System Specification Document (SSD v2)

Version 2.0

06 December 2013

Prepared by

TU Wien, VUA, GeoVille, ETH Zürich, AWST, FMI, UCC and NILU

This document forms deliverable D5.3 and was compiled for the ESA Climate Change Initiative Phase 1 Soil Moisture Project (ESRIN Contract No. 4000104814/11/I-NB). For more information on the CCI programme of the European Space Agency (ESA) see http://www.esa-cci.org/.

Number of pages: 111

Authors:	M. Ertl, A. Boresch		
Circulation (internal):	Project consortium and science partners		
External:	ESA		
Issue	Date	Details	Editor
0.1	2012-11-29	Document template with annotated TOC.	A. Boresch
0.2	2013-06-14	Added introduction and references (sections 1 and 2). Included descriptions of architectural factors and design decisions from architectural design workshops (sections 3 and 4).	M. Ertl , A. Boresch
0.3	2013-06-18	Peer review of sections 1-4.	A. Boresch
0.4	2013-06-24	Added specification of the System's architecture in terms of architectural views (section 5, views 1-5 and 7).	M. Ertl
0.5	2013-06-24	Addressed comments received from TUW.	M. Ertl
0.6	2013-06-25	Peer review of section 5.	A. Boresch
0.7	2013-06-27	Addressed comments received from peer review and from EP4SM project.	J. Strasser, M. Ertl
1.0	2013-06-28	Finalization for delivery of version 1.	M. Ertl, A. Boresch
1.1	2013-10-16	Prepared baseline for version 2.	M. Ertl
1.2	2013-11-11	Added traceability matrix (section 6).	M. Ertl
1.3	2013-11-18	Added life cycle sustainment activities (sec. 5.8).	A. Boresch
1.4	2013-11-20	Added information view (section 5.6) and submitted to TUW for review.	M. Ertl
1.5	2013-11-25	Added ECV change process (section 5.4.4) and conceptual schemata (sections 5.6.4 and 5.6.5). Peer review of section 5.8.	M. Ertl
1.6	2013-11-26	Peer reviews of sections 5.4.4, 5.6 and 6.	A. Boresch
1.7	2013-11-27	Prepared draft version for presentation at TUW on November 28, 2013.	M. Ertl, A. Boresch
1.8	2013-12-04	Addressed comments received from TUW.	M. Ertl
2.0	2013-12-06	Finalization for delivery of version 2.	M. Ertl, A. Boresch

For any clarifications please contact Martin Ertl (ertl@awst.at).

Table of Contents

1 INTRODUCTION 1
1.1 PURPOSE AND SCOPE OF DOCUMENT 1
1.2 DOCUMENT OVERVIEW 1
1.3 TARGET AUDIENCE 2
1.4 TERMINOLOGY USED IN THIS DOCUMENT 2
2 REFERENCES 5
2.1 APPLICABLE DOCUMENTS 5
2.2 REFERENCE DOCUMENTS 5
3 ARCHITECTURAL FACTORS 10
3.1 SYSTEM PURPOSE AND INTENDED USE 10
3.2 OVERVIEW OF SYSTEM CONTEXT 12
3.3 MAIN EXTERNAL INTERFACES 14
3.4 INSTALLATION SITE CHARACTERISTICS 14
3.5 DESIGN CONSTRAINTS 15
3.6 CANDIDATE SYSTEM ARCHITECTURES 19
3.7 DECISION METHODS 20
4 DOCUMENTATION OF ARCHITECTURAL DESIGN DECISIONS 23
4.1 GENERAL CONSIDERATIONS 23
4.2 CCI COMMON ELEMENTS AND COLLABORATION 29
4.3 DATA PROCESSING SOFTWARE 32
5 SYSTEM ARCHITECTURE 36
5.1 FUNCTIONAL VIEW 38
5.2 ORGANIZATION VIEW 40
5.3 SOFTWARE VIEW 42
5.3.1 Application Software 42
5.3.2 Platform Software 44
5.3.3 Platform Composition 46
5.4 PROCESS VIEW 51
5.4.1 Data Acquisition Process 51
5.4.2 ECV Production Process 53
5.4.3 Software Change Process 55
5.4.4 ECV Change Process 58
5.5 USE-CASE VIEW 61
5.6 INFORMATION VIEW 64
5.6.1 Context Description: Soil Moisture Data from Earth Observation Satellites 65
5.6.2 External Schema: Terminology of Remote Sensing and Earth Observation 78
5.6.3 External Schema: Data Processing Levels 83
5.6.4 Conceptual Schema of Soil Moisture Data Items 85
5.6.5 Conceptual Schema of System Monitoring & Control Information Items 88
5.7 INFRASTRUCTURE VIEW 91
5.8 LIFE CYCLE SUSTAINMENT ACTIVITIES 94
5.8.1 Sustainment of System Operations 94
5.8.2 Sustainment of System Maintenance 94
5.8.3 Sustainment of System Evolution 95
6 TRACEABILITY MATRIX 97

List of Figures

FIGURE 1: SYSTEM FUNCTIONS HIERARCHY 38
FIGURE 2: ORGANIZATIONAL CHART OF THE SOIL MOISTURE ECVPS 40
FIGURE 3: SOFTWARE CONFIGURATION HIERARCHY FOR APPLICATION SOFTWARE 43
FIGURE 4: SOFTWARE CONFIGURATION HIERARCHY FOR PLATFORM SERVICES 44
FIGURE 5: SOFTWARE CONFIGURATION HIERARCHY FOR GEO LIBRARIES 46
FIGURE 6: PLATFORM COMPOSITION BY ENVIRONMENTS (ENVS) AND VIRTUAL MACHINES (VMS) 47
FIGURE 7: ACTIVITY FLOW FOR ACQUISITION OF L2 DATA AND ANCILLARY DATA 52
FIGURE 8: ACTIVITY FLOW FOR THE PRODUCTION OF ECV DATA PRODUCTS 54
FIGURE 9: ACTIVITY FLOW FOR RESOLUTION OF SOFTWARE ISSUES, SOFTWARE DEVELOPMENT AND TESTING 56
FIGURE 10: ACTIVITY FLOW FOR RESOLUTION OF ECV DATA PRODUCT ISSUES, AND PRODUCT DEVELOPMENT 59
FIGURE 11: MICROWAVE INSTRUMENTS USED FOR THE GENERATION OF SOIL MOISTURE ECV DATA PRODUCTS 66
FIGURE 12: BEAMS AND WIDE SWATHS OF AMI-WS AND ASCAT INSTRUMENTS 68
FIGURE 13: WIDE SWATHS AND DAILY GLOBAL COVERAGE OF AMI-WS (LEFT) AND ASCAT (RIGHT) 68
FIGURE 14: CONSTRUCTION OF THE WARP 5 GLOBAL GRID USED FOR ASCAT L2 DATA 72
FIGURE 15: A) GEODESIC ISEA4H GRID OF SMOS; B) REGULAR LATITUDE-LONGITUDE GRID USED FOR SM ECV 72
FIGURE 16: PRINCIPLE DATA FLOW FOR THE GENERATION OF SOIL MOISTURE ECV DATA PRODUCTS 74
FIGURE 17: SOIL MOISTURE CLIMATOLOGY FOR THE MONTH AUGUST 76
FIGURE 18: CONCEPTUAL SCHEMA OF SOIL MOISTURE DATA ITEMS 86
FIGURE 19: CONCEPTUAL SCHEMA OF SYSTEM MONITORING & CONTROL INFORMATION ITEMS 88
FIGURE 20: CONCEPTUAL SCHEMA FOR ACTIVITY AND PROCESS MANAGEMENT 90
FIGURE 21: CLOUD INFRASTRUCTURE FOR DEPLOYMENT OF THE SOIL MOISTURE ECVPS 92

List of Tables

Table 1: Technical Terms ... 2
Table 2: Mission of the Soil Moisture ECV Production System (SM ECVPS) ... 10
Table 3: Objectives of the Soil Moisture ECVPS ... 10
Table 4: List of Architectural Factors for consideration in the architectural analysis ... 15
Table 5: List of Trade-off Issues and Candidate Solutions ... 19
Table 6: Architectural Design Team ... 21
Table 7: Template for Technical Memos (TMs) ... 23
Table 8: (TM-001) System Specification Approach ... 23
Table 9: (TM-002) Functional Scope ... 24
Table 10: (TM-003) Requirements Maturity Status ... 25
Table 11: (TM-004) Development Capability ... 26
Table 12: (TM-005) ECV Dependency ... 27
Table 13: (TM-006) Standardization ... 28
Table 14: (TM-007) Sustainment ... 28
Table 15: (TM-008) CCI Web Portal ... 29
Table 16: (TM-009) Processing Environment ... 30
Table 17: (TM-010) L1 Data Archive ... 31
Table 18: (TM-011) Version Control ... 31
Table 19: (TM-012) ECV Processor Architecture ... 32
Table 20: (TM-013) ECV Processor Programming Language ... 33
Table 21: (TM-014) ECV Processor Performance ... 35
Table 22: Main Use Cases of the Soil Moisture ECVPS ... 61
Table 23: Template for Use Cases (UCs) ... 61
Table 24: (UC-001) Download ECV Data Product ... 62
Table 25: (UC-002) View ECV Data Product ... 62
Table 26: (UC-003) Compare ECV Data Products ... 63
Table 27: Color Code used in ER Schemata for different Kinds of Entity Types ... 65
Table 28: Grid Systems used for Calibrated Measurement Data and SSM Geophysical Parameter Data ... 73
Table 29: Business Domain Terminology: Remote Sensing, Earth Observation, Soil Moisture ... 78
Table 30: Description of Data Processing Levels ... 83
Table 31: Allocation of System Requirements to System Functions ... 97

