

# Water Vapour Climate Change Initiative (WV\_cci) - CCI+ Phase 1



System Requirements Document (SRD)

Ref: D3.1

Date: 1 October 2020

Issue: 2.1

For: ESA / ECSAT

Ref: CCIWV.REP.012



a LEONARDO and THALES company



aeronomie.be



UNIVERSITY OF  
TORONTO



UNIVERSITY OF  
LEICESTER

UNIVERSITÉ DE  
VERSAILLES  
SAINT-QUENTIN-EN-YVELINES



Science & Technology Facilities Council  
Rutherford Appleton Laboratory

Universidade de Vigo

***This Page is Intentionally Blank***

**Project** : **Water Vapour Climate Change Initiative (WV\_cci) - CCI+ Phase 1**

**Document Title:** **System Requirements Document (SRD)**

**Reference** : **D3.1**

**Issued** : **1 October 2020**

**Issue** : **2.1**

**Client** : **ESA / ECSAT**

**Authors** : Olaf Danne (Brockmann Consult), Michaela Hegglin and Hao Ye (University of Reading)

**Copyright** : Water\_Vapour\_cci Consortium, ESA.

## Document Change Log

<b>Issue/ Revision</b>	<b>Date</b>	<b>Comment</b>
0.1	23 Apr 2019	First Draft
0.9	3 June 2019	First version for internal review
1.0	7 June 2019	First submitted version
1.1	31 August 2019	Revision based on v1.0 RIDs
2.0	31 July 2020	Scheduled document update
2.1	1 October 2020	Revision based on v2.0 RIDs

# TABLE OF CONTENTS

<b>1. INTRODUCTION .....</b>	<b>8</b>
1.1 Purpose and Scope .....	8
<b>2. WV_CCI SYSTEM OVERVIEW.....</b>	<b>9</b>
2.1 WV Data Products .....	9
2.2 WV System Context.....	11
2.3 User Requirements .....	11
2.4 Assumptions on Common CCI System Requirements .....	14
2.5 High Level System Requirements .....	15
<b>3. WV_CCI SCENARIOS .....</b>	<b>18</b>
3.1 TCWV Processing System (BC) .....	18
3.1.1 Sensor constellations and temporal coverage .....	18
3.1.2 Products and variables .....	19
3.1.3 Reprocessing capability.....	19
3.1.4 Improvement cycle .....	20
3.1.5 Validation and interaction with users .....	21
3.2 VRWV Processing System (UoR).....	21
3.2.1 Sensor constellations and temporal coverage .....	21
3.2.1.1 CDR-3.....	21
3.2.1.2 CDR-4.....	22
3.2.2 Products and variables .....	23
3.2.3 Reprocessing capability.....	23
3.2.3.1 CDR-3.....	23
3.2.3.2 CDR-4.....	23
3.2.4 Improvement cycle .....	24
3.2.4.1 CDR-3.....	24
3.2.4.2 CDR-4.....	24
3.2.5 Validation and interaction with users .....	25
<b>4. WV_CCI PRODUCTION WORKFLOW .....</b>	<b>26</b>
4.1 TCWV Processing System (BC) .....	26
4.1.1 Processing chain and processors .....	26
4.1.2 Main inputs and outputs .....	26
4.1.3 Estimation of system key parameters .....	27
4.2 VRWV Processing System (UoR).....	28
4.2.1 Processing chain and processors .....	28
4.2.2 Main inputs and outputs .....	28
4.2.3 Estimation of system key parameters .....	30
4.2.3.1 CDR-3.....	30
4.2.3.2 CDR-4.....	31
<b>5. DETAILED SYSTEM REQUIREMENTS.....</b>	<b>32</b>
5.1 Functional Requirements .....	34
5.1.1 Preprocessing.....	35
5.1.2 WV processing .....	36
5.1.3 Validation .....	37
5.1.4 Data Management .....	38
5.1.5 User Interaction .....	38
5.2 Operational Requirements .....	39

5.2.1	Preprocessing.....	40
5.2.2	WV processing .....	40
5.2.3	Data Management .....	41
5.2.4	User Interaction .....	41
5.3	Performance and Sizing Requirements .....	41
5.3.1	Preprocessing.....	42
5.3.2	WV processing .....	42
5.3.3	Data Management .....	42
5.3.4	User Interaction .....	42
5.4	Interface Requirements.....	42
5.4.1	Preprocessing.....	43
5.4.2	WV processing .....	44
5.4.3	Data Management .....	44
5.4.4	User Interaction .....	45
5.5	Reliability, Availability and Maintainability Requirements .....	45
5.5.1	Preprocessing.....	46
5.5.2	WV processing .....	46
5.5.3	Data Management .....	46
5.5.4	User Interaction .....	46
<b>6.</b>	<b>REQUIREMENTS TRACEABILITY .....</b>	<b>47</b>
	<b>APPENDIX 1: REFERENCES.....</b>	<b>54</b>
	<b>APPENDIX 2: GLOSSARY .....</b>	<b>55</b>

## INDEX OF TABLES

Table 4-1:	Main inputs of the TCWV processing chain .....	26
Table 4-2:	Overall data volume and estimation of processing time.....	28
Table 4-3:	Main inputs of the VRWV CDR-3 processing chain .....	28
Table 4-4:	Main inputs of the VRWV CDR-4 processing chain .....	29
Table 4-5:	Overall data volume and estimation of processing time for CDR-3 .....	30
Table 4-6:	Overall data volume and estimation of processing time for CDR-4 .....	31

## INDEX OF FIGURES

Figure 2-1:	MERIS/MODIS/CM SAF HOAPS TCWV L3 daily merge for July 15 <sup>th</sup> , 2011 (greyscale, 0-70 kgm <sup>-2</sup> ). Yellow indicates no data. Taken from [9]. .....	9
-------------	---	---

Figure 2-2: Zonal monthly mean cross-section of VRWV from CDR-3 for July 2008. The cross-section reveals dehydration in the lower stratosphere polar vortex region during southern hemisphere wintertime.....10

Figure 2-3: Mean VRWV from CDR-4 in 2010-2012 at pressure levels (from top to bottom: 100 hPa, 150 hPa, and 200 hPa). .....10

Figure 5-1: Detailed system requirements and their hierarchy, starting from the high-level requirements given in section 2.5. The numbers refer to the requirements defined further in this section. The empty boxes (...) are further refined in Figure 5-2 for TCWV and Figure 5-3 for VRWV.....32

Figure 5-2: Detailed system requirements for the TCWV processing system part.....33

Figure 5-3: Detailed system requirements for the VRWV processing system part. ....33

# 1. INTRODUCTION

## 1.1 Purpose and Scope

As requested by the CCI general Statement of Work [1]:

*"Create a System Requirements Document (SRD) specifying the requirements of a Processing System capable of generating ECV data products as specified in the applicable technical annex (A-I). The SRD shall include verifiable requirements on the following: Data processing function of each step of its processing chain, including data volumes; Platform specification; Compliance to all processing needs defined by the Task 3 inputs."*

this SRD defines a set of requirements for the Water Vapour (WV) processing system for the European Space Agency (ESA) Climate Change Initiative (CCI). WV is one of the Essential Climate Variables (ECV) studied by CCI [2].

From the statement above it follows that the SRD needs to be a complete, structured collection of individual requirements of the ECV production system. The other inputs are deliverables of the WV\_cci project itself. They request for a system:

- to generate the product specified in the Product Specification Document [3] using data specified in the Data Access Requirement Document [4];
- to apply the algorithms specified in the Algorithm Theoretical Basis Document [5] for this purpose;
- to make the product accessible to users as specified in the User Requirement Document [6].

The approach used for this document is to develop a typical scenario for the WV retrieval with minimal and desirable variants. Together with the description of the processing workflow, this serves as an additional starting point for the derivation of requirements. As far as possible, the requirements are traced to inputs, scenarios and the processing workflow.

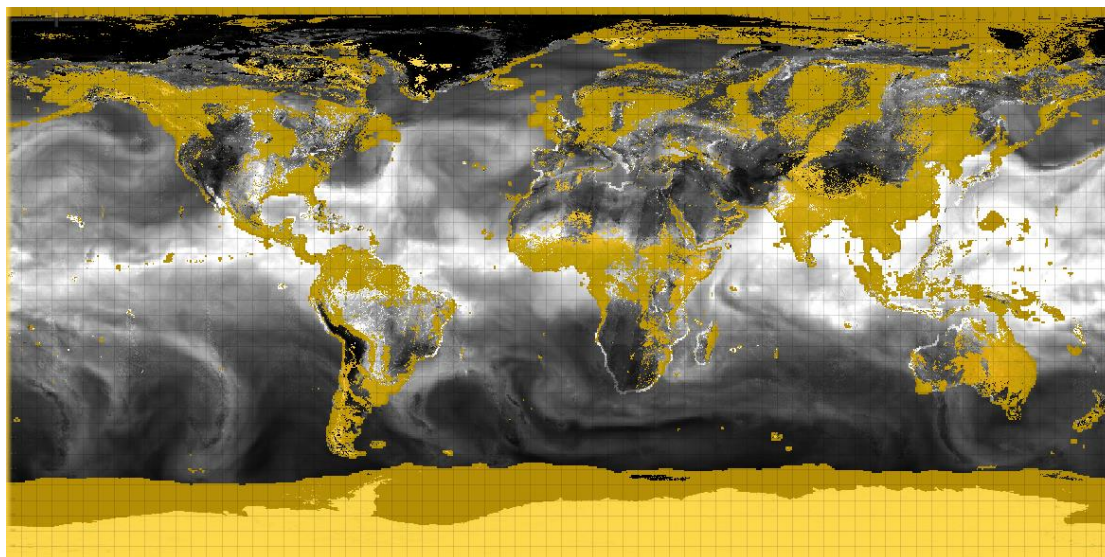
The defined WV\_cci system requirements serve as input to the WV\_cci system specification as given in the System Specification Document [7].



