modelling User Group
CRG	Climate Research Group
DLR	Deutschen Zentrums für Luft- und Raumfahrt
DWD	Deutscher Wetterdienst (German MetService)
ECCC	Environment and Climate Change Canada
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
EDA	ERA5 - reduced resolution ten member ensemble
EMiR	ERS/Envisat MWR Recalibration and Water Vapour Thematic Data Record Generation

Term	Definition
Envisat	Environmental Satellite
ERA5	ECMWF Re-Analysis 5
ERA-Interim	ECMWF Re-Analysis Interim
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GCOS	Global Climate Observing System
GEOS-5	Goddard Earth Observing System Model, Version 5
GMI	Global Precipitation Microwave Imager
GNSS	Global Navigation Satellite System
GOMOS	Global Ozone Monitoring by Occultation of Stars
GOZCARDS	Global OZone Chemistry And Related trace gas Data records for the Stratosphere
GPS	Global Positioning System
GRUAN	GCOS Reference Upper-Air Network
HARMOZ	HARMonized dataset of Ozone profiles
HALOE	Halogen Occultation Experiment
HIRDLS	High Resolution Dynamics Limb Sounder
HOAPS	Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data
IAGOS	In-service Aircraft for a Global Observing System
IASI	Infrared Atmospheric Sounder Interferometer
ILAS-II	Improved Limb Atmospheric Spectrometer-II
IMS	Infrared Microwave Sounding
IPSL-CM	Institut Pierre Simon Laplace Climate Model
IR	Infrared
LMD	Laboratoire Météorologie Dynamique
LMS	Lowermost stratosphere
LST	Land Surface Temperature
LWP	Vertically integrated liquid water
MERIS	Medium Resolution Imaging Spectrometer Instrument
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
MHS	Microwave Humidity Sounder
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder

Term	Definition
MODIS	Moderate Resolution Imaging Spectrometer
MOZAIC	Measurement of OZone by Airbus In-service airCraft
MPI-M	Max-Planck Institute for Meteorology
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEO	National Centre for Earth Observation
NCEP	National Centers for Atmospheric Prediction
NDVI	Normalized Difference Vegetation Index
NIR	Near IR
NOAA	National Oceanic & Atmospheric Administration
NWP	Numerical Weather Prediction
OLCI	Ocean and Land Colour Instrument
PCs	Principle components
POAM	Polar Ozone and Aerosol Measurement
PSD	Product Specification Document
RAL	Rutherford Appleton Laboratory
RMS	Root mean square
RR	Reduced resolution
RTTOV	Radiative Transfer for TOVS
SAGE	Stratospheric Aerosol and Gas Experiment
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SCISAT	Scientific Satellite
SE	Spectral Earth
SMILES	Solar wind Magnetosphere Ionosphere Link Explorer
SMR	Software Modification Report
SNR	Signal-to-noise ratio
SOFIE	Solar Occultation For Ice Experiment
SPARC	Stratosphere-troposphere Processes And their Role in Climate
SPURT	Spurenstofftransport in der Tropopausenregion, trace gas transport in the tropopause region
SSM/I	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager Sounder
SST	Sea Surface Temperature

Term	Definition
SuomiNet	Global ground based GPS network (named after Verner Suomi)
SWOOSH	Stratospheric Water and OzOne Satellite Homogenized data set
TBD	To be determined
TCWV	Total Column Water Vapour
TMI	Tropical Rainfall Measuring Mission's Microwave Imager
TOA	Top Of Atmosphere
UKMO	United Kingdom Meteorological Office
UoL	University of Leicester
UoR	University of Reading
URD	User Requirements Document
UT	Upper troposphere
UTLS	Upper Troposphere and Lower Stratosphere
UV	Ultraviolet
vis	Visible
VMR	Volume mixing ratio
VRes	Vertically resolved
WACCM	Whole Atmosphere Community Climate Model
WAVAS-I	Water Vapour Assessment
WAVAS-II	Water Vapour Assessment 2
WCRP	World Climate Research Programme
WGS	World Geodetic System 1984
WMO	World Meteorological Organization
WV	Water Vapour
WV_cci	Water Vapour climate change initiative

End of Document