List of Acronyms

AWST	Angewandte Wissenschaft Software und Technologie GmbH
CCB	Configuration Control Board
CCI	Climate Change Initiative
CDR	Climate Data Record
CEOS	Committee on Earth Observation Satellites
CF	Climate and Forecasting
CMUG	Climate Modelling User Group
CNES	Centre National d'Etudes Spatiales
CPU	Central Processing Unit
DMZ	Demilitarized Zone
DVCS	Distributed Version Control System
DVL	Development
ECMWF	European Centre for Medium Range Weather Forecasting
ECV	Essential Climate Variable
ECVPS	ECV Production System
ENV	Environment
ESA	European Space Agency
FAQ	Frequently Asked Questions
FCDR	Fundamental Climate Data Record
FTE	Full-Time Equivalent
FTP	File Transfer Protocol
GbE	Gigabit Ethernet
GCOS	Global Climate Observing System
GEO	Group on Earth Observation
GEWEX	Global Energy and Water Cycle Experiment
GTOS	Global Terrestrial Observing System
IaaS	Infrastructure as a Service
IDL	Interactive Data Language
IGWCO	Integrated Global Water Cycle Observations

ISMWG	International Soil Moisture Working Group
ISO	International Organization for Standardization
IT	Information Technology
L1b	Level 1b
L2	Level 2
L3s	Level 3 super-collated
LAN	Local Area Network
LPRM	Land Parameter Retrieval Model
LVP GEO	Land Product Validation
MTBF	Mean Time Between Failures
NASA	National Aeronautics and Space Administration
OPS	Operations
OS	Operating System
PaaS	Platform as a Service
PC	Personal Computer
PCB	Product Control Board
PECVPS	Prototype ECV Production System
PSD	Product Specification Document
RAM	Random Access Memory
SaaS	Software as a Service
SAD	Software Architecture Document
SM	Soil Moisture
SM ECVPS	Soil Moisture ECV Production System
SM PECVPS	Soil Moisture Prototype ECV Production System
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture and Ocean Salinity
SRD	System Requirements Document
SSD	System Specification Document
TBD	To Be Determined
TCDR	Thematic Climate Data Record
TM	Technical Memo

TOPC	Terrestrial Observation Panel on Climate
TST	Test
TUW	TU Wien
UP	Unified Process
URD	User Requirements Document
VCS	Version Control System
VM	Virtual Machine
VOD	Vegetation Optical Depth
VPN	Virtual Private Network
VUA	Vrije University of Amsterdam
WARP	soil WAter Retrieval Package
WCRP	World Climate Research Programme
WGS	World Geodetic System
WOAP	WCRP Observations and Assimilation Panel

1 Introduction

1.1 Purpose and Scope of Document

This System Specification Document (SSD) specifies the architecture of the Soil Moisture (SM) Essential Climate Variable (ECV) Production System (conveniently abbreviated "SM ECVPS" and generally termed "the System" in this document). While the SM ECVPS will be developed and implemented in the ESA Climate Change Initiative (CCI) Phase 2, the SSD is a deliverable under the ESA CCI Phase 1 Soil Moisture project [PROP1].

The document fulfills the following purposes:

- The SSD describes the operational context and main interfaces of the SM ECVPS, and specifies the System's architecture in terms of complementary architectural views on the basis of system and user requirements, as specified in [SRD] and [URD].
- The SSD documents the architectural factors (requirements, constraints and assumptions) that influenced architectural design decisions. It provides records of the trade-off analyses and the rationales of design decisions.
- The SSD forms the main technical input to the System construction and Subsystem development to be performed under the planned ESA CCI Phase 2 Soil Moisture project, and provides guidance for the project setup and planning.

Scope considerations:

- The SSD does not include a work breakdown, or an implementation plan or funding plan for the System's development, construction and operations. These are expected to be covered elsewhere.
- The document does not repeat the technical specifications of Soil Moisture ECV Products and the ECV Processor Subsystem, as those are specified in [PSD], [DPM] and [IODD].

1.2 Document Overview

The document contains six major sections:

- Section 1: Introduction to the document (this section).
- Section 2: Listing of applicable and referenced documents.
- Section 3: Descriptions of purpose and intended use of the System, its operational context and the Architectural Factors (requirements, constraints and assumptions) that influenced architectural design decisions.

- Section 4: Records of the architectural trade-off analyses and design decisions in the form of Technical Memos (TMs).
- Section 5: Specification of the System's architecture in terms of complementary architectural views.
- Section 1: Tracing of requirements to System architecture elements.

1.3 Target Audience

The prime audience are the CCI management team and the Soil Moisture CCI Phase 2 project team.

Within the systems engineering process of the Soil Moisture CCI project, it will form the basis for the construction of the SM ECVPS and the development of its Subsystems under CCI Phase 2.

During these phases and beyond the SSD will form a lasting reference of the original architectural design, and will help future system maintainers and developers to understand the systems origins and the rationale behind the design decisions made.

Another target audience are the system engineering teams for ECV Production Systems of other CCI projects, who may learn about differences and similarities between the ECV Production Systems in order to judge the feasibility of designing and implementing common components that could be used by more than one ECVPS.

1.4 Terminology Used in this Document

Terms that are used with a specific technical meaning in this document are defined in Table 1.

Table 1: Technical Terms

Application Software	Subset of a vertically broadly scoped Information System: combination of all software components required to fulfill the System's responsibilities specific to its business domain.
Architectural Factor	Requirements, constraints or assumptions that have an impact on the architectural design.
Architectural View	Reduced design description of a System's architecture that puts the focus on a specific aspect like the System's static or logical structure, its main control flows, or its essential components, etc; the complete architectural design is then obtained by mentally matching concurrent Architectural Views together.

Assumption	Statements about a property of or a process in the Environment, on which a System depends.
Configuration Control	Process of controlling System Elements, covering hardware, software, configuration, data and documentation, their identification by versioning (revision control); and the process of evaluating change requests for approval or disapproval.
Configuration Control Board	Group of Users who are responsible for evaluating and approving proposed changes to a System.
Constraint	Statement constraining the design of a System that derives from a User or a given property of the Environment.
Data Product	Packaged data with identity.
Data Provider	User providing input data to a System. External to the System.
Data User	User who is using a System's Data Products. External to the System.
Design Scope	Extent of design control when designing a System.
ECV Data Product	ECV Data with identity. Disseminated as a coherent set of data files, as specified in [DSWG].
ECV Data	Observations combined from multiple instruments into a space-time grid. Synonyms are Super-collated (L3s) Data and TCDR.
ECV Processor	System Element providing data processing capabilities.
End-User	User who is using a System but not in a role of administrating, maintaining or developing it.
Environment	Everything outside a System's Boundary.
Infrastructure	Basic services and facilities necessary for a System to function.
Interface	Point of contact between a System and its Environment, reaching across the System Boundary.
L1b Data	Satellite data that have been processed to sensor units. Synonym is FCDR.
L2 Data	Retrieved environmental variables at the same resolution and location as the L1b Data.
Platform	System composed of a set of Software Environments and Virtual Machines (VMs) providing services for software development, testing and operation.
Processing Algorithms	Algorithms implemented by a System's Processor Modules.

Product Control Board	Group of Users who are responsible for evaluating and approving proposed changes to Data Products.
Product Specification	Specification of a Data Product forming the basis for production and use of the Data Product.
Regulatory Organizations	Elements of a System's Environment with controlling or enabling impact on the System.
Requirement	Statement constraining the design of a System that derives from an expressed need of a User of the System.
Resampling	Interpolation and extrapolation of data points as needed to pass data sampled on a specific temporal and spatial grid to another one.
Rescaling	Normalization technique employed in processing of Soil Moisture data.
Scientific Methods	Methods describing the science aspects of Processing Algorithms.
Software Component	System Element constituted by computer software.
Software Configuration Hierarchy	Hierarchical Configuration of Software Components.
Software Environment	Environment for developing, testing or running software applications.
Subsystem	System Element at the first decomposition level of a System.
System	Set of interacting or interdependent components to be specified, designed and constructed, separated from its Environment by a boundary (the "System Boundary").
System Boundary	Boundary between a System and the rest of the world (its Environment).
System Element	Part of a System at any decomposition level.
System Function	Range of activities to be performed by a System; a System Function may be accomplished by one or more System Elements comprised of personnel, hardware, software, data and facilities; System Functions are identified by functional analysis, and the whole set of System Functions must match up to the System's purpose.
System Team	System Element constituted by the group of Users who operate, maintain or develop the System.

2 References

2.1 Applicable Documents

The documents outlined here detail the scope and focus for the work that is reported in this document.

[CNTR1] Phase 1 of the ESA Climate Change Initiative Soil-Moisture-cci. ESRIN Contract No: 4000104814/11/I-NB.

[KO1] Phase 1 of the ESA Climate Change Initiative Soil-Moisture-cci. Kick-Off Meeting, Minutes of Meeting, 18.01.2012, ESA_CCI_SoilMoisture_KO_17_180112.doc

[NEG1] Phase 1 of the ESA Climate Change Initiative Soil-Moisture-cci. Contract Negotiation and Clarification Meeting, Minutes of Meeting, 21.10.2011, ESA_CCI_SoilMoisture_MINUTES_contract_meeting_21102011.doc

[PROP1] Proposal (Parts 1-5) in response to ESA Climate Change Initiative Phase 1 ESRIN/AO/1-6782/11/I-NB, Vienna University of Technology.

[SOW1] ESA Climate Change Initiative Phase 1, Statement of Work for Soil Moisture and Ice Sheets, European Space Agency, EOEP-STRI-EOPS-SW-11-0001.