## 2. WV\_CCI SYSTEM OVERVIEW

### 2.1 WV Data Products

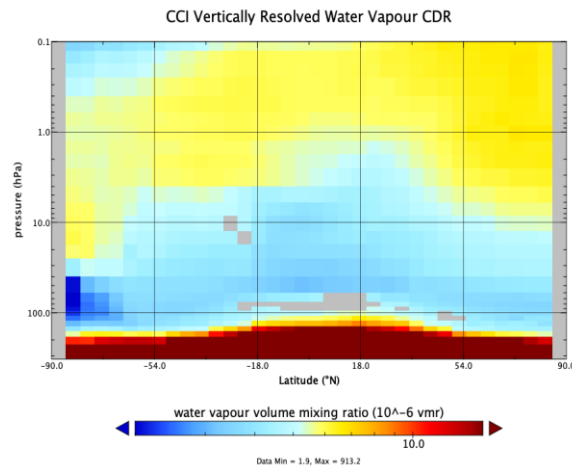
The Water Vapour CCI (WV\_cci) project focuses on the Water Vapour ECV comprising total column of water vapour (TCWV) and vertically resolved water vapour (VRWV). As input for TCWV retrieval, the project will use data from ESA's MERIS and OLCI instruments, and from NASA's MODIS Terra. The derived TCWV will be complemented by CM SAF HOAPS TCWV data over ice-free ocean, mainly derived from SSM/I and SSMIS microwave data. For VRWV, input data for the stratospheric zonal monthly mean climate data record (CDR-3) consist of data obtained from the satellite limb sounders SAGE II, HALOE, UARS-MLS, POAM III, SAGE III, SMR, SCIAMACHY, MIPAS, Aura MLS, ACE-FTS, ACE-MAESTRO, and SAGE III/ISS. For the three-dimensional prototype version of the upper tropospheric/lower stratospheric (UTLS) climate data record (CDR-4), input data consist of observations from the satellite limb sounders MIPAS, Aura-MLS, ACE-MAESTRO, ACE-FTS, and a combined retrieval product (IMS) from the IASI/MHS/AMSU satellite instruments. Figure 2-1 shows an example of global daily TCWV, and Figure 2-2 and Figure 2-3 show the vertical cross-section and horizontal distribution of VRWV from CDR-3 and CDR-4, respectively.



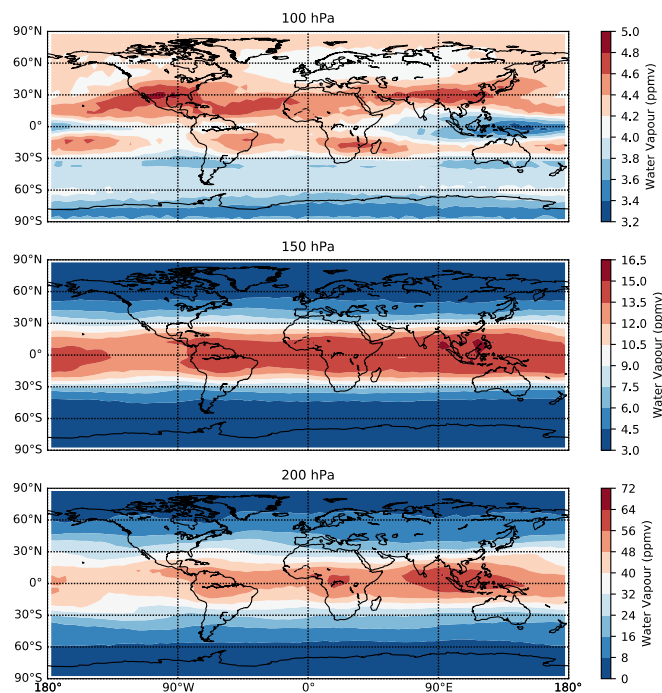
**Figure 2-1: MERIS/MODIS/CM SAF HOAPS TCWV L3 daily merge for July 15<sup>th</sup>, 2011 (greyscale, 0-70 kgm<sup>-2</sup>). Yellow indicates no data. Taken from [9].**

As stated in the WV\_cci Statement of Work [8], a main goal of the project is to provide a homogeneous and stable TCWV data record with stability on global scale close to GCOS requirements. An essential step is the merging of the TCWV data of the

available sensors in order to fill all data gaps in the global products as far as possible, and to identify regional, temporal, or systematic differences and discontinuities when bringing the data together. This should ideally result in an improvement of the underlying algorithms and the elimination of existing problems.



**Figure 2-2: Zonal monthly mean cross-section of VRWV from CDR-3 for July 2008. The cross-section reveals dehydration in the lower stratosphere polar vortex region during southern hemisphere wintertime.**



**Figure 2-3: Mean VRWV from CDR-4 in 2010-2012 at pressure levels (from top to bottom: 100 hPa, 150 hPa, and 200 hPa).**

As for the TCWV and also stated in the WV\_cci Statement of Work [8], another main goal is to produce a well characterized and homogeneous long-term VRWV data record in the stratosphere. A key aspect in the construction of such a CDR is the requirement

for a homogenization procedure to account for systematic measurement biases, which can vary in time and are often altitude- and latitude-dependent. Employment of appropriate merging procedures (such as the use of a model as transfer function between different satellite data records) is expected to improve our knowledge of the spatio-temporal biases in the observations and their uncertainties, ultimately guiding future efforts to improve these data records. In addition, as stated in [TR-5] of the WV\_cci Statement of Work [8], a prototype of a harmonised limb/nadir data record of water vapour in the UTLS is required. To achieve this goal, better knowledge of the consistency between water vapour products from limb-sounders and nadir instruments in this region will have to be gained.

## 2.2 WV System Context

The WV\_cci system interacts with various other entities as described in detail in the SSD [7].

## 2.3 User Requirements

The input requirements for this document stem from:

- the CCI general Statement of Work [1]
- the WV\_cci Statement of Work [8] (annex 1 of [1])
- the WV\_cci project documents URD [6], PSD [3], DARD [4], ATBD [5]
- CCI Project Guidelines [11]

From [1], there are the following requirements

- SoW-R-1 ("Addressing one of 9 ECVs")
- SoW-R-2 ("Execution within 36 months")
- SoW-R-3 ("Meeting GCOS requirements for particular ECV")
- SoW-R-4 ("Taking advantage of existing results from European programs")
- SoW-R-5 ("Taking advantage of existing results from non-European programs ")
- SoW-R-6 ("Continuity of action")
- SoW-R-7 ("Legacy of the CCI in the project")
- SoW-R-8 ("Assess documents, systems and data from CCI in CCI+")
- SoW-R-9 ("Take account of key technical constraints")

- SoW-R-10 ("Engage key international scientists")
- SoW-R-11 ("Participate in Climate Science Working Group")
- SoW-R-12 ("Cooperate with the CMUG team")
- SoW-R-13 ("Liaise with the Knowledge Exchange team")
- SoW-R-14 ("Publication of results in peer-review journals")
- SoW-R-15 ("Execute all tasks as specified in Phase 1")
- SoW-R-16 ("Integration of data from Copernicus satellites")
- SoW-R-17 ("Dimensions of the processing system"),

together with a list of cardinal requirements, which are tailored in the Annexes of [1] for the various ECVs (here Annex A [8] for Water Vapour ECV).

From [8], further requirements can be defined as:

- SoW\_WV-1 on GCOS requirements
- SoW\_WV-2 on user engagement
- SoW\_WV-3 on product framework
- SoW\_WV-4 on key issues.

The URD defines requirements mainly on the product to be produced and on tools:

- UR-1 on products/variables
- UR-2 on spatial and temporal resolution
- UR-3 on accuracy and precision
- UR-4 on error specification
- UR-5 on product formats and access.

The first four of them can be translated into the single requirement for the system to generate products as specified in the PSD [3]. UR-5 requests for outputs that conform with the CCI Data Standards (see <http://cci.esa.int/working-groups>), and for data access e.g. via the CCI Open Data Portal ([cci.esa.int/data](http://cci.esa.int/data)).

The PSD [3] defines content and format of the WV products. These are products of TCWV and VRWV. They are referred to as:

- PSD-1 (TCWV CDR-1)
- PSD-2 (TCWV CDR-2)

- PSD-3 (VRWV CDR-3)
- PSD-4 (VRWV CDR-4)

in this document. Like the URD requirements they translate into requirements to generate these products.

The CCI Data Engineering Working Group is also frequently generating/updating requirements that apply to all ECVs. They include detailed requirements on metadata, both usage (CF-compliance) and discovery (Unidata metadata discovery standard), and on versioning and traceability. These requirements can be expected to be considered in the PSD for the given ECV.

The DARD [4] specifies the set of input products used for WV\_cci. They comprise the main sensor data inputs, further ancillary data for processing, and various validation datasets:

- DARD-1 on MERIS RR level 1, 4<sup>th</sup> reprocessing, 2002-2011
- DARD-2 on MODIS Terra (MOD021KM), collection 6, 2011-2017
- DARD-3 on OLCI RR level 1, 1<sup>st</sup> reprocessing, 2016-2017
- DARD-4 on CM SAF HOAPS TCWV
- DARD-5 on WAVAS-II
- DARD-6 on ACE-FTS
- DARD-7 on ACE-MAESTRO
- DARD-8 on SPARC Data Initiative
- DARD-9 on IMS
- DARD-10 on AURA-MLS
- DARD-11 on MIPAS
- DARD-12 on auxiliary data for TCWV retrieval
- DARD-13 on auxiliary data for VRWV retrieval
- DARD-14 on EMiR
- DARD-15 on GOME-Evolution Climate
- DARD-16 on GOZCARDS (NASA)
- DARD-17 on IASI

- DARD-18 on ARSA
- DARD-19 on CARIBIC/MOZAIC/IAGOS
- DARD-20 on GNSS
- DARD-21 on GRUAN
- DARD-22 on SPURT aircraft data
- DARD-23 on SuomiNet
- DARD-24 on LST from ESA CCI
- DARD-25 on SST and TCWV from ESA CCI.

This translates into requirements for the system to maintain this data, to protect restricted data against unauthorised access, and to provide it as input to the production. The DARD further specifies the volume of input data. This translates into sizing requirements.

The ATBD [5] identifies algorithms to be used in the generation of the WV products. Together with the processing workflow described in the SSD [7], this translates into requirements to integrate processors for these algorithms. WV\_cci intends to re-evaluate the selected algorithms following the generation of temporal subsets of the full dataset, which is an illustration of the system requirement to support reprocessing and a process for evaluating and changing algorithms.