[SOW1-SM] ESA Climate Change Initiative Phase 1, Statement of Work for Soil Moisture and Ice Sheets, Annex L: Soil Moisture ECV (Soil_Moisture_cci), European Space Agency, annex to EOEP-STRI-EOPS-SW-11-0001.

2.2 Reference Documents

This section provides a list of reference documents upon which this document is either based, or to which this document refers.

[Ahson11] Cloud Computing and Software Services: Theory and Techniques, editors Syed A. Ahson and Mohammad Ilyas, CRC Press, 2011.

[ATBD] Algorithm Theoretical Baseline Document, Version 1, 14 February 2013, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20130214_CCI_Soil_Moisture_ATBD1.v1.0.pdf and 20130214_CCI_Soil_Moisture_ATBD1_Annex1.v1.0.pdf

[CCI-SRD] ESA CCI-wide System Requirements Document, at the time of writing in production by ESA and not available to the authors; as substitute referring to collection of loose papers distributed to the SEWG during April to September 2012: (FR) CCI Functional Requirements etc v2.docx, (IR) Interface

Requirements_0_3.docx, (RR) Reliability Requirements_0_2.doc, (N) SEWG 04_09_12 notes.docx

[CECR] Comprehensive Error Characterisation Report, Version 0.7, 21 May 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20120521_CCI_Soil_Moisture_D1.2.1_CECR_v.0.7.pdf

[Chen76], Peter Chen, The Entity-Relationship Model – Toward a Unified View of Data, ACM Transactions on Database Systems, Vol. 1, No. 1, March 1976, pp. 9-36, http://www.csc.lsu.edu/news/erd.pdf

[Colloc1] ESA CCI First Collocation. Presentation: System Engineering Recommendations from Drafting Team, SystemEngineeringRecom.pdf.

[Colloc2] ESA CCI Second Collocation Report, CCI-PRGM-EOPS-TN-11-0029, Issue 1, Revision 0, Pascal Lecomte and Cat Downy, February 7th, 2012, CCISecondCollocationReport.pdf

[DARD] Data Access Requirements Document, Version 1.2, 13 April 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20120413_CCI_Soil_Moisture_D1.3_DARD_v.1.2.pdf

[DSWG] ESA, Guidelines for Data Producers – Climate Change Initiative Phase 1, CCI-PRGM-EOPS-TN-11-0003, Issue 3, Revision 1, Victoria Bennett, 20/07/2012, CCI_Data_Standards_Guidelines_v5_July2012.pdf

[DPM] Detailed Processing Model, Version 1.2, 26 November 2013, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20131126_CCI_Soil_Moisture_D2.7_DPM_v.1.2.pdf

[ECSS-E-ST-40C] European Cooperation for Space Standardization (ECSS), Space Engineering Software, ECSS-E-ST-40C, ECSS Secretariat, ESA-ESTEC, Requirements & Standards Division, Noordwijk, The Netherlands, 6 March 2009.

[EP4SM-SAD] Exploitation Platform Demonstrator for Soil Moisture (EP4SM), System Architecture Document, AWST-TR 13/04, ESA-ESRIN, Version 1.0, April 8, 2013.

[Hollmann13] R. Hollmann and coauthors: The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. Bull. Amer. Meteor. Soc., 94, 2013, 1541–1552, http://www.esa-soilmoisture-cci.org/node/163

[IDEF1X] Integration Definition for Information Modeling (IDEF1X), Federal Information Processing Standards (FIPS) Publication 184, December 21, 1993, http://www.itl.nist.gov/fipspubs/idef1x.doc

[IODD] Input/Output Data Definition Document, Version 1.0, 21 December 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20121221_CCI_Soil_Moisture_D2.8_IODD_v1.0.pdf

[JCGM-GUM] Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM), 1st edition, September 2008, JCGM, JCGM 100:2008, http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

[JCGM-VIM] International vocabulary of metrology – Basic and general concepts and assocated terms (VIM), 3rd edition, 2012, JCGM, JCGM 200:2012, http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf

[Jeu12a] R. de Jeu, W. Dorigo, R. Parinussa, W. Wagner and D. Chung: Soil Moisture, State of the Climate in 2011, Special supplement to the Bulletin of the American Meteorological Society, Vol. 93, No.7, 2012, pp.30-34, http://www.esa-soilmoisture-cci.org/node/128

[Jeu12b] R. de Jeu, W. Dorigo, R. Parinussa, W. Wagner, Y Liu, D. Chung and D. Fernandez-Prieto: Sidebar 2.2: Building a climate record of soil moisture from historical satellite observations, State of the Climate in 2011, Special supplement to the Bulletin of the American Meteorological Society, Vol. 93, No.7, 2012, pp.32-33, http://www.esa-soilmoisture-cci.org/node/128

[Kidd05] R. Kidd: NWP Discrete Global Grid Systems, ASCAT Soil Moisture Report Series No. 4, Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria, December 20, 2005.

[Kruchten95] Architectural Blueprints – The 4+1 View Model of Software Architecture, Philippe Kruchten, IEEE Software 12 (6), pp. 42-50, November 1995.

[Larman02], Applying UML and Patterns – An Introduction to Object-Oriented Analysis and Design and the Unified Process, Craig Larman, PH PTR, 2nd Ed., 2002.

[NASA-EOS86] Report of the EOS data panel, Earth Observing System, Data and Information System, Data Panel Report, Vol. IIa., NASA Technical Memorandum 87777, June 1986, NASA, http://hdl.handle.net/2060/19860021622

[NETCDF] NetCDF Climate and Forecast (CF) Metadata Conventions: Version 1.6, 5 December, 2011, Brian Eaton, Jonathan Gregory, Bob Drach, Karl Taylor, and Steve Hankin.

[PSS-05-0] ESA PSS-05-0 Software Engineering Standards, Issue 2, ESA Board for Software Standardisation and Control (BSSC), ESA-ESTEC, February 1991.

[PSS-05-03] ESA PSS-05-03 Guide to the software requirements definition phase, Issue 1, Revision 1, ESA Board for Software Standardisation and Control (BSSC), ESA-ESTEC, March 1995.

[PSD] Product Specification Document, Version 1.0, 26 November 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20121126_ECV_SM_Product_Specification_Document_1.0.pdf

[PVP] Product Validation Plan, Version 1.1, 30 August 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20120830_CCI_Soil_Moisture_D2.1_PVP_v.1.1.pdf

[RM-ODP98] Information technology – Open Distributed Processing: Reference Model – Part 1: Overview, ISO/IEC 10746-1:1998, International Organization for Standardization, Geneva, Switzerland, 1998.

[Shames06]. Toward a Framework for Modeling Space Systems Architectures, Peter Shames and Joseph Skipper, NASA, JPL, 2006.

[SMDC] Proposal in response to Cooperative R&D Projects, Industrial Research, Experimental Development of the Austrian Research Promotion Agency (FFG): Soil Moisture Data Cubes, AWST, Version 1.3, 17 October 2012.

[SPD] System Prototype Description, Version 1.0, 04 July 2013, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, CCI_Soil_Moisture_D3.1_SPD_v.1.0.pdf

[SRD] System Requirements Document, Version 1.0, 21 December 2012, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, 20121221_CCI_Soil_Moisture_D5.1_SRD.pdf

[SVR] System Verification Report, Version 1.0, 02 August 2013, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, CCI_Soil_Moisture_D3.2_SVR_v.1.0.pdf

[UML] OMG Unified Modeling Language (OMG UML), Superstructure, Object Management Group, Version 2.4.1, August 2011, http://www.omg.org/spec/UML/2.4.1/Superstructure/

[URD] User Requirements Document, Version 2.0, 17 September 2013, ESA Climate Change Initiative Phase 1 Soil Moisture Project, http://www.esa-soilmoisture-cci.org/node/119, CCI_Soil_Moisture_D1.1_URD_v.2.0.pdf

[Wagner12] W. Wagner, W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, M. Ertl: Fusion of active and passive microwave observations to create an Essential Climate Variable data record on Soil Moisture, ISPRS Annals of the

Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals), Volume I-7, XXII ISPRS Congress, Melbourne, Australia, 25 August – 01 September 2012, pp.315-321, http://www.esa-soilmoisture-cci.org/node/137

[Zoltan09] B. Zoltan: Spaceborne Scatterometers for Change Detection over Land, Dissertation, Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Vienna, December 14, 2009.

3 Architectural Factors

3.1 System Purpose and Intended Use

The Soil Moisture ECV Production System (SM ECVPS, termed “the System” in this document) serves a single purpose expressed by the statement of Table 2, which we adopt as the System’s Mission.

Table 2: Mission of the Soil Moisture ECV Production System (SM ECVPS)

Mission Statement
Produce the most complete and consistent within scientific and observational limits, global, long-time continuous, multi-sensor, error characterized and verified Soil Moisture Data Products, from satellite and in-situ measurements, for use in Climate Research and Modeling.
Note: In the context of Soil Moisture “global” means covering all Earth land surfaces, and “long-time continuous” means beginning with the first available satellite measurements in 1978, since then passing from one satellite generation with radiometer (passive) or scatterometer (active) sensors to the next and aiming to extend to future satellite generations.

This mission statement was inspired from [SOW1], CR-2. On its basis we develop a short list of prime functions complementing the Mission Statement and necessary for the System to provide in order to achieve its Mission. These functions are stated in Table 3 and constitute the System Objectives.

Table 3: Objectives of the Soil Moisture ECVPS

OBJ-#	System Objective
OBJ-01	Produce Soil Moisture ECV Data Products that meet the requirements as specified in [URD], “Key requirements for an ECV soil moisture data set”.
OBJ-02	Continuously develop and improve ECV Data Products, Scientific Methods and Processing Algorithms, a) to update to changing and steadily becoming more demanding requirements by data users, and b) to extend the ECV Data Products by the data from planned future satellite missions.
OBJ-03	Acquire input data of sufficient quality for the production and development of ECV Data Products.
OBJ-04	Preserve input data, intermediate data and produced ECV Data Products for future generations.
OBJ-05	Preserve know how and production capability for future generations.

This short list of System Objectives will be extended to the full System Functions Hierarchy in section 5.1, which provides a logical, functional view onto the System, defining the System in terms of System Functions.

In order to achieve the full System specification the stated System Objectives are decomposed into System Functions in section 5.1, and the latter are mapped to System Elements, as specified in the other Architectural Views of section 5. This decomposition and specification process is based on the requirements as specified in [SRD] and [URD], which provide input at the required level of detail for this purpose.

In addition to the logical view of the static System Functions there are two architectural views describing the dynamic behavior of the System in sections 5.4 and 5.5. The first one, called the Use-Case View is a purely outside view describing interactions of outside Actors with the System, while the second, the Process View considers processes that are running inside the System and may or may not have interactions with the outside world (the System Environment). Naturally, the Use-Case View depicts the System as a single entity (as a "black box") while it's interior is exposed (as a "white box") in the Process View.

There are three main Use Cases:

- Download ECV Data Product
- View ECV Data Product
- Compare ECV Data Products

In each Use Cases the Main Actor is a Data User, i.e. any person who intends to use ECV Data Products. For this purpose the Data User interacts with the System to retrieve a particular version, a subset, or several of the ECV Data Products that have been produced, published and kept available by the System.

In order to serve Data Users with the ECV Data Products the following main System processes (Operational Scenarios) are required:

- Data Acquisition Process
- ECV Production Process
- Software Change Process
- ECV Change Process

The main System processes are under control of two boards, Product Control Board (PCB) and Configuration Control Board (CCB), which are composed of internal and external System Stakeholder representatives, who received their mandate by the System Sponsors. The boards are the main decision makers, the PCB with a focus on Data User requirements and on ECV Data Products, the CCB on keeping the System operational in the day-to-day business and on maintaining its operational capabilities in the long run.

The Operational Scenarios and Use Cases are further detailed in the Process and Use-Case Views in sections 5.4 and 5.5, respectively.

3.2 Overview of System Context

This section provides a description of the System's operational context. Context elements are generally considered to be outside of the System. In the following paragraphs we present those context elements that are noteworthy because of their specific relationship with the System.

Data Users. They constitute the target group and prime stakeholders for Soil Moisture ECV Data Products produced by the System. Besides being consumers of disseminated products and any derived observations, Data Users add value to System products – and they are the prime source of product related requirements. A comprehensive survey and analysis of Data Users differentiating by Societal Benefit Area (SBA) was conducted in [URD].

Science Communities. Specifically relevant are the international Earth Observation (EO) (GEO/CEOS), Climate Research and Modeling (CMUG, GCOS, etc.), Soil Moisture (TUW, VUA, NASA and CNES), and Remote Sensing communities. Their relationships to the System are manifold: They are consumers of ECV Data Products and thus constitute a part of the Data User group described above, but they also act as regulatory and standardization organizations and thus are an additional source for product related requirements. And most importantly, they hold the scientific assets, namely the data, scientific methods, algorithms, program code and know how required for the production and correct use of ECV Data Products as well as additional observations and conclusions that may be derived thereof.

Software Developers and IT Specialists. Management and processing of satellite data draws heavily upon capabilities of Information Technology (IT). Current scientific achievements and future improvements in the area of ECV Data Products with respect to performance, quality and stability of their production are intimately linked with the available software and hardware technologies. IT was in the past decades and will continue to be in the future a quickly evolving field, bringing with it new, today undiscovered possibilities – but these can only be used if the System adapts and incorporates the state-of-the art and future developments from this field. Software Developers and IT Specialists are those who turn the scientific assets developed by the Science Communities into professional processing systems for operational use.

Data Providers. They provide the L2 Data and Ancillary Data which are the necessary input for the production of Soil Moisture ECV Data Products. Data Providers are EUMETSAT, NASA, TUW and VUA, after integration of SMOS data also ESA and CNES[1]. L2 Data Providers typically also process the data from lower levels, e.g. from L1b to L2, a process known as "Soil Moisture Retrieval" [ATBD]. Data Providers are highly specialized on specific sensors and retrieval technologies, they have operational systems and infrastructures in place and are distributed to different locations and countries. Data Providers in the wider sense are also Satellite Ground Segments and Satellite Missions who are feeding data into the Soil Moisture Retrieval systems[2]. The System as specified herein extends the Soil Moisture Retrieval systems and any feeding Satellite Ground Segments in a modular fashion to a super system, benefitting from existing infrastructures and experiences of established Data Providers but otherwise performing independently.

ECV Production Systems (ECVPS). Thirteen ECVPS are currently being specified under ESA CCI Phase 1, one of them being the Soil Moisture ECVPS ("the System"). Additional ECVPS may be initiated in the future. The System as specified herein is designed to be independent from other ECVPS, their internal data and output data products[3]. Various architecturally relevant factors in the context of ECV systems were identified[4] and are addressed in the architectural design memos in section 4. The future operational relationships between the ECVPS are unknown as of today.

Hosting Organization. The organization to host (operate) the System during its operational lifetime, in accordance with [SRD] SR-0680, System Hosting. Possible Hosting Organizations are the L2 Data Providers (TUW, VUA), National Data Centers (EODC-Water, ZAMG), and operators of Satellite Ground Segments (ESA, EUMETSAT). All except EODC-Water are established organizations with existing operational infrastructures. The selection will be made by ESA in coordination with its member states and is not under control of the project team. Once selected the Hosting Organization may impose additional system requirements. The specific technical environment of the installation site and further constraints imposed by the Hosting Organization are currently unknown. All established candidate organizations have made relevant investments in the past in building up their operational infrastructures and standards. The System will therefore have to be integrated with the existing infrastructure (SR-0580, System Integration). This issue was identified as an architecturally relevant factor (cf. Table 4, AF-005), and is addressed in the architectural design memos in section 4.

Sponsors. The Sponsors will ensure long-term funding and sustainment of the System's operational lifecycle. Soil Moisture ECV Data Products will be freely and publicly accessible in

---

[1] For a complete list of the input data see [SRD] section 5.1, for a detailed description see [DARD].
[2] See [DARD] for more detailed information.
[3] See [SRD] SR-0660 on ECV independence. In competition thereto are SR-0570, Cross-ECV Synergies, and SR-0590, SEWG Standards.
[4] These are AF-007, AF-008, AF-010, AF-011, AF-012, AF-013, AF-015, AF-024, AF-025, AF-026 in Table 4.

accordance with [SRD] SR-0640, Open Data. It is expected that the System will be sponsored either by the EU, by National States or by dedicated organizations to ensure long-term sustainment even for future generations (cf. Table 6, AF-003). The Sponsor will also be a relevant source of new requirements with an influence on the System design (cf. Table 6, AF-004). This aspect is addressed in the architectural design memos in section 4.

3.3 Main External Interfaces

This section provides a description of the System's main external interfaces.

The System has two main organizational interfaces to connect it with context elements at an organizational level. These interfaces are integrated with the System's steering boards:

- Product Control Board (PCB) – to survey and address the needs of those who have an interest in Soil Moisture ECV Data Products, i.e. Data Users, Science Community, Data Providers, Sponsors, and possibly other ECVPS.
- Configuration Control Board (CCB) – to involve those who share operational and maintenance responsibility, i.e. Hosting Organization (if not considered internal), external Software Developers and IT Specialists, and possibly other ECVPS.

Furthermore, the System has a set of technical interfaces for the following purposes:

- Access to ECV Data Products by Data Users (FTP download, on-line viewing, off-line or on-line data comparison and analysis tools).
- Data acquisition from Data Providers (human interfaces for manual transfers of L2 Data and Ancillary Data; software interfaces for automated transfers of L2 Data).
- Integration with Hosting Organization's operational infrastructure and/or other ECVPS (these may range from loose coupling to tight integration with existing hardware and software items).
- Tools for collaboration with Science Community, Software Developers and other ECVPS (CCI Web Portal, version control systems, etc.)

3.4 Installation Site Characteristics

The installation site has not yet been determined. The System is designed to be hosted (installed, operated and maintained) at a Hosting Organization, which may impose additional System requirements (see discussion in section 3.2). It will later have to be integrated with the future operational infrastructure of the selected organization.

Conversely, to get the requirements that the System imposes on the Hosting Organization and installation site, consider [SRD] SR-0429, Cooperation with Hosting Organization/s; SR-

0562, Technical Platform; and the requirements listed in [SRD] sections 6.5, 6.9, 6.11, 6.12 and 6.13.

3.5 Design Constraints

This section lists requirements and constraints with significant impact on the System's architecture. The main source of system requirements is [SRD], which was the starting point for the search and selection of Architectural Factors, presented in Table 4. The table collects and classifies system requirements and a few additional sources into Architectural Factors which are considered in the architectural analysis described in section 4.

Each entry of Table 4 has a unique AF-number, followed by the description of the Architectural Factor, and the references to the relevant system requirements (SR-number) and additional sources. Entries are sorted and categorized by thematic similarity. The complete requirements are described in [SRD].

Each Architectural Factor was assessed with respect to: (a) its potential impact on the system architecture ("Impact"), and (b) the difficulty to address it at the system architecture level and current project phase ("Difficulty"). Table 4 shows the results using the classifications H(igh) , M(edium) and L(ow).

The last column in Table 4 contains forward references to the Technical Memos (TMs) of section 4. The TMs provide a discussion of the issues and document the architectural decisions. Each TM typically refers to several Architectural Factors, though some Factors are addressed by several TMs.