## 2.4 Assumptions on Common CCI System Requirements

One important question for the design of the WV\_cci system is the sharing of infrastructure, resources or functions with other ECVs, or their independence [10]. This question has been addressed in earlier years of the CCI in the SRDs of e.g. OC and SST ([13], [14]), resulting in a basic approach which overall still applies also for WV\_cci. The main features of this approach are:

- There is a common data access layer shared among all ECVs which provides harmonised access to CCI data for climate modellers. This shall lower the barrier for climate users to use several of the ECVs.
- The individual production and data environments per ECV are close to the scientific groups to support an agile, continuous development and nimble reaction to issues with short cycles. The production environments are optimised for re-processing and validation. Strict versioning ensures production of stable product releases. Optionally, sharing of an environment by production and development allows for access to long time series also for the scientific improvement cycle.

The common data access layer has been realized in the meantime by the CCI open data portal [15]. The concept of agile and continuous development with the generation and reprocessing of defined data (sub)sets is explicitly foreseen in WV\_cci and outlined in the relevant work packages.

This approach also leads to a few more 'abstract requirements (also taken from [13], called 'assumptions' there):

- Assumption-1: Climate users shall get harmonised access to different ECV outputs with common data formats and a single point of access;
- Assumption-2: EO science teams shall be part of the system and shall improve algorithms, react on issues raised by users of the outputs, extend the process to new inputs;
- Assumption-3: Product versions shall co-exist. A stable version shall be accessible by users and shall be extended to the future by new data. Older versions will be available for the long term (without this, publications based on previous processing will not be repeatable nor can the impact of reprocessing be evaluated);
- Assumption-4: The different ECV production systems shall be autonomous and independent of each other, and founded in the relevant communities;
- Assumption-5: The ECV production systems shall be optimised for repeated re-processing and validation. This should not preclude their use in near real time processing, but the focus shall be on quality.

## 2.5 High Level System Requirements

This section anticipates the definition of requirements and identifies a few high-level requirements for the WV system.

As a starting point, WV-SR-0010 is the top of a hierarchy of requirements with WV-SR-0100 to -0400 as next layers. These requirements will be broken down further in section 5.

<b>WV-SR-0010</b>	<b>Generate and disseminate WV products</b>
<p>The WV_cci system shall generate, publish and disseminate water vapour products. It shall be responsive allowing for evolution with the ability to re-generate the</p>	

products repeatedly with improved inputs or improved generation methods within the frame of the project.

<b>WV-SR-0100</b>	<b>Implement processing workflows and scenarios</b>
<p>The WV_cci system shall implement the processing workflows and the scenarios for WV_cci.</p>	

<b>WV-SR-0200</b>	<b>Output products and stability</b>
<p>The WV_cci system shall generate the outputs in a stable way as required in the URD and as specified in the PSD.</p>	
<p>Note: This requirement defines what shall be produced. Its source is the URD and the PSD. It also focuses on stability which is one aspect of Assumption-3 from section 2.4. The ATBD in combination with the description of the processing workflow serves as source how the products shall be generated.</p>	

<b>WV-SR-0300</b>	<b>Data dissemination</b>
<p>The WV_cci system shall interact with climate modelers and WV community users and shall inform about WV products, provide access to the data, as subsets or in bulk, and facilitate feedback on products and issues.</p>	
<p>Note: Basically, the data dissemination shall be provided by the Open Data Portal once the data has been generated by the WV_cci system.</p>	

<b>WV-SR-0400</b>	<b>System evolution</b>
<p>The WV_cci system shall be nimble with respect to the improvement of algorithms, i.e. bringing new algorithm versions into production. Separate and frequent re-processing shall be possible for development and testing. Suitable automatic procedures shall be in place to perform validation and other steps to enable the transfer to production.</p>	



Note: This requirement is in agreement with assumption-2 and assumption-5.

Before detailing these high-level requirements, the following two sections describe scenarios and production workflows as an additional basis for the breakdown.

### 3. WV\_CCI SCENARIOS

This section identifies possible WV\_cci production scenarios for CCI+ and optionally beyond. Aspects are the sensors considered, the products and variables generated, interaction between ECVs and with users, and features to improve the products and the process. For each aspect up to three variants of a minimum, an intermediate and an ideal scenario are considered.

#### 3.1 TCWV Processing System (BC)

##### 3.1.1 Sensor constellations and temporal coverage

The minimum TCWV production scenario covers as temporal period the three years 2010–2012 including the year 2011 with available input data from both MERIS and MODIS Terra (TCWV retrieval over land and gap filling over oceans) and CM SAF HOAPS (TCWV over ice-free oceans from SSM/I, SSMIS). This scenario represents the generation of 'Dataset 1' as specified in work package 4 of the project.

The intermediate scenario covers the processing period 2002–2017 with all available input data from MERIS, MODIS Terra, OLCI, and CM SAF HOAPS. This scenario represents the generation of the 'full' dataset as specified in work package 4 of the project.

The maximum scenario covers the same as the intermediate scenario, but in addition uses input data from the MODIS AQUA instrument from the period 2002–2017. This scenario is considered as 'option data set' for WV\_cci. Its realisation will be decided at a later date.

Minimum	MERIS + MODIS Terra + CM SAF HOAPS, 2010-2012
Intermediate	MERIS + MODIS Terra + OLCI + CM SAF HOAPS, 2002-2017
Maximum	Same as intermediate, plus MODIS Aqua 2002-2017

The impact of the variants to the requirements is on the functional and the performance level with respect to input data ingestion, pre-processing, TCWV retrieval, data transfer and storage capacity for intermediate and final products.

### 3.1.2 Products and variables

The minimum is what has been generated as 'Dataset 1'. The other variants extend this by an improved representation of uncertainty variables and quality flags in the products, resulting from a more comprehensive consideration in subsequent versions of the TCWV retrieval algorithm and corresponding ATBD.

Minimum	TCWV products contain a layer for uncertainty, but filled with dummy values only. No data quality flag is included
Intermediate	Uncertainty layer is filled as retrieved from improved TCWV retrieval algorithm
Maximum	Same as intermediate, but full data quality flag is provided, following definition in further extended TCWV retrieval algorithm and ATBD

The variants should not have significant impact on performance. The quality flag can be added to the products as byte or short integer variable and will only slightly increase storage demand.

### 3.1.3 Reprocessing capability

The capability to re-process TCWV datasets or subsets is essential for quality improvement. The variants below represent different levels of performance of the whole processing chain.

Minimum	Reprocessing of 1 year TCWV in 2 months
---------	---

Intermediate	Reprocessing of 1 year TCWV in 1 month
Maximum	Reprocessing of 1 year TCWV in 0.5 months

The impact of the variants on requirements is mainly on the performance level. Technically, the set of variants call for scalability to allow going from one variant to the other. As learned from the 'Dataset 1' generation, the maximum variant can obviously just be achieved with the given resources as described in the SSD. The minimum and intermediate variants would provide some more spare time to identify and address possible issues.

### 3.1.4 Improvement cycle

Improvements might result from a better TCWV algorithm, better pre-processing (e.g. cloud classification), and by eliminating specific problems (e.g. discontinuities) when merging TCWV retrieved from the different sensors.

Minimum	Generation of 'Dataset 1' as specified in work package 4 of the project
Intermediate	Generation of 'Dataset 2' as specified in work package 4 of the project, which includes a reprocessing of 'Dataset 1' with improved TCWV algorithm
Maximum	Generation of the full dataset as specified in work package 4 of the project, which will likely include further algorithm improvements and elimination of problems

The impacts of the variants on the requirements are direct and indirect. Assuming that the currently known options for improvements are final, the system shall directly support their development. Assuming that there will be more improvements over time, the system shall indirectly support the listed improvements by supporting an improvement cycle.

### 3.1.5 Validation and interaction with users

The minimum corresponds to what has been delivered with 'Dataset 1'. Additional features are improved and more convenient means for data access/exchange, and communication means towards the user and from users back to the science team.

Minimum	Delivery of standard L3 products via CCI Open Data Portal, interaction with key users
Intermediate	Same as minimum, but delivery of additional products (e.g. TCWV L2) on demand. Interaction with a growing number of users
Maximum	Final data access highly performant and possible from both FTP and via CCI Open Data Portal. Provide feedback interface for strengths and weaknesses of the products, for climate modellers and WV community

The impact of the variants to requirements is on the functional and interface level. The maximum variant may be translated into e.g. a customer relationship management with traceable communication, issue tracking, and improvements for users to simplify their use of WV products.

## 3.2 VRWV Processing System (UoR)

### 3.2.1 Sensor constellations and temporal coverage

#### 3.2.1.1 CDR-3

The minimum and maximum VRWV CDR-3 production scenarios distinguish themselves in the temporal coverage, with 34 (1985–2018) and 35 years (1985–2019), respectively. Input data for both are from SAGE II, HALOE, UARS-MLS, POAM III,

SAGE III, SMR, SCIAMACHY, MIPAS, Aura-MLS, ACE-FTS, and ACE-MAESTRO. SAGE III-ISS is considered in the maximum variant. The minimum and maximum scenarios represent the generation of a 'Prototype dataset' and a 'Final dataset' as specified in work package 4 of the project to be delivered after year 2 and year 3 of the WV\_cci project, respectively.

Minimum	SAGE II, HALOE, UARS-MLS, POAM III, SAGE III, SMR, SCIAMACHY, MIPAS, Aura-MLS, ACE-FTS, and ACE-MAESTRO covering different time periods between 1985 and 2018
Maximum	SAGE II, HALOE, UARS-MLS, POAM III, SAGE III, SMR, SCIAMACHY, MIPAS, Aura-MLS, ACE-FTS, ACE-MAESTRO, and SAGE III/ISS, covering different time periods between 1985 and 2019

The variants are not expected to have any significant impact on performance.

### 3.2.1.2 CDR-4

The minimum VRWV CDR-4 production scenario covers as temporal period the three years 2010–2012 with input data from MLS, ACE-FTS, ACE-MAESTRO, IMS, and MIPAS (till to the end of 2011). This scenario represents the generation of 'Prototype dataset' as specified in work package 4 of the project.

The maximum scenario covers the processing period 2010–2014 with all available input data from MLS, ACE-FTS, ACE-MAESTRO, IMS, and MIPAS. This scenario represents the generation of the 'Full prototype dataset' as specified in work package 4 of the project.

Minimum	MLS, ACE-FTS, ACE-MAESTRO, IMS 2010-2012  MIPAS 2010-2011
Maximum	MLS, ACE-FTS, ACE-MAESTRO, IMS 2010-2014  MIPAS 2010-2012/04

The variants should not have significant impact on performance. The impact of the variants to the requirements is on the functional and the performance level with respect to input data ingestion, pre-processing, VRWV retrieval, and data transfer and storage capacity for intermediate and final products.

### 3.2.2 Products and variables

The minimum for both CDR-3 and CDR-4 is what has been generated as 'Prototype datasets'. The maximum extends this by a representation of uncertainty variables and quality flags in the products, resulting from a more comprehensive consideration in subsequent versions of the VRWV merging algorithm and corresponding ATBD.

Minimum	VRWV CDR-3 and CDR-4 prototype products with no uncertainty and data quality flag included
Maximum	VRWV CDR-3 and CDR-4 products with uncertainty and data quality flag included, following definition in further extended VRWV merging algorithms (as specified in the ATBD)

The impact of the variant on requirements is mainly on the addition of uncertainty and quality flags in CDR-3 and CDR-4 data. Including uncertainty in the two CDRs should increase the processing time and storage demand, but is required to improve the product quality.

### 3.2.3 Reprocessing capability

The capability to re-process VRWV datasets or subsets is essential for quality improvement. The variants below represent different levels of performance of the whole processing chain.

#### 3.2.3.1 CDR-3

Minimum	Reprocessing of full dataset VRWV CDR-3 in 1 month
Maximum	Reprocessing of full dataset VRWV CDR-3 in 2 weeks

#### 3.2.3.2 CDR-4

Minimum	Reprocessing of 1 year VRWV CDR-4 in 2 months
Maximum	Reprocessing of 1 year VRWV CDR-4 in 0.5 months

The impact of the variants on requirements is mainly on the performance level. As learned from the 'Prototype dataset' generation for both CDR-3 and CDR-4, the maximum variant can be achieved with the given resources as described in the SSD. The minimum variant would provide some additional time to identify and address possible issues.

### 3.2.4 Improvement cycle

Improvements might result from a better VRWV merging algorithm (e.g. between limb- and nadir-view retrievals), new version input data (e.g., for Aura-MLS v5.1 or ACE-FTS v4.1), and the implementation of new input data (e.g., AIRS, SAGE III-ISS) when merging VRWV retrievals from multiple sensors.

#### 3.2.4.1 CDR-3

Minimum	Generation of 'Prototype Dataset' as specified in work package 4 of the project
Maximum	Generation of the 'Final dataset' as specified in work package 4 of the project, which will likely include further algorithm improvements and addition of input data from the SAGE III-ISS

#### 3.2.4.2 CDR-4

Minimum	Generation of 'Prototype Dataset' as specified in work package 4 of the project
Maximum	Generation of the full prototype dataset as specified in work package 4 of the project, which will likely include further algorithm improvements and elimination of problems

The impacts of the variants on the requirements are direct and indirect. Assuming that the currently known options for improvements are final, the system shall directly support their development. Assuming that there will be more improvements over time, the system shall indirectly support the listed improvements by supporting an improvement cycle.



### 3.2.5 Validation and interaction with users

The minimum corresponds to what has been delivered for CDR-3 and CDR-4 with 'Prototype Datasets'. Additional features are improved and more convenient means for data access/exchange, and communication means towards the user and from users back to the science team.

Minimum	Delivery of prototype L3 products via FTP and interaction with key users (e.g., the Climate Research Group, CMUG)
Maximum	Final data access is possible both via FTP and CCI Open Data Portal. Provide feedback interface for strengths and weaknesses of the products, for climate modellers and WV community

As for CDR-1 and CDR-2, the impact of the variants on requirements is on the functional and interface level. The maximum variant may be translated into e.g. a customer relationship management with traceable communication, issue tracking, and improvements for users to simplify their use of VRWV products.

## 4. WV\_CCI PRODUCTION WORKFLOW

### 4.1 TCWV Processing System (BC)

#### 4.1.1 Processing chain and processors

The TCWV processing chain and the related processors for TCWV L2 and L3 retrieval, the merge of L3 products, and the generation of the final CF- and CCI compliant products are described in detail in the SSD [7].

#### 4.1.2 Main inputs and outputs

The generation of the TCWV products as specified in the PSD [3] makes use of the complete MERIS RR L1B archive (2002-2012), MODIS Terra TOA reflectance products (MOD021KM) and MODIS L2 cloud products (MOD35\_L2) from 2011–2017, OLCI RR L1B products from 2016–2017, supplemented by the use of CM SAF HOAPS TCWV L3 products (2002–2017) over ice-free ocean derived from SSM/I and SSMIS data. As an option, the ingestion of MODIS Aqua TOA reflectance products (MYD021KM). The DARD [4] lists all these inputs in detail, a summary is given in Table 4-1. All outputs are listed in detail in the PSD [3].

**Table 4-1: Main inputs of the TCWV processing chain**

Product	Coverage	Volume	Comment
MERIS RR L1B	2002–2012	~ 27.5 TB	Output of 4th reprocessing <sup>1</sup>
MODIS Terra L1B	2011–2017	~ 105 TB	MOD021KM product
MODIS L2 Cloud	2011–2017	~ 7 TB	MOD35_L2 product
OLCI RR L1B	2016–2017	~ 30 TB	Output of 1st reprocessing
CM SAF HOAPS TCWV L3	2002–2017	~120 GB	Data from SSM/I, SSMIS. Pre-processed by DWD
MODIS Aqua L1B (optional) <sup>2</sup>	2002–2017	~240 TB	MYD021KM product.

<sup>1</sup> If possible, it is planned to use data from the 4th reprocessing for the generation of the final TCWV datasets. At compilation time of this version 1 of the SRD, the MERIS L1b data from the 4th reprocessing are still under review at ESA and not yet publicly available.

<sup>2</sup> The ingestion of MODIS Aqua was foreseen as an option in the WV\_cci project. At compilation time of this version 2 of the SRD, a final decision on this option and the period of MODIS Aqua data to be used has not been made.

### 4.1.3 Estimation of system key parameters

WV\_cci TCWV products will be created in subsequent production cycles:

- 'Dataset 1' for 2010–2012, using MERIS, MODIS, and CM SAF HOAPS data
- 'Dataset 2' for 2015–2017, using OLCI, MODIS, and CM SAF HOAPS. This production cycle will also include a reprocessing of 'Dataset 1'
- 'Full dataset' for 2002–2017, using all available input data
- 'Option': same as 'Full dataset', but in addition using MODIS Aqua data from 2002–2017. As mentioned, a final decision about this option has not yet been made.

This subsection tries to preliminary estimate the overall data volume and processing power required for the TCWV production. The numbers are given in Table 4-2.

The overall input data volume is about 170 TB (plus optional MODIS Aqua data). The estimation of the overall processing time volume is given in wall clock time as this gives a better impression in relation to overall project progress. However, these numbers are rough estimates as they are dependent on available computing resources which may have to be shared with other users on both JASMIN and Calvalus systems.

**Table 4-2: Overall data volume and estimation of processing time**

Item	Description	Overall Volume
Input data	as listed in Table 4-1	~ 170 TB (+ 240TB MODIS Aqua option)
Intermediate data	Pixel classification and ERA Interim products	~60 TB (estimate, dynamically deleted when no longer needed)
Output data	All TCWV L3 products to be delivered and as specified in PSD. Does not include any L2 products	~ 12 GB per year to process
L2 pre-processing time	Pixel classification at level 2 and auxiliary data ingestion. Parallel processing of MERIS at Calvalus and MODIS at JASMIN	~ 2 days (wall clock time) per year to process
TCWV L2 processing time	TCWV retrieval from pre- processing outcome. Parallel processing of MERIS at Calvalus and MODIS at JASMIN	~ 2 days per year to process
TCWV L3 processing time	Daily and monthly aggregation, sensor merge, ingestion of HOAPS products, final product formatting. All L3 processing at Calvalus	~ 1 day per year to process

## 4.2 VRWV Processing System (UoR)

### 4.2.1 Processing chain and processors

The VRWV processing chain covering the merging of L3 products and the generation of the final CF- and CCI compliant products are described in detail in the SSD [7].

### 4.2.2 Main inputs and outputs

The generation of the VRWV CDR-3 and CDR-4 products as specified in the PSD [3] makes use of a range of vertically resolved L2 and L3 WV products. The DARD [4] lists all these inputs in detail; a summary is given in Table 4-3 and Table 4-4 for CDR-3 and CDR-4, respectively. All outputs are listed in detail in the PSD [3].

**Table 4-3: Main inputs of the VRWV CDR-3 processing chain**

Product	Coverage	Volume	Comment
SAGE II V7.0	1985–2005	~50 MB	

Product	Coverage	Volume	Comment
UARS MLS V6.0	1991–1993		All L3 input data (the SPARC Data Initiative zonal mean climatologies) [10] can be accessed via the Pangea data archive (doi TBD).
HALOE V19	1991–2005		
SAGE III V4.0	2002–2005		
POAM III V4.0	1998–2005		
SMR 544 GHz-band V2-0 SMR 489 GHz-band V2-1	2001–2018		
Aura MLS V4.2	2004–2019		
ACE-FTS V3.6	2004–2019		
ACE-MAESTRO V31	2004–2019		
MIPAS V3o_H2O_21 MIPAS V5r_H2O_224	2002–2004 2005–2012		
SCIAMACHY V4.2	2002–2012		
SAGE III-ISS V5.1	2017–2019		

**Table 4-4: Main inputs of the VRWV CDR-4 processing chain**

Product	Coverage	Volume	Comment
Aura MLS L2	2010–2014	~6 GB	V4.2 from NASA JPL
ACE-FTS L2	2010–2014	~500 MB	V3.5 from WAVAS_SAHAR
ACE-MAESTRO L2	2010–2014	~300 MB	V31 from WAVAS_SAHAR
MIPAS L2	2010–2011	~5 GB	V7R MIPAS-ESA
IMS L2	2010–2014	~6 TB	V2.1 from RAL

## 4.2.3 Estimation of system key parameters

### 4.2.3.1 CDR-3

WV\_cci VRWV CDR-3 products will be created in three subsequent production cycles:

- 'Prototype Dataset v0' for 1985-2018, using SAGE II, UARS-MLS, HALOE, MIPAS, Aura-MLS, POAM III, SMR, SCIAMACHY, ACE-MAESTRO, and ACE-FTS
- 'Prototype Dataset v1' for 1985-2018, using SAGE II, UARS-MLS, HALOE, MIPAS, Aura-MLS, POAM III, SMR, SCIAMACHY, ACE-MAESTRO, ACE-FTS, SAGE III-ISS
- 'Full dataset' for 1985-2019, using using SAGE II, UARS-MLS, HALOE, MIPAS, Aura-MLS, POAM III, SMR, SCIAMACHY, ACE-MAESTRO, ACE-FTS, SAGE III-ISS.