Architectural Factors of difficulty H(igh) could not be satisfactorily resolved at the current project phase. The "solution" described in the pertinent TMs is only a preliminary workaround in these cases. A proper solution will have to be determined at a later stage and will typically require involvement of additional stakeholders.

Table 4: List of Architectural Factors for consideration in the architectural analysis

AF-#	Architectural Factor	Req.-Ref.	Impact	Difficulty	TM-Ref.
System Scope					
AF-001	System engineering approach for extended system scope.	[Colloc2]	H	L	TM-001

AF-#	Architectural Factor	Req.-Ref.	Impact	Difficulty	TM-Ref.
AF-002	Broad scope of L1 to ECV vs. focus of L2 to ECV processing chain.	SR-0670 SR-0040 SR-0420 SR-0421 SR-0050 SR-0425 SR-0060	H	L	TM-002
Long-Term Funding and System Sustainment					
AF-003	Long-term system sustainability, technically and financially.	[Colloc2 SR-0680 SR-0999	M	H	TM-007
Requirements Status					
AF-004	Sponsor of the operational system and thus sponsor requirements unknown.	[Colloc2]	H	H	TM-003
AF-005	Integration with a hosting organization.	SR-0429 SR-0680 SR-0580 SR-0562 SR-0870	H	H	TM-003 TM-013
AF-006	Missing formal review of ESA requirements.	[NEG1]	M	H	TM-001
AF-007	Common CCI system requirements [CCI-SRD] not yet complete.	[CCI-SRD]	L	M	TM-003
AF-008	National data security regulations applicable to hosting organization may vary.	SR-0869	L	L	TM-003
ECV Systems Relationships					
AF-009	Independence between ECVs.	SR-0660	H	L	TM-005
AF-010	Cross-ECV synergies.	SR-0570	M	M	TM-005 TM-003
AF-011	Common portal for ECV Data Products dissemination.	SR-0570	L	L	TM-008
AF-012	SEWG may provide recommendations for common tools and for a shared processing environment on a low level. The processing environment may also be provided by the future hosting organization and requirements may clash.	SR-0570 SR-0562	M	M	TM-009
AF-013	SEWG may provide recommendations for an L1 Data Archive.	[Colloc3]	L	L	TM-010
Standardization					

AF-#	Architectural Factor	Req.-Ref.	Impact	Difficulty	TM-Ref.
AF-014	ECSS standards not applicable for extended system engineering approach including organizational requirements; they lack description of processes and require tailoring.	SR-0610	L	L	TM-006
AF-015	Recommendations and standards from different sources need consolidation. Standardization among ECVs uncertain.	SR-0600 SR-0590 SR-0610	M	L	TM-005 TM-006
Changeability					
AF-016	Developing and perfecting ECV Data Products will be an ongoing and continuous effort.	SR-0969	H	L	TM-004
AF-017	Development of Scientific Methods and Processing Algorithms for L2 Data and ECV Data Products is ongoing and not expected to complete in the near future.	SR-0020 SR-0030 SR-0070 SR-0080 SR-0090 SR-0560 SR-0950 SR-0970 SR-0971 SR-0980 SR-0990	H	L	TM-004
AF-018	Changes in methods and algorithms need to be implemented as Processor Modules and transitioned into operations quickly and with the desired quality. New or improved algorithms shall become operational within 6 months.	SR-0560 SR-0971 SR-0190 SR-0520 [Colloc2]	H	M	TM-004
AF-019	Flexible processing chains (workflows) and standardization of workflow elements.	SR-0561 SR-0560 SR-0541	H	M	TM-013
AF-020	Future satellite missions need to be integrated and will put a steady stress on updating methods, algorithms and processing software, and producing new or revised ECV Data Products.	SR-0040 SR-0420 SR-0421 SR-0530	H	M	TM-004
AF-021	Capability for System Evolution required.	SR-0868 SR-0999 SR-1000 SR-1010 SR-1020 SR-1030	H	L	TM-004
Constraints					
AF-022	Preference for open source software.	SR-0650	M	M	TM-009

AF-#	Architectural Factor	Req.-Ref.	Impact	Difficulty	TM-Ref.
AF-023	Diverging requirements on programming language for the implementation of processor modules.	SR-0541 SR-0690 SR-0691	L	H	
Collaboration					
AF-024	Capabilities for open collaboration.	SR-0910	M	L	TM-011
AF-025	Configuration control for software and configuration items.	SR-0920 SR-0932 SR-0930 SR-0909	H	L	TM-011 TM-013
AF-026	Version control for data and data products.	SR-0931 SR-0500 SR-0510 SR-0520 SR-0180 SR-0181 SR-0969	H	L	TM-011 TM-013
Operations					
AF-027	Tools for process monitoring & control and anomaly recovery.	[SRD] Secs-6.1.9, 6.1.10 SR-0562	H	L	TM-013
AF-028	Support for several processing modes: incremental processing, full reprocessing, continuous interval processing (NRT), on-demand processing.	SR-0120 SR-0130 SR-0131	H	M	TM-013
AF-029	Quality, reliability, maintainability and evolution of processing software.	SR-0882 SR-0889 SR-1010 SR-1020 SR-0868 SR-0999	M	L	TM-013
Performance					
AF-030	Revised performance requirement for Processor Modules: 1 week for 1 year extension, and 3 weeks for full reprocessing for the current 32 year data set.	SR-0370 SR-0376	H	H	TM-014 TM-013

3.6 Candidate System Architectures

This section lists the candidate solutions which were considered in order to address the Architectural Factors listed in section 3.5. The Architectural Factors are thematically grouped into sets of architectural Trade-off Issues, and for each Issue potentially suitable Candidate Solutions are identified. The Candidate Solutions are evaluated in the Technical Memos (TMs) in section 4. Table 5 shows the list of Trade-off Issues and corresponding Candidate Solutions. The TM-numbers in the left column refer to the TMs in section 4 holding the evaluation results and stating the motivations (rationales) and design decisions made for each Trade-off Issue.

Table 5: List of Trade-off Issues and Candidate Solutions

TM-#	Trade-off Issue	Candidate Solutions
General considerations		
TM-001	System Specification Approach	A. Specification in terms of software and hardware only B. Full system engineering
TM-002	Functional Scope	A. L1 to ECV processing chain B. L2 to ECV processing chain tied up to existing systems
TM-003	Requirements Maturity Status	A. Use [SRD] and [URD] as requirements baseline B. Include requirements from Hosting Organizations: B1. TUW B2. VUA B3. EODC-Water B4. ZAMG B5. ESA B6. EUMETSAT C. Include requirements from Sponsors D. Include requirements from [CCI-SRD]
TM-004	Development Capability	A. Add development, maintenance and evolution capabilities B. Operations capability only
TM-005	ECV Dependency	A. Independency from other ECVPS B. Independency from others but allow cross-ECV synergies C. Reuse of other ECVPS elements
TM-006	Standardization	A. Setup standards under CCI Phase 1 B. Setup standards at beginning of CCI Phase 2 C. Setup standards at System startup
TM-007	Sustainment	A. Find long-term Sponsor: A1. EU A2. National States A3. Science Communities A4. Beneficiary Organizations B. Have system elements supporting sustainment C. Develop a Sustainment Plan under CCI Phase 2

TM-#	Trade-off Issue	Candidate Solutions
CCI common elements and collaboration		
TM-008	CCI Web Portal	A. Stand-alone Soil Moisture Web Portal B. Linkage integration with a light-weight CCI Web Portal C. Heavy-weight CCI Web Portal
TM-009	Processing Environment	A. Dedicated hardware and software infrastructure B. Shared infrastructure and common tools among ECVPS-s C. GRID processing infrastructure D. Use commercial Cloud processing infrastructure E. Heterogeneous Cloud processing architecture
TM-010	L1 Data Archive	N/A
TM-011	Version Control	Configuration control for software and configuration items: A. Centralized VCS (CVS, CVSNT, Subversion et al.) B. Distributed VCS (e.g. Git) Version control for data: C. File naming convention D. VCS for data E. Distributed and synchronized databases
Data processing software		
TM-012	ECV Processor Architecture	A. Monolithic ECV Processor B. Modular architecture with framework and modules
TM-013	ECV Processor Programming Language	A. Pure IDL implementation B. Pure C/C++ implementation C. Pure Python implementation D. Processing framework with multi-language support
TM-014	ECV Processor Performance	A. Stay with relaxed requirements of [SRD] B. Setup project to overcome performance barriers

3.7 Decision Methods

This subsection describes the process that was followed under Tasks 5.2 and 5.3 of the CCI Phase 1 Soil Moisture project in order to establish the System's architecture as specified herein by a structured and documented process for making the necessary architectural decisions.

The process for the creation of SSD version 1 under Task 5.2 has been as follows:

1. Created an annotated document template and submitted it for review to the consortium project managers and the ESA Climate Office.
2. Established the Architectural Design Team as shown in Table 6.

3. Performed an architectural analysis of system requirements specified in [SRD] and user requirements specified in [URD], and determined the architecturally relevant factors. They are listed in section 3.5.
4. Devised Candidate Solutions to address the AFs. Prepared Technical Memos (TMs) relating AFs to Candidate Solutions and aspects of the logical system architecture (i.e. the system functions and software configuration hierarchies) for discussion by the Architectural Design Team.
5. Conducted two architectural workshops at TUW on May 7 and 14, 2013 for discussion of Candidate Solutions on the basis of the TMs and an initial draft of the logical system architecture. The trade-off analysis and selection of preferred solutions was performed by the Architectural Design Team in the discussion of each TM at the workshops.
6. Updated the TMs to include the discussions, rationales and decisions from the architectural workshops. Included the eventual set of processed TMs in section 4.
7. Specified the System's architecture in terms of logical, process and use-case views in section 5.
8. Delivered SSD version 1 for review by consortium partners and ESA Climate Office.