Here we provide a preliminary estimate of the overall data volume and processing power required for the production of VRWV CDR-3. The numbers are given in Table 4-5.

The overall input data volume is 50 MB. The estimation of the overall processing time is 2 weeks. These estimates are only weakly dependent on available computing resources, but may be an underestimate depending on the quality of the input data obtained.

**Table 4-5: Overall data volume and estimation of processing time for CDR-3**

Item	Description	Overall Volume
Input data	as listed in Table 4-3	~ 50 MB
Intermediate data	Chemistry-climate model WV time series from a specified dynamics simulation	~ 2 GB
Output data	All VRWV CDR-3 L3 products to be delivered as specified in PSD [3]. Does not include input data products.	~ 15 MB
VRWV L3 processing time	Bias-correction of input L3 SPARC Data Initiative climatologies using chemistry-climate model as transfer function and merging of VRWV CDR-3 (L3)	~ 2 weeks

## 4.2.3.2 CDR-4

WV\_cci VRWV CDR-4 products will be created in subsequent production cycles:

- 'Dataset 0' for 2010–2012, using MLS, ACE-FTS, ACE-MAESTRO, MIPAS, and IMS data
- 'Full dataset' for 2010–2014, using MLS, ACE-FTS, ACE-MAESTRO, MIPAS, and IMS data
- 'Option': same as 'Full dataset', but in addition using AIRS data from 2010–2014. As mentioned, a final decision about this option has not yet been made.

This subsection tries to provide a preliminary estimate of the overall data volume and processing power required for the VRWV CDR-4 production. The numbers are given in Table 4-6.

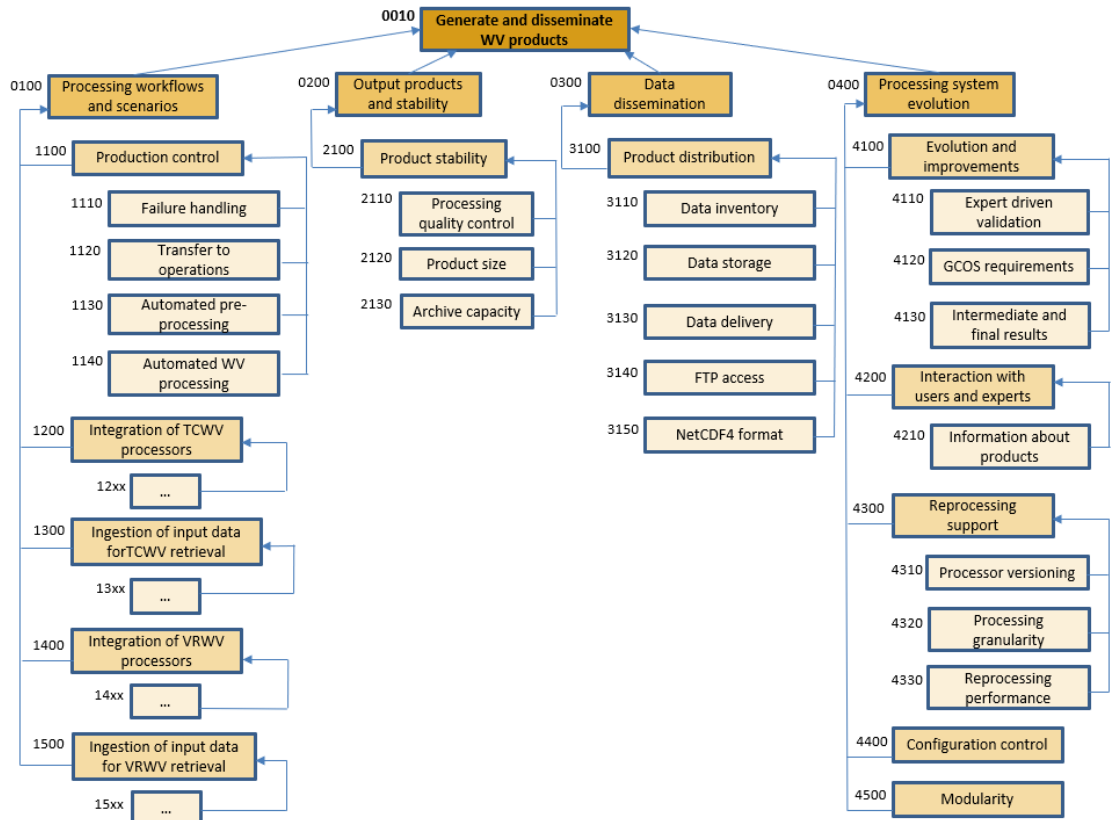
The overall input data volume is about 6 TB. The estimation of the overall processing time volume is given in wall clock time as this gives a better impression in relation to overall project progress. However, these numbers are rough estimates as they are dependent on available computing resources which may have to be shared with other users on the Met RACC system.

**Table 4-6: Overall data volume and estimation of processing time for CDR-4**

Item	Description	Overall Volume
Input data	as listed in Table 4-4	~ 6 TB
Intermediate data	Derived meteorological products (DMPs) for each input data obtained from NASA JPL	~6 TB
Output data	All VRWV CDR-4 L3 products to be delivered and as specified in PSD. Does not include any L2 products	~ 500 MB for 2010-2014
L2 pre-processing time	Combination of profiles and DMPs for each input data, excluding outliers	~ 2 days
VRWV L2 processing time	Bias-correction of VRWV L2 profiles according to GRUAN data	~ 1 day per year to process
VRWV L3 processing time	Monthly aggregation, sensor merge, ingestion of HOAPS products, final product formatting	~ 1 day

## 5. DETAILED SYSTEM REQUIREMENTS

With the source requirements from section 2.3, the WV\_cci scenarios from section 3 and the production workflows from section 4 as basis, this section derives a set of requirements for the WV\_cci system from these inputs.



**Figure 5-1: Detailed system requirements and their hierarchy, starting from the high-level requirements given in section 2.5. The numbers refer to the requirements defined further in this section. The empty boxes (...) are further refined in Figure 5-2 for TCWV and Figure 5-3 for VRWV.**



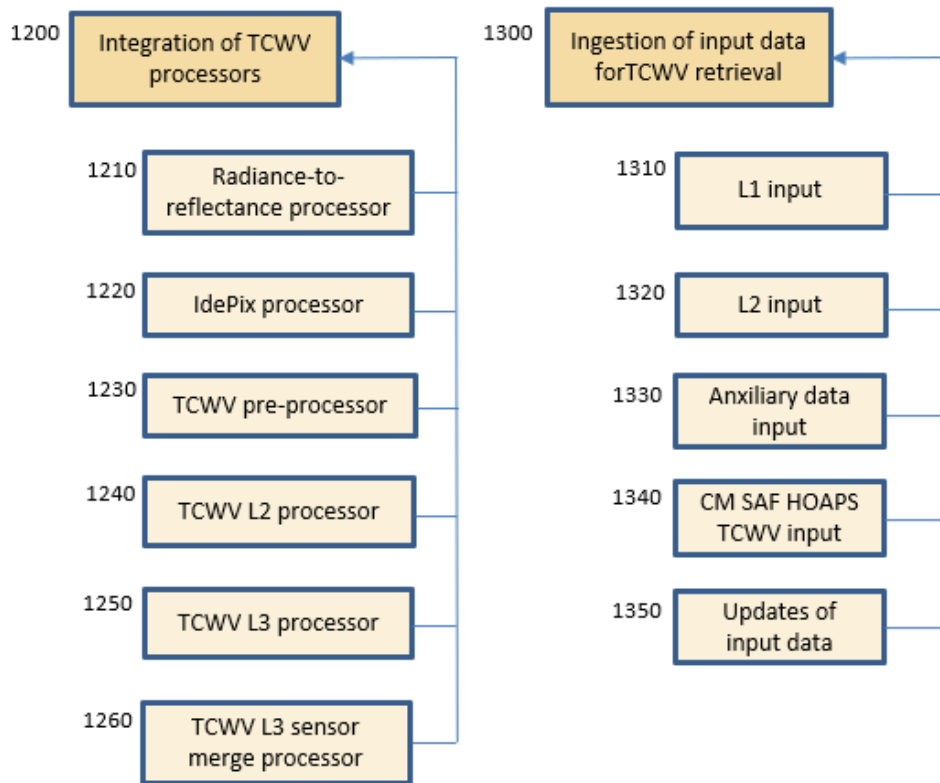


Figure 5-2: Detailed system requirements for the TCWV processing system part.

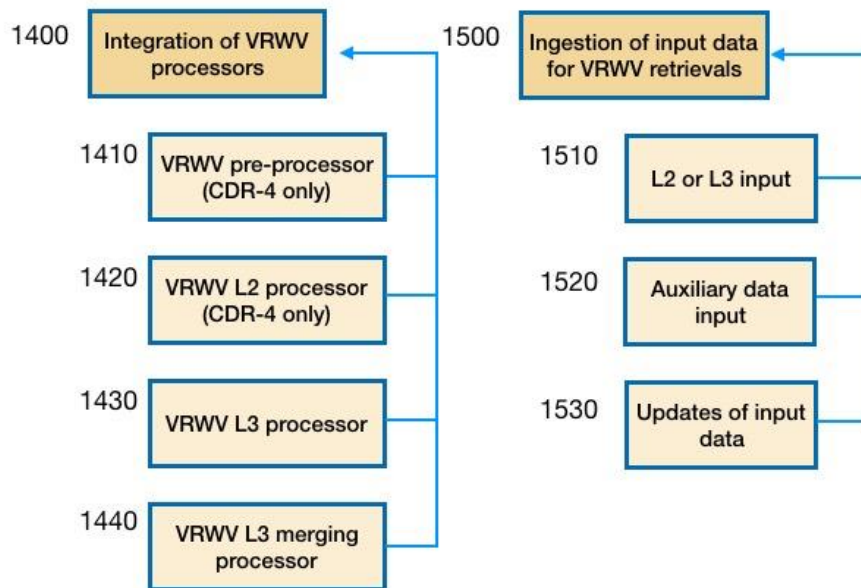


Figure 5-3: Detailed system requirements for the VRWV processing system part.

The hierarchy splits the high level requirements into requirements on

- Processing workflows and scenarios (1xxx)
- Output products and stability (2xxx)
- Data dissemination (3xxx)
- System evolution (4xxx),

as illustrated in Figure 5-1, Figure 5-2, and Figure 5-3. The numbering scheme follows this hierarchy while the following subsections and thus the sequence of listing requirements follows the ECSS categories.

## 5.1 Functional Requirements

<b>WV-SR-1100</b>	<p><b>Production control</b></p> <p>The WV_cci system shall control production according to the processing workflows for pre-processing, TCWV and VRWV retrieval.</p>
-------------------	---

<b>WV-SR-1200</b>	<p><b>Integration of TCWV processors</b></p> <p>The WV_cci system shall integrate the processors required for TCWV retrieval.</p>
-------------------	---

<b>WV-SR-1400</b>	<p><b>Integration of VRWV processors</b></p> <p>The WV_cci system shall integrate the processors required for VRWV retrieval.</p>
-------------------	---

<b>WV-SR-2100</b>	<p><b>Stability and quality of outputs</b></p> <p>The WV_cci system shall provide means to ensure the stability and quality of its outputs.</p>
-------------------	---

<b>WV-SR-2110</b>	<p><b>Processing quality control</b></p> <p>The WV_cci system shall provide means to do quality checks and manual inspection of intermediate results during processing.</p>
-------------------	---

<b>WV-SR-3100</b>	<p><b>Product distribution</b></p> <p>The WV_cci system shall provide means to distribute the generated WV products to the user community and to the Open Data Portal.</p>
-------------------	--

<b>WV-SR-4100</b>	<p><b>Processing system evolution and improvement</b></p> <p>The WV_cci system shall be designed in a way that allows its continuous evolution and improvement.</p>
-------------------	---

### 5.1.1 Preprocessing

<b>WV-SR-1210</b>	<p><b>Radiance-to-Reflectance processor</b></p> <p>The WV_cci TCWV processing system shall provide a processor for conversion of MERIS and OLCI TOA radiance to TOA reflectance.</p>
-------------------	--

<b>WV-SR-1220</b>	<p><b>IdePix pixel classification processor</b></p> <p>The WV_cci TCWV processing system shall integrate the IdePix processor for pixel classification of MERIS, MODIS and OLCI L1b products.</p>
-------------------	---

<b>WV-SR-1230</b>	<p><b>TCWV pre-processor</b></p> <p>The WV_cci TCWV processing system shall provide a processor which generates a colocation product with all input variables required for TCWV L2 retrieval from L1b TOA reflectances.</p>
-------------------	---

<b>WV-SR-1410</b>	<p><b>VRWV pre-processor</b></p> <p>The WV_cci VRWV processing system shall provide a processor which combines VRWV profiles from all input data with the DMPs from reanalyses (CDR-4 only).</p>
-------------------	--

### 5.1.2 WV processing

<b>WV-SR-1240</b>	<p><b>TCWV L2 retrieval processor</b></p> <p>The WV_cci TCWV processing system shall provide a processor which generates a TCWV L2 product from the final pre-processed colocation product.</p>
-------------------	---

<b>WV-SR-1250</b>	<p><b>TCWV L3 processor</b></p> <p>The WV_cci TCWV processing system shall provide a processor which generates a TCWV L3 daily aggregation product from the TCWV L2 products for a given day.</p>
<p><b>As it was agreed during the project, this functionality shall be extended to monthly L3 products. This was not foreseen in the original production plan, but has turned out to be a very useful add-on.</b></p>	

<b>WV-SR-1260</b>	<p><b>TCWV L3 sensor merge processor</b></p>
-------------------	--

	The WV_cci TCWV processing system shall provide a processor to merge TCWV L3 products obtained from the different sensors.
--	--

<b>WV-SR-1420</b>	<p><b>VRWV L2 retrieval processor</b></p> <p>The WV_cci VRWV processing system shall provide a processor that generates a bias-corrected VRWV L2 product from the input VRWV profiles (CDR-4 only).</p>
-------------------	---

<b>WV-SR-1430</b>	<p><b>VRWV L3 processor</b></p> <p>The WV_cci VRWV processing system shall provide a processor which generates bias-corrected VRWV L3 products from the input VRWV L3 data (CDR-3) using a chemistry-climate model as transfer function.</p> <p>The WV_cci VRWV processing system shall provide a processor which generates a VRWV L3 monthly aggregation product from the VRWV L2 products (CDR-4).</p>
-------------------	--

<b>WV-SR-1440</b>	<p><b>VRWV L3 sensor merging processor</b></p> <p>The WV_cci VRWV processing system shall provide a processor to merge VRWV L3 products obtained from the different sensors.</p>
-------------------	--

### 5.1.3 Validation

<b>WV-SR-4110</b>	<p><b>Expert-driven validation</b></p> <p>The WV_cci system shall support the validation of output products by the validation team and the re-integration of the validation results into the system (i.e. processor updates).</p>
-------------------	---

<b>WV-SR-4120</b>	<p><b>GCOS requirements</b></p> <p>The WV_cci validation system shall assess to what extent the products meet all GCOS requirements for the WV ECV.</p>
-------------------	---

<b>WV-SR-4130</b>	<p><b>Intermediate and final processing results</b></p> <p>The WV_cci system shall provide means to do quality checks for all types of intermediate and final results during processing.</p>
<p>These quality checks are provided by a set of verification steps which are described in detail in the SVR [9].</p>	

#### 5.1.4 Data Management

<b>WV-SR-3110</b>	<p><b>Data inventory</b></p> <p>The WV_cci system shall provide an inventory of its data products. The inventory shall support product versioning.</p>
-------------------	--

<b>WV-SR-3120</b>	<p><b>Structured data storage</b></p> <p>The WV_cci shall maintain its data in a structured way. The structure shall be organised along product types, versions, and temporal coverage. The structure shall allow to keep several versions of outputs at the same time.</p>
-------------------	---

#### 5.1.5 User Interaction

<b>WV-SR-3130</b>	<p><b>On-demand data delivery</b></p>
-------------------	---------------------------------------

	The WV_cci system shall provide subsets of specific WV data products on demand.
In collaboration with the users, this function may also include intermediate products which go beyond the agreed datasets delivered in the various production cycles (e.g. TCWV L2 products).	

<b>WV-SR-4200</b>	<b>Interaction with users and groups of scientific experts</b>  The WV_cci systems, i.e. their developers/operators, shall establish a continuous interaction with users and relevant groups of scientific experts.
-------------------	---

<b>WV-SR-4210</b>	<b>Information about products</b>  The WV_cci systems shall provide means to inform users (including the CCI Open Data Portal) about new versions of products.
-------------------	--

## 5.2 Operational Requirements

<b>WV-SR-1110</b>	<b>Failure handling</b>  The WV_cci system shall provide means for operators to detect and handle failures.
-------------------	---

<b>WV-SR-1120</b>	<b>Transfer to operations</b>  The WV_cci system shall provide an operational procedure to transfer a processor version from development to operations after validation with minimal effort.
-------------------	--

<b>WV-SR-4300</b>	<p><b>Reprocessing support</b></p> <p>The WV_cci system shall support re-processing of parts or all of the dataset in defined production cycles.</p>
-------------------	--

<b>WV-SR-4310</b>	<p><b>Processor Versioning</b></p> <p>The WV_cci system shall be designed in a way that allows to install and run different versions of processors in the same environment.</p>
-------------------	---

<b>WV-SR-4320</b>	<p><b>Processing granularity</b></p> <p>The WV_cci system shall be designed in a way that operators handle requests for larger datasets of 1 year or one month of inputs. Handling single products should still be possible for recovery from failure.</p>
-------------------	--

### 5.2.1 Preprocessing

<b>WV-SR-1130</b>	<p><b>Automated pre-processing</b></p> <p>The pre-processing chains of the WV_cci system shall be automated such that operators are required to submit requests and monitor the system for the processing steps.</p>
-------------------	--

### 5.2.2 WV processing

<b>WV-SR-1140</b>	<b>Automated WV processing</b>
-------------------	--------------------------------



	The WV processing chains of the WV_cci system shall be automated such that operators are required to submit requests and monitor the system for the processing steps.
--	---

### 5.2.3 Data Management

There are currently no specific operational requirements to data management.

### 5.2.4 User Interaction

There are currently no specific operational requirements to user interaction.

## 5.3 Performance and Sizing Requirements

<b>WV-SR-1600</b>	<p><b>Production performance</b></p> <p>The WV_cci system shall be able to generate the WV datasets as specified in the PSD within the foreseen production cycles.</p>
-------------------	--

<b>WV-SR-1610</b>	<p><b>Processing Concurrency</b></p> <p>The WV_cci system shall provide means for massive parallel processing.</p>
-------------------	--

<b>WV-SR-4330</b>	<p><b>Reprocessing performance</b></p> <p>The WV_cci system shall be able to reprocess dedicated WV datasets within their foreseen production cycles.</p>
-------------------	---

<b>WV-SR-2120</b>	<b>Size of generated products</b>
-------------------	-----------------------------------

	<p>The WV_cci system shall be able to process WV datasets in suitable blocks of sub-periods in which intermediate data are automatically deleted to meet the storage capacity limits.</p>
--	---

### 5.3.1 Preprocessing

There are currently no specific operational requirements to preprocessing.