Table 6: Architectural Design Team

Initial	Name	Affiliation
AB	Alexander Boresch	AWST
DC	Daniel Chung	TUW
ME	Martin Ertl	AWST
RK	Richard Kidd	TUW

The SSD was updated in the course of the development process under Task 5.3 of the CCI Phase 1 Soil Moisture project. Additional views and contents specifically added in version 2 are:

- Section 5.4.4: ECV Change Process,
- Section 5.6: Information View[5],
- Section 5.8: Life Cycle Sustainment Activities, and
- Section 6: Traceability Matrix.

[5] In collaboration with the Soil Moisture Data Cubes project [SMDC].

The process for the creation of SSD version 2 under Task 5.3 has been as follows:

1. Created a baseline document from SSD version 1 which had been approved by the ESA Climate Office.
2. Added the new sections listed above and conducted internal peer reviews of these sections.
3. Presented the draft SSD version 2 document on November 28, 2013 at TUW.
4. Addressed review comment from TUW.
5. Delivered SSD version 2 for review by consortium partners and ESA Climate Office.

4 Documentation of Architectural Design Decisions

This section provides the records from the architectural trade-off analyses. Each Trade-off Issue addresses a subset of the Architectural Factors (AF) listed in section 3.4. Typically, one Trade-off Issue addresses several Architectural Factors, however, some Architectural Factors are addressed by several Trade-off Issues. Records are provided in the form of Technical Memos (TM), as shown in Table 7. Each Technical Memo addresses one Trade-off Issue identified in section 3.6.

Table 7: Template for Technical Memos (TMs)

Technical Memo **Trade-off Issue: *Issue Name***	
Solution Summary:	*Brief summary of solution.*
ID	TM-#
Factors:	*Relevant Architectural Factors (AF-# referring to Table 4).*
Solution:	*Selected solution from the Candidate Solutions listed in Table 7, and additional details describing the solution.*
Motivation:	*List of motivations and constraints related to the solution.*
Unresolved Issues:	*List of outstanding issues at time of writing, if any.*
Alternatives Considered:	*Other Solutions that were considered and rejected (initial capital letters referring to Table 5).*
Discussion:	*Any other points.*

4.1 General Considerations

Table 8: (TM-001) System Specification Approach

Technical Memo **Trade-off Issue: System Specification Approach**	
Solution Summary:	Apply system engineering approach to extended system scope including organizational requirements on the basis of [SRD] and [URD].
ID	TM-001
Factors:	AF-001: System engineering approach for extended system scope. AF-006: Missing formal review of ESA requirements.
Solution:	B. Apply system engineering approach: Specify system functions at an abstract level and perform a functional decomposition that does not depend on the realization of a system function via human resources, software, hardware or other technical means or even a combination thereof. Specify the system in terms of architectural views. Equally address aspects of software, hardware and human organization as

Technical Memo **Trade-off Issue: System Specification Approach**	
	appropriate at the architectural level. Use the [SRD] and [URD] as basis for system specification.
Motivation:	There were different opinions on the desired approach to system specification: SOW and ESA TO generally favor a system scope, which is limited to the technical domain (i.e. software and hardware) and does not include organizational system aspects. ESA standard ECSS-E-ST-40C proposed by the ESA TO specifically applies to software systems. A larger system scope was requested by the CCI programme manager at [Colloc2] in order to create a system which includes organizational aspects and processes and has a built-in self-development capability. We follow this recommendation here and in [SRD]. Requirements from ESA were received in [SOW1] and at the Collocation Workshops [Colloc1] and [Colloc2], but a systematic requirements review process with ESA as a stakeholder was excluded by the CCI programme manager at [NEG1]. Our requirements gathering process collected relevant requirements into the [SRD] without explicit consultation of the ESA sources. [URD] and [SRD] have complementary content rather than [SRD] being derived from [URD] requirements, the system specification will therefore be based on both.
Unresolved Issues:	[SRD] and the system specification in this document may need to be revised if ESA provides feedback early in CCI Phase 2. A formal and systematic review process is recommended.
Alternatives Considered:	A. System specification solely in terms of software and hardware aspects. Such an approach would be too narrow and would miss organizational aspects and processes. The resulting system would not have built-in capabilities for self-development, maintenance and evolution, which are deemed necessary.
Discussion:	N/A

Table 9: (TM-002) Functional Scope

Technical Memo **Trade-off Issue: Functional Scope**	
Solution Summary:	System performs L2 to ECV processing and interfaces with L2 Data Providers.
ID	TM-002
Factors:	AF-002: Broad scope of L1 to ECV vs. focus of L2 to ECV processing chain.
Solution:	B. L2 to ECV processing chain tied up to existing systems: Implement the processing chain from L2 to ECV [SR-0670]. Interface with L2 Data Providers [SR-0040], [SR-0420], [SR-0421]. Interface with providers of ancillary data [SR-0050], [SR-0425], and initial system data [SR-0060].
Motivation:	Modularization of processing chain as described in [PROP1] part 3, section 3.6.2.1, and scope trade-offs made in [SRD] section 4.1.

Technical Memo **Trade-off Issue: Functional Scope**	
Unresolved Issues:	L1 to L2 processing chains to be handled elsewhere.
Alternatives Considered:	A. L1 to ECV processing chain; was excluded in [PROP1].
Discussion:	N/A

Table 10: (TM-003) Requirements Maturity Status

Technical Memo **Trade-off Issue: Requirements Maturity Status**	
Solution Summary:	Use [SRD] and [URD] as requirements baseline and assume no particular hosting organization or operator. Specify an abstract system architecture where specific requirements are incomplete.
ID	TM-003
Factors:	AF-004: Sponsor of the operational system and thus sponsor requirements unknown. AF-005: Integration with a hosting organization. AF-006: Missing formal review of ESA requirements. AF-007: Common CCI system requirements [CCI-SRD] not yet complete. AF-008: National data security regulations applicable to hosting organization may vary.
Solution:	A. Use [SRD] and [URD] as requirements baseline and make no additional assumptions of a specific hosting organization or operator: Specify a high-level architecture which does not rely on pre-existing systems. Leave it open if system elements will be developed, adapted from existing solutions, or integrated with the operational infrastructure of the future hosting organization. For the architectural description use virtualization abstraction, Cloud computing and web technologies. Even if the actual realization later may be different, such a generic architecture will be sufficiently flexible to be adapted. Rely on [SRD] where it contains assumptions on operator or hosting requirements.
Motivation:	Integration with a hosting organization is intended but the actual organization is unknown and its requirements are not available at this stage [SR-0429], [SR-0680], [SR-0580], [SR-0562]. Cross-ECV synergies [SR-0570] and [CCI-SRD] are uncertain. The system will be integrated with a hosting organization both technically and organizationally. The hosting organizations may have strict standards and constraints on technology choices, and may already have developed or acquired operational infrastructure (hardware, processing platform or framework, operational monitoring & control software, interfaces, etc.). These may be incompatible with design choices made herein. On the other hand making specific assumptions on the future operator may turn our wrong at the end. Operational use of the system is still years ahead and in the time to come potential hosting organizations will be further developing their operational infrastructure. They will gradually refurnish or even replace it with modern

Technical Memo **Trade-off Issue: Requirements Maturity Status**	
	technologies, therefore, choose a modern state of the art approach.
Unresolved Issues:	The design may not survive unmodified the integration with a hosting organization or operator.
Alternatives Considered:	B1, B2. TUW or VUA as hosting organization (specific requirements and constraints known to be incompatible with those of other potential hosting organizations). B3. EODC-Water as hosting organization (envisaged Austrian national data center for earth observation, possibly as an element of ESA's collaborative ground segment; not yet fully established). B4. ZAMG as hosting organization (Austrian organization for metrological and geophysical services; requirements are not available from this organization). B5, B6. ESA or EUMETSAT as hosting organization (requirements are not available from these organizations). C. National state/s or EU as sponsors (requirements are not available from these entities). Science communities and beneficiary organizations as sponsors (specific entities not yet identified). D. Include requirements from [CCI-SRD] (incomplete, out-of-scope to perform requirements analysis thereupon).
Discussion:	N/A

Table 11: (TM-004) Development Capability

Technical Memo **Trade-off Issue: Development Capability**	
Solution Summary:	Specify system functions or elements for development, maintenance, evolution and QA.
ID	TM-004
Factors:	AF-016: Developing and perfecting ECV Data Products will be an ongoing and continuous effort. AF-017: Development of scientific methods and processing algorithms for L2 Data and ECV Data Products is ongoing and not expected to complete in the near future. AF-018: Changes in methods and algorithms need to be implemented as Processor Modules and transitioned into operations quickly and with the desired quality. New or improved algorithms shall become operational within 6 months. AF-020: Future satellite missions need to be integrated and will put a steady stress on updating methods, algorithms and processing software, and producing new or revised ECV Data Products. AF-021: Capability for System Evolution required.
Solution:	A. Specify development, maintenance and evolution as system elements: Specifically, support product development, software development and transition into operations. Specify a QA regime for continuous improvement of standards and processes. Assume separate life-cycle phases for initial system development and for

Technical Memo **Trade-off Issue: Development Capability**	
	subsequent operational use and evolution.
Motivation:	Development of scientific methods and processing algorithms will be ongoing in order to meet requirements from data users, which we can expect to change and to become more stringent in the future. Data from new satellite missions will need to be integrated. This situation will remain similar throughout the lifetime of the System. Changing data, products, methods, algorithms and technologies will cause corresponding updates of software. Some software components will directly interface with data users and therefore be exposed to changing user requirements. Application and platform software will be affected by updates and replacements of IT infrastructure and technologies, which can be expected to take place in cycles of 3 to 5 years.
Unresolved Issues:	N/A
Alternatives Considered:	B. Operations without capabilities for development and evolution (like GMES services, which did not initially consider this).
Discussion:	N/A