### 5.3.2 WV processing

<p><b>WV-SR-1620</b></p>	<p><b>Map-reduce for level 3 binning and aggregation</b></p> <p>The WV_cci system should support distributed processing of level 3 products by an infrastructure that supports sorting of large parallel data streams, e.g. by the map-reduce approach.</p>
--------------------------	---

### 5.3.3 Data Management

<p><b>WV-SR-2130</b></p>	<p><b>Archive capacity</b></p> <p>The WV_cci system shall provide storage capacity in its production archive for the generation of the WV datasets as specified in the PSD.</p>
--------------------------	---

### 5.3.4 User Interaction

There are currently no specific operational requirements to user interaction.

## 5.4 Interface Requirements

<b>WV-SR-1300</b>	<b>Ingestion of input data</b>  The WV_cci system shall ingest all input data sets required for TCWV retrieval.
-------------------	---

<b>WV-SR-1500</b>	<b>Ingestion of input data</b>  The WV_cci system shall ingest all input data sets required for VRWV retrieval.
-------------------	---

#### 5.4.1 Preprocessing

<b>WV-SR-1310</b>	<b>Level 1 input for TCWV retrieval</b>  The WV_cci system for TCWV retrieval shall ingest and process level 1 products from MERIS, MODIS Terra and OLCI.
-------------------	---

<b>WV-SR-1320</b>	<b>Level 2 input for TCWV retrieval</b>  The WV_cci system for TCWV retrieval shall ingest MOD35_L2 cloud products from MODIS.
-------------------	--

<b>WV-SR-1330</b>	<b>Anxiliary data input for TCWV retrieval</b>  The WV_cci system for TCWV retrieval shall ingest anxiliary datasets required for preprocessing.
-------------------	--

<b>WV-SR-1510</b>	<p><b>Level 3 (CDR-3) or Level 2 (CDR-4) input for VRWV merge</b></p> <p>The WV_cci system for VRWV CDR-3 shall ingest VRWV profiles from SAGE II, UARS-MLS, HALOE, MIPAS, Aura-MLS, POAM III, SMR, SCIAMACHY, ACE-MAESTRO, and ACE-FTS.</p> <p>The WV_cci system for VRWV CDR-4 shall ingest VRWV profiles from MLS, ACE-FTS, ACE-MAESTRO, MIPAS, and IMS.</p>
-------------------	---

<b>WV-SR-1520</b>	<p><b>Anxiliary data input for VRWV retrieval</b></p> <p>The WV_cci system for VRWV retrieval shall ingest anxiliary datasets required for preprocessing. This includes chemistry-climate model data for CDR-3 and DMP data for CDR-4.</p>
-------------------	--

#### 5.4.2 WV processing

<b>WV-SR-1340</b>	<p><b>TCWV Level 3 input</b></p> <p>The WV_cci system for TCWV retrieval shall ingest CM SAF HOAPS TCWV L3 products generated from SSM/I and SSMIS over ice-free ocean.</p>
-------------------	---

#### 5.4.3 Data Management

<b>WV-SR-1350</b>	<p><b>Updates of input data for TCWV retrieval</b></p> <p>The WV_cci system shall have a function to ingest new or updated input data from media.</p>
-------------------	---

<b>WV-SR-1530</b>	<p><b>Updates of input data for VRWV retrieval</b></p>
-------------------	--

	<p>The WV_cci system shall have a function to ingest new or updated input data from media.</p>
--	--

#### 5.4.4 User Interaction

<p><b>WV-SR-3140</b></p>	<p><b>FTP access</b></p> <p>The WV_cci system shall provide access to its output products by FTP-pull.</p>
--------------------------	--

<p><b>WV-SR-3150</b></p>	<p><b>NetCDF output format</b></p> <p>The WV_cci system shall provide outputs in a format that conforms to the CCI Data Standards.</p>
--------------------------	--

#### 5.5 Reliability, Availability and Maintainability Requirements

<p><b>WV-SR-4400</b></p>	<p><b>Configuration control</b></p> <p>The WV_cci system shall maintain its software components and in particular its processors and configurations under configuration control</p>
--------------------------	---

<p><b>WV-SR-4500</b></p>	<p><b>Modularity</b></p> <p>The WV_cci system shall follow a modular approach in its design, in particular for the processors that shall be interchangeable by other versions and different implementations without changing other parts of the system.</p>
--------------------------	---

### 5.5.1 Preprocessing

The reliability and maintainability requirements listed above are applicable here.

### 5.5.2 WV processing

The reliability and maintainability requirements listed above are applicable here.

### 5.5.3 Data Management

The reliability and maintainability requirements listed above are applicable here.

### 5.5.4 User Interaction

The reliability and maintainability requirements listed above are applicable here.

## 6. REQUIREMENTS TRACEABILITY

This section traces input requirements to sections of this document (§) or to particular WV\_cci system requirements (WV-SR-xxxx) that implement them.

Input Req. ID	Task	Reference
SoW-TR-1	GCOS requirements for WV ECV products	WV-SR-4120
SoW-TR-2	Interaction with end-user community	Assumption-1, WV-SR-4110 WV-SR-4200 WV-SR-4210
SoW-TR-3	Raising the profile of WV_cci	Assumption-1, Assumption-4
SoW-TR-4	Demonstration that work in WV_cci is complementary to other water vapour ECV activities	Assumption-2, Assumption-5 WV-SR-4110
SoW-TR-5	Focus of scientific issues	Assumption-2, WV-SR-1310
SoW-TR-6	Collect requirements from climate scientists	§2.3
SoW-TR-7	User survey	URD v2.0
SoW-TR-8	Quantitative requirements specification	§2.3
SoW-TR-9	User requirements for product uncertainties	WV-SR-1240 WV-SR-4110
SoW-TR-10	User requirements for water vapour isotopologue information	URD v2.0
SoW-TR-11	Selection of algorithms to use	WV-SR-1200 WV-SR-1300 WV-SR-4110

Input Req. ID	Task	Reference
SoW-TR-12	Consideration of product uncertainties in algorithms	WV-SR-1310 WV-SR-1330 WV-SR-4110 WV-SR-4120
SoW-TR-13	Instrument selection for TCWV retrievals	WV-SR-1310
SoW-TR-14	Provide long-term multi-mission TCWV ECV product	WV-SR-1200 WV-SR-3100
SoW-TR-15	Provide harmonised multi-mission stratospheric vertical profile ECV product	WV-SR-1300 WV-SR-3100
SoW-TR-16	Liaise with ozone and aerosol CCIs for access to stratospheric auxiliary data products	DARD v2.0
SoW-TR-17	Survey algorithms for retrieving multiple products simultaneously	
SoW-TR-18	Round-robin intercomparison	PVASR v2.0
SoW-TR-19	Test harmonisation strategies	
SoW-TR-20	Intercompare nadir/limb retrievals	PVASR v2.0
SoW-TR-21	Assemble validation data set	WV-SR-4110
SoW-TR-22	Validation shall quantify WV accuracy and uncertainties	WV-SR-4110
SoW-TR-23	Homogeneity analysis	WV-SR-4110
SoW-TR-24	Comparison with GCOS stability requirement	WV-SR-4110 WV-SR-4120



<b>Input Req. ID</b>	<b>Task</b>	<b>Reference</b>
SoW-TR-25	Compare WV from VIS/IR with microwave sounders and in situ data	WV-SR-4110
SoW-TR-26	Compare WV products with ECVs provided by external organisations	WV-SR-4110
SoW-TR-27	Compare WV products with climate model and reanalysis data	WV-SR-4110
SoW-TR-28	Develop automated high performance processing chain	WV-SR-1100 WV-SR-4100
SoW-TR-29	Improvements of products from processing iterations	WV-SR-2100 WV-SR-4100
SoW-TR-30	Generation of ECV data records	WV-SR-1100 WV-SR-3100
SoW-TR-31	Global coverage of products, full mission records of instruments	WV-SR-1100 WV-SR-3100
SoW-TR-32	Liaise with internal and external users	WV-SR-4100 WV-SR-4200
SoW-TR-33	User case studies	WV-SR-4200
SoW-TR-34	Assess consistency and variability of ECV time series	WV-SR-4110
SoW-TR-35	Interaction with international frameworks and projects	WV-SR-4200
<b>DATA ACCESS REQUIREMENTS (from DARD)</b>		
<b>Satellite Data for Production</b>		

<b>Input Req. ID</b>	<b>Task</b>	<b>Reference</b>
DARD-SAT-1	Access data from MERIS, MODIS, OLCI	WV-SR-1310
DARD-SAT-2	Access data from CM SAF HOAPS	WV-SR-1340
DARD-SAT-3	Access data from WAVAS-II	WV-SR-1510
DARD-SAT-4	Access data from ACE-FTS	WV-SR-1510
DARD-SAT-5	Access data from ACE-MAESTRO	WV-SR-1510
DARD-SAT-6	Access data from SPARC	WV-SR-1510
DARD-SAT-7	Access data from IMS	WV-SR-1510
DARD-SAT-8	Access data from AURA-MLS	WV-SR-1510
DARD-SAT-9	Access data from MIPAS	WV-SR-1510
<b>Auxiliary and Other Data</b>		
DARD-AUX-1	Access ECMWF data from OLCI products	WV-SR-1310
DARD-AUX-2	Access data from chemistry climate models	WV-SR-1520
DARD-AUX-3	Access MERRA-2 and/or ERA5 temperature fields	WV-SR-1520
<b>Satellite Data for Comparison</b>		
DARD-COMP-1	Access data from EMiR	WV-SR-4110
DARD-COMP-2	Access data from GOME Evolution Climate product	WV-SR-4110
DARD-COMP-3	Access data from GOZCARDS (NASA)	WV-SR-4110
DARD-COMP-4	Access data from IASI	WV-SR-4110