Table 12: (TM-005) ECV Dependency

Technical Memo **Trade-off Issue: ECV Dependency**	
Solution Summary:	Avoid dependencies on other ECVs but support common ECV system elements if existing.
ID	TM-005
Factors:	AF-009: Independence between ECVs. AF-010: Cross-ECV synergies. AF-015: Recommendations and standards from different sources need consolidation. Standardization among ECVs uncertain.
Solution:	B. Stay independent from other ECV systems but allow cross-ECV synergies: Do not rely on other ECV systems' data or data products. Support common ECV system elements, but for the system specification do not assume they will be available. Specify in terms of abstract elements which can later be concretized. Allow the system to scale to produce additional data products (other ECVs or even different product types like SWI).
Motivation:	The cross-ECV collaboration but independence requirements potentially appear contradictory. The independence requirement was relaxed by the SEWG effort to seek cross-ECV synergies. Results from SEWG are not yet conclusive. Three potential common ECV system elements have been specified (see below).
Unresolved Issues:	If collaborative work on system engineering by the individual ECV projects should lead into joint deliverables (common standards, procedures or system

Technical Memo **Trade-off Issue: ECV Dependency**	
	elements), then two options are available: a) joint and dedicated projects parallel to the ECV projects in CCI Phase 2 (like CMUG project), or b) dedicated tasks with deliverables across Phase 2 ECV projects.
Alternatives Considered:	A. Stay independent from other ECV systems not allowing any cross-ECV synergies. This would reduce the overall efficiency of the CCI programme. C. Directly reuse elements of other ECV systems. This would create undesired dependencies.
Discussion:	N/A

Table 13: (TM-006) Standardization

Technical Memo **Trade-off Issue: Standardization**	
Solution Summary:	Consolidate and tailor standards at beginning of CCI Phase 2.
ID	TM-006
Factors:	AF-014: ECSS standards not applicable for extended system engineering approach including organizational requirements; they lack description of processes and require tailoring. AF-015: Recommendations and standards from different sources need consolidation. Standardization among ECVs uncertain.
Solution:	B. Setup standards at beginning of CCI Phase 2, performing consolidation and tailoring at that time.
Motivation:	ESA standard ECCS-E-ST-40C proposed by ESA TO specifically applies to software development, but not to the extended system scope including organizational requirements.
Unresolved Issues:	Standards documents [SR-0600] and [SR-0590] are not available.
Alternatives Considered:	A. Setup standards under CCI Phase 1 (out-of-scope). C. Setup standards at System startup (better to start in the development phase).
Discussion:	N/A

Table 14: (TM-007) Sustainment

Technical Memo **Trade-off Issue: Sustainment**	
Solution Summary:	Have system elements supporting sustainment; and develop a Sustainment Plan specifying the required resources for long term system sustainment under CCI Phase 2.
ID	TM-007
Factors:	AF-003: Long-term system sustainability, technically and financially.
Solution:	B. Have system elements supporting sustainment:

Technical Memo **Trade-off Issue: Sustainment**	
	Specify system functions for long-term planning and system evolution. Specify system functions for development, maintenance and evolution of data, products and software. Specify a system architecture which does not impose specific technologies. Defer implementation decisions to CCI Phase 2. C. Develop a Sustainment Plan under CCI Phase 2: Request a critical mass of human resources with know-how, and maintain it. Include resources so that the team can perform development to respond to changing requirements and technologies.
Motivation:	ECV Data Products shall cover all times from the beginning of the first satellite measurements and shall be extended into the future without interruption. They shall further be preserved for the next human generations.
Unresolved Issues:	Financial sustainability requires a long-term sponsor and funding plan. Is free access in SR-0640 intended to mean no cost? If so it is unknown if the future sponsor and hosting organization will support this.
Alternatives Considered:	A. Find long-term Sponsor (not an architectural design task): A1. EU A2. National States A3. Science Communities A4. Beneficiary Organizations
Discussion:	N/A

4.2 CCI Common Elements and Collaboration

Table 15: (TM-008) CCI Web Portal

Technical Memo **Trade-off Issue: CCI Web Portal**	
Solution Summary:	Specify Soil Moisture system elements, which can be integrated into a common CCI portal.
ID	TM-008
Factors:	AF-011: Common portal for ECV Data Products dissemination.
Solution:	B. Linkage integration with a light-weight CCI Web Portal: Specify system elements for a Soil Moisture Web Portal and a Soil Moisture ECV Data Products Dissemination System. Assume that a common CCI Web Portal will provide a dedicated area for Soil Moisture and will be able to link to the Soil Moisture Web Portal and Dissemination System.
Motivation:	SEWG may provide recommendations for a common portal for ECV Data Products dissemination. Discussions at SEWG indicate a preference for a light-weight CCI Web Portal to act as a common entry point for data users, which may simply be a web page.

Technical Memo **Trade-off Issue: CCI Web Portal**	
Unresolved Issues:	The CCI Web Portal is currently undefined.
Alternatives Considered:	A. Stand-alone Soil Moisture Web Portal. C. Heavy-weight CCI Web Portal that comprehensively implements user support, public relations, collaboration features and services, and ECV Data Products dissemination for all ECVs.
Discussion:	N/A

Table 16: (TM-009) Processing Environment

Technical Memo **Trade-off Issue: Processing Environment**	
Solution Summary:	Specify a heterogeneous Cloud processing architecture.
ID	TM-009
Factors:	AF-012: SEWG may provide recommendations for common tools and for a shared processing environment on a low level. The processing environment may also be provided by the future hosting organization and requirements may clash. AF-022: ESA CCI and scientists prefer use of open source software.
Solution:	E. Heterogeneous Cloud processing architecture: Specify in terms of Cloud processing architecture, use virtualization abstraction and leave concrete realization of hardware and software infrastructure and environments open. Allow for a heterogeneous Cloud to use the hardware and software infrastructure of the hosting organization and to extend it with commercial (public) Cloud services for collaboration and high resource demand situations. Plan cost evaluation and selection of service providers as a recurring task with a planning cycle of 1-3 years. Initially required facility and hardware resources for Soil Moisture are specified in [SRD] section 6.5.
Motivation:	Current commercial Cloud processing services are reported to be three times more expensive than dedicated hardware in case of a steady processing workload [Colloc3]. On the other hand, in peak demand situations commercial Clouds are more cost effective than largely idle local IT resources. Further advantages of public Clouds are their scalability/elasticity, and the possibility to share computing environments. Clouds also have advantages for scientific collaboration as they provide a collaborative environment for sharing of data, algorithms and software. A dedicated European Science Cloud may become available and may lower costs for scientific applications, but this is undetermined at the time of writing. We assume that future information systems will be Cloud-like. A demonstrator project (EP4SM) is currently ongoing to implement and evaluate a Cloud platform as a proof-of-concept for Soil Moisture processing. The SM ECVPS will be tested on EP4SM in comparison to the deployment in the existing IT infrastructure at TUW. The IT infrastructure of the eventual hosting organization or operator is not

Technical Memo **Trade-off Issue: Processing Environment**	
	yet known.
Unresolved Issues:	Open source software is required in [SR-0650] and was stated for application software, which includes common tools that may be shared between ECVs. This will need to be reviewed with respect to the DSWG standards and SEWG proposals when they are available.
Alternatives Considered:	A. Dedicated and isolated hardware and software infrastructure that is not prepared for extension into the Cloud. B. Shared infrastructure and common tools among ECVs (unknown at the time of writing). C. GRID processing infrastructure (lacks virtualization layer, aged technology). D. Use a commercial (public) Cloud processing infrastructure (too expensive).
Discussion:	N/A

Table 17: (TM-010) L1 Data Archive

Technical Memo **Trade-off Issue: L1 Data Archive**	
Solution Summary:	N/A
ID	TM-010
Factors:	AF-013: SEWG may provide recommendations for an L1 Data Archive.
Solution:	N/A to Soil Moisture.
Motivation:	The Soil Moisture ECV system will implement the processing chain from L2 to ECV but not from L1 to L2. L1 to L2 processing is performed by the L2 Data Providers TUW and VUA.
Unresolved Issues:	ESA CCI may want to contact L1 Data Providers on access to L1 data for use in the L1 Data Archive (ESA, EUMETSAT, NASA), and Soil Moisture L2 Data Providers (EUMETSAT, NASA, TUW, VUA) on shared use of the L1 Data Archive.
Alternatives Considered:	None.
Discussion:	N/A

Table 18: (TM-011) Version Control

Technical Memo **Trade-off Issue: Version Control**	
Solution Summary:	Use existing open-source DCVS solutions and integrate them with the CCI or Soil Moisture Web Portal.
ID	TM-011
Factors:	AF-024: Capabilities for open collaboration.