Input Req. ID	Task	Reference
DARD-COMP-5	Access data from SWOOSH (NOAA)	WV-SR-4110
<b>Reanalysis, Climate Model and Other Data for Comparison</b>		
DARD-REA-1	Access data from CMIP6	WV-SR-4110
DARD-REA-2	Access data from ERA5	WV-SR-4110
<b>Ground-Based, In-Situ and Other Data for Validation</b>		
DARD-VAL-1	Access data from ARSA	WV-SR-4110
DARD-VAL-2	Access data from CARIBIC/MOZAIC/IAGOS	WV-SR-4110
DARD-VAL-3	Access data from GNSS	WV-SR-4110
DARD-VAL-4	Access data from GRUAN	WV-SR-4110
DARD-VAL-5	Access data from SPURT	WV-SR-4110
DARD-VAL-6	Access data from SuomiNet	WV-SR-4110
<b>Other Data Records</b>		
DARD-OTH-1	Access data from LST CCI	WV-SR-4110
DARD-OTH-2	Access data from SST CCI	WV-SR-4110
<b>USER REQUIREMENTS (from URD)</b>		
<b>Requirements for ECV Water Vapour</b>		
URD-GCOS-200, URD-GCOS-154	TCWV frequency, resolution, uncertainty, stability	WV-SR-1200 WV-SR-1300 WV-SR-4110

<b>Input Req. ID</b>	<b>Task</b>	<b>Reference</b>
URD-WMO-OSCAR	TCWV frequency, resolution, uncertainty, stability for Climate-AOPC, Global NWP, Climate Modeling (WCRP), Atmospheric Chemistry	WV-SR-1200 WV-SR-1300 WV-SR-4110
<b>Requirements from User Feedback</b>		
URD-FDB-1	CDR length	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-2	CDR 1+2 frequency	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-3	CDR 1+2 resolution	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-4	CDR 1+2 systematic accuracy	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-5	CDR 1+2 random accuracy	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-6	CDR 1+2 stability	WV-SR-1200 WV-SR-1300 WV-SR-4200
URD-FDB-7	CDR 1+2 frequency	WV-SR-1200 WV-SR-1300 WV-SR-4200
<b>PRODUCT SPECIFICATION REQUIREMENTS (from PSD)</b>		
<b>Specifications of Primary Water Vapour Data Products</b>		

<b>Input Req. ID</b>	<b>Task</b>	<b>Reference</b>
PSD-1	TCWV L3 over land from MERIS, MODIS and OLCI, 2002-2017 ('CDR-1')	WV-SR-1200 WV-SR-2100 WV-SR-3100
PSD-2	TCWV L3 over land and ocean from MERIS, MODIS, OLCI and SSMI, 2002-2017 ('CDR-2')	WV-SR-1200 WV-SR-2100 WV-SR-3100
PSD-3	Vertically resolved stratospheric water vapour climatologies, 1985-2019 ('CDR-3')	WV-SR-1300 WV-SR-2100 WV-SR-3100  WV-SR-1500
PSD-4	Vertically resolved water vapour profiles, 2010-2014 ('CDR-4')	WV-SR-1300 WV-SR-2100 WV-SR-3100  WV-SR-1500
PSD-5	Variables in TCWV L3 products	WV-SR-1200 WV-SR-3150
PSD-6	Spatial and temporal resolution of TCWV L3 products	WV-SR-1200 WV-SR-3150
PSD-7	Accuracy and stability of TCWV L3 products	WV-SR-1200 WV-SR-2100
PSD-8	Variables in vertically resolved water vapour products	WV-SR-1400 WV-SR-3150
PSD-9	Spatial and temporal resolution of vertically resolved water vapour products	WV-SR-1400 WV-SR-3150
PSD-10	Accuracy and stability of vertically resolved water vapour products	WV-SR-2100

## APPENDIX 1: REFERENCES

- [1]: ESA Climate Change Initiative Extension (CCI+) Phase 1 – New Essential Climate Variables – Statement of Work. Ref. ESA-CCI-PRGM-EOPS-SW-17-0032, Issue 1, Rev. 3, 22 August 2017.
- [2]: World Meteorological Organization: [Essential Climate Variables](#). © 2019 World Meteorological Organization (WMO).
- [3]: ESA CCI Water Vapour: Product Specification Document. M. Schröder, M. Hegglin, O. Danne, J. Fischer, A. Lang, R. Siddans, C. Sioris, G. Stiller, and K. Walker. Issue 2.0, 1 November 2019.
- [4]: ESA CCI Water Vapour: Data Access Requirement Document. M. Schröder, M. Hegglin, H. Brogniez, J. Fischer, D. Hubert, A. Laeng, R. Siddans, C. Sioris, G. Stiller, T. Trent, and K. Walker. Issue 1.0, 11 December 2018.
- [5]: ESA CCI Water Vapour: Algorithm Theoretical Basis Document. J. Fischer et al., Issue 1.1, 3 April 2019.
- [6]: ESA CCI Water Vapour: User Requirements Document. M. Hegglin, M. Schröder, H. Brogniez, and D. Hubert. Issue 1.1, 28 Januar 2019.
- [7]: ESA CCI Water Vapour: System Specification Document. O. Danne and M. Hegglin, Issue 2.0, 8 May 2020.
- [8]: ESA Climate Change Initiative Extension (CCI+) Phase 1 – New Essential Climate Variables – Statement of Work. Annex A: Water Vapour ECV (Water\_Vapour\_cci). Ref. ESA-CCI-PRGM-EOPS-SW-17-0032, Issue 1, Rev. 3, 22 August 2017.
- [9]: ESA CCI Water Vapour: System Verification Report. O. Danne and M. Hegglin, Issue 2.0, 8 May 2020.
- [10]: 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.
- [11]: ESA CCI Project Guidelines. Ref. EOP-DTEX-EOPS-TN-10-0002, Issue 1.0, 5 November 2010.
- [12]: ESA Climate Change Initiative Phase I - Scientific User Consultation and Detailed Specification Statement of Work (SoW). Issue 1.4, 9 November 2009.
- [13]: Ocean Colour Climate Change Initiative (OC\_CCI) - Phase One – System Requirements Document. Peters, M., Boettcher, M., Fomferra, N., and M. Grant. Ref. AO-1/6207/09/I-LG, Issue 1.0, 30 November 2011.
- [14]: Sea Surface Temperature Climate Change Initiative (SST\_CCI) - Phase One – System Requirements Document. Quast, R., Fomferra, N., Boettcher, M., Merchant, C., and P. Spinks, Issue 1.1, 30 April 2012.
- [15]: ESA CCI Open Data Portal. <http://cci.esa.int/content/cci-open-data-portal>

## APPENDIX 2: GLOSSARY

<b>Term</b>	<b>Definition</b>
<i>ACE-FTS</i>	Atmospheric Chemistry Experiment - Fourier Transform Spectrometer
<i>ACE-MAESTRO</i>	Atmospheric Chemistry Experiment - Measurement of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
<i>ARSA</i>	Analyzed RadioSOUNDings Archive
<i>ATBD</i>	Algorithm Theoretical Basis Document
<i>BC</i>	Brockmann Consult
<i>Calvalus</i>	Cal/Val and User Services
<i>CARIBIC</i>	Civil Aircraft for the Regular Investigation of the atmosphere Based on an Instrument Container
<i>CCI</i>	Climate Change Initiative
<i>CM SAF</i>	Satellite Application Facility on Climate Monitoring
<i>CMUG</i>	Climate Modelling User Group of the CCI
<i>DARD</i>	Data Access Requirement Document
<i>ECSAT</i>	European Centre for Space Applications and Telecommunications
<i>ECSS</i>	European Cooperation for Space Standardization
<i>ECV</i>	Essential Climate Variable
<i>EMIR</i>	ERS/Envisat MWR Recalibration and Water Vapour Thematic Data Record Generation
<i>EO</i>	Earth Observation
<i>ESA</i>	European Space Agency
<i>FTP</i>	File Transfer Protocol
<i>GCOS</i>	Global Climate Observing System
<i>GNSS</i>	Global Navigation Satellite System
<i>GOME</i>	Global Ozone Monitoring Experiment
<i>GOZCARDS</i>	Global Ozone Chemistry And Related trace gas Data records for the Stratosphere
<i>GRUAN</i>	GCOS (Reference) Upper-Air Network
<i>HOAPS</i>	Hamburg Ocean Atmosphere Parameters and Fluxes
<i>IAGOS</i>	In-service Aircraft for a Global Observing System
<i>IASI</i>	Infrared Atmospheric Sounding Interferometer
<i>IMS</i>	Infrared Microwave Sounding

<b>Term</b>	<b>Definition</b>
<i>JASMIN</i>	Joint Analysis System Meeting Infrastructure
<i>LST</i>	Land Surface Temperature
<i>MERIS</i>	Medium Resolution Imaging Spectrometer
<i>MIPAS</i>	Michelson Interferometer for Passive Atmospheric Sounding
<i>MLS</i>	Microwave Limb Sounder
<i>MODIS</i>	Moderate Resolution Imaging Spectroradiometer
<i>MOZAIC</i>	Measurement of OZone by Airbus In-service airCraft
<i>NASA</i>	National Aeronautics and Space Administration
<i>NetCDF</i>	Network Common Data Form
<i>OLCI</i>	Ocean and Land Colour Instrument
<i>PSD</i>	Product Specification Document
<i>RR</i>	Reduced Resolution
<i>SoW</i>	Statement of Work
<i>SPARC</i>	Stratosphere-troposphere Processes And their Role in Climate
<i>SPURT</i>	Spurenstofftransport in der Tropopausenregion (Trace gas transport in the tropopause region)
<i>SR</i>	System Requirement
<i>SRD</i>	System Requirements Document
<i>SSD</i>	System Specification Document
<i>SSM/I</i>	Special Sensor Microwave Imager
<i>SSMIS</i>	Special Sensor Microwave Imager Sounder
<i>SST</i>	Sea Surface Temperature
<i>SVR</i>	System Verification Report
<i>TCWV</i>	Total Column of Water Vapour
<i>TOA</i>	Top of Atmosphere
<i>UoR</i>	University of Reading
<i>URD</i>	User Requirements Document
<i>VRWV</i>	Vertically Resolved Water Vapour
<i>WAVAS</i>	Water Vapour Assessment
<i>WV</i>	Water Vapour



***End of Document***