Technical Memo **Trade-off Issue: Version Control**	
	AF-025: Configuration control for software and configuration items. AF-026: Version control for data and data products.
Solution:	B. Use an open-source DCVS (e.g. Git) for configuration control of software and configuration items. C. Resort to file naming conventions for version control of data. Specify a CCI and/or Soil Moisture Web Portal and integrate collaboration tools into the portal. Put all system software, configuration items (e.g. processing parameters) and data products under configuration control and establish a configuration control board (CCB) and a product control board (PCV). Define standards, policies and procedures for versioning and configuration control of data, software and configuration items. Detailed configuration control policies and procedures will depend on the choice of a particular VCS. Aim for a common CCI approach to open collaboration to reach a wider acceptance within the science community.
Motivation:	Generally prefer open-source solutions for collaboration tools. Mature open-source Version Control Systems (VCSs) are available. Other ECV projects recommend Git or Subversion. Preference is given to the newest generation of Distributed Version Control Systems (DVCS) over centralized ones like CVS, CVSNT or Subversion. A suitable example for collaboration in a distributed community is Git. Assume that L2 Data, ancillary data, internal data products, and ECV data products will be versioned at the file level, and will adhere to DSWG and [PSD]. DSWG agreed on file naming conventions.
Unresolved Issues:	Digital Object Identifiers (DOI) may be used to identify the Soil Moisture ECV Data Product/s (under discussion).
Alternatives Considered:	Configuration control for software and configuration items: A. Centralized VCS (CVS, CVSNT, Subversion and the like) (not appropriate for scientific community which is non-centralistic). Version control for data: D. VCS for data (L2 Data, intermediate data, or ECV Data Products) (data volumes too high). E. Distributed and synchronized database approach for L2 Data and ECV Data Products (too ambitious).
Discussion:	N/A

4.3 Data Processing Software

Table 19: (TM-012) ECV Processor Architecture

Technical Memo **Trade-off Issue: ECV Processor Architecture**	
Solution Summary:	Specify a processing framework with a standard interface to individual processing modules.

Technical Memo **Trade-off Issue: ECV Processor Architecture**	
ID	TM-012
Factors:	AF-005: Integration with a hosting organization. AF-019: Flexible processing chains (workflows) and standardization of workflow elements. AF-022: Preference for open-source software. AF-025: Configuration control for software and configuration items. AF-026: Version control for data and data products. AF-027: Tools for process monitoring & control and anomaly recovery. AF-028: Support for several processing modes: incremental processing, full reprocessing, continuous interval processing (NRT), on-demand processing. AF-029: Quality, reliability, maintainability and evolution of processing software.
Solution:	B. Modular architecture with processing framework and specialized processing modules: Modularize processing software and assign processing functions to dedicated modules. Specify a standard interface between the framework and individual modules. Specify the following components: process chaining (workflows), individual processing steps (specialized modules implementing algorithms), monitoring & control applications. Specify dedicated and separate environments for development, testing and operations. Wait with the definitive selection of the processing framework until a decision has been made about a possibly shared CCI processing environment.
Motivation:	The processing software will be more stable and better maintainable if it can be decomposed into a framework and modules, where the framework does not depend on the specifically implemented algorithms (abstraction), and the modules implementing individual processing algorithms do not need to be aware of the general processing functions performed by the framework. A stable software architecture will also facilitate versioning and configuration control. Further, modularization enables separate performance tuning and optimization of individual modules.
Unresolved Issues:	N/A
Alternatives Considered:	A. Implementation of a single large ECV processor without separation of individual functions and without standardized internal interfaces (not achievable with the required quality criteria).
Discussion:	N/A

Table 20: (TM-013) ECV Processor Programming Language

Technical Memo **Trade-off Issue: ECV Processor Programming Language**	
Solution Summary:	Select a general processing framework which supports a range of

Technical Memo Trade-off Issue: ECV Processor Programming Language	
	programming languages.
ID	TM-013
Factors:	AF-023: Diverging requirements on programming language for the implementation of processor modules.
Solution:	D. Choose a processing framework with multi-language support: Modularize processing software and separate functions for chaining (workflows) from individual processing steps (dedicated modules). Use a suitable processing framework that supports chaining and calling library functions, executables, and services programmed in C, C++, Java, Python and IDL. Use a processing framework that supports parallelization, and for the Cloud environment also scaling and elastic computing. Properly implement processing modules using a suitable programming language for operations after rapid algorithm prototyping using a prototyping language.
Motivation:	The current software environment at TUW is based on IDL and requires applications being implemented in IDL. The may be different for the eventual system operator and hosting organization. Modern SOA architectures allow heterogeneous systems to be built using a range of programming languages. Libraries written in IDL can be linked into a processing framework. IDL can be called from C/C++, and C/C++ can be called from IDL. Similar for Java and Python. Calling Python from IDL requires a commercial license, calling IDL from Python is free but requires Linux. In order to realize SOA or Cloud architectures, web service interfaces can be implemented with IDL using ION (IDL On the Net). ION requires a commercial license. Whatever utilities, means or methods are used for rapid algorithm development, subsequently, a proper implementation process will be needed to establish a code base that has operational quality, is maintainable and can further evolve.
Unresolved Issues:	Use of IDL in an open source environment is restricted. There may be license conflicts if open source software, which is distributed under the GPL license, shall be integrated.
Alternatives Considered:	A. Implementation of ECV processor in IDL (suggested by TUW). B. Implementation of ECV processor in C/C++ (requested by ESA in [NEG1]). C. Implementation of ECV processor in Python (under evaluation by scientific community).
Discussion:	It is not required that the processing framework supports MATLAB. Python is experiencing a hype in the scientific community and has already entered operational systems. UK met office is porting legacy IDL code to Python [Colloc3]. IDL has low runtime performance compared to compiled languages like C/C++, but is faster than byte-compiled languages like Java. IDL is closed commercial software and disagrees with the open source requirement [SR-0650]. Suitable IDL licenses for Cloud environments do not yet exist.

Table 21: (TM-014) ECV Processor Performance

Technical Memo **Trade-off Issue: ECV Processor Performance**	
Solution Summary:	Develop specialized Processor Modules that implement the individual steps of the processing chain, separately subject them to performance tuning and optimization, and drive them with a Processing Framework. Make use of parallelization, scaling, or elastic compute. Start a technology research project and get support by professional software developer.
ID	TM-014
Factors:	AF-030: Revised performance requirement for Processor Modules: 1 week for 1 year extension, and 3 weeks for full reprocessing for the current 32 year data set.
Solution:	B. Setup project to overcome performance barriers: Assume a spatial resolution of 0.25 degree for ECV Data Products at the beginning. Design processing software so that it can evolve, i.e. use a modular architecture. Perform profiling and performance testing on prototype software to analyze performance bottlenecks. Check processing efficiency and I/O throughput. Select and adapt an existing processing framework solution that can provide the required performance. Compare the performance of the selected solution to alternative solutions. Perform performance tuning and optimization for individual processor modules.
Motivation:	Current operational and future expected performance requirements will be much higher than stated in current requirements [SR-0370] and [SR-0376]. To overcome the performance issue a systematic approach with dedicated technology research and development will be needed.
Unresolved Issues:	N/A
Alternatives Considered:	A. Stay with the relaxed requirements [SR-0370] and [SR-0376].
Discussion:	The performance of the current prototype processor modules is not sufficient for full reprocessing of 30+ year time series data. Requirements [SR-0370] and [SR-0376] state relaxed scenarios that are not applicable to full reprocessing (i.e. with regridding of all L2 data, which will be necessary after L2 data reprocessing, e.g. after algorithm changes). A revised estimate for full reprocessing of 30 year data indicates a processing time of 30 weeks rather than the 12 weeks stated in [SR-0370]. Data volumes will increase considerably with the inclusion of data from new satellite missions (cf. [SRD] section 4.3). Higher spatial resolution data products may be required in the future [URD]. Both new satellite missions and higher resolution will put an increasingly higher demand on the processor performance.

5 System Architecture

This section specifies the System's architecture in terms of several complementary Architectural Views. An Architectural View is a reduced design description that puts the focus on a specific aspect like the System's static or logical structure, its control flows, its components, etc. The complete architectural design is then obtained by considering all complementary Architectural Views together.

A number of established frameworks are available to specify consistent and complete sets of Architectural Views for certain types of systems. Relevant examples are the "4+1 View Model of Software Architecture" for software systems [Kruchten95], the ISO "Reference Model for Open Distributed Processing" for distributed hardware-software systems [RM-ODP98], an evolution of the 4+1 View Model by the Software Architecture Document (SAD) of the Unified Process (UP) Design Model [Larman02], and the "Framework for Modeling Space Systems Architectures" defining a nominal set of views derived from earlier frameworks [Shames06].

Generally, Architectural Views are tools for system design and communication about design aspects and are expressed in text and diagrams. Most frameworks provide a set of optional Architectural Views for tailoring to the needs of the specific system to be designed.

For the purpose of specifying the Soil Moisture ECVPS we tailored the nominal set of views of [Shames06]. This set is comprehensive and allows for the consideration of organizational aspects in addition to hardware and software aspects of a system. However, none of the frameworks we are aware of specifically considers the architectural aspects of virtualization technology or Cloud Computing (cf. the Infrastructure View in section 5.7). In order to incorporate all relevant architectural aspects we tailored the set of Architectural Views as follows:

- **Functional View:** Definition of the System in terms of individual System Functions which are obtained from a functional decomposition process. The Functional View is the starting point for the other Architectural Views which are responding to it by describing System Elements or System behavior implementing the System Functions.
- **Organization View:** Description of the human organization, establishing the System Team and its organizational structure.
- **Software View:** Description of the Software Components among the System Elements, establishing the Software Configuration Hierarchy.
- **Process View:** Description of System behavior, specifying the System's main internal processes executing during the System's operational use (Operational Scenarios), and their interaction with the (external) System Environment.

- **Use-Case View:** Alternative description of System behavior, specifying the System's response to actions initiated by Actors from the (external) System Environment.
- **Information View:** Description of data and information items among the System Elements, establishing their structure and relationships.
- **Infrastructure View:** Description of the basic services and facilities necessary for the System to be deployed and operated.

5.1 Functional View

In section 3.2 we provided a description of the System's operational context, i.e. of those elements that have a relationship with the System but remain outside the System. In this section we define the System Functions. These are the tasks or activities the System needs to perform in order to achieve its Mission (cf. section 3.1, Table 2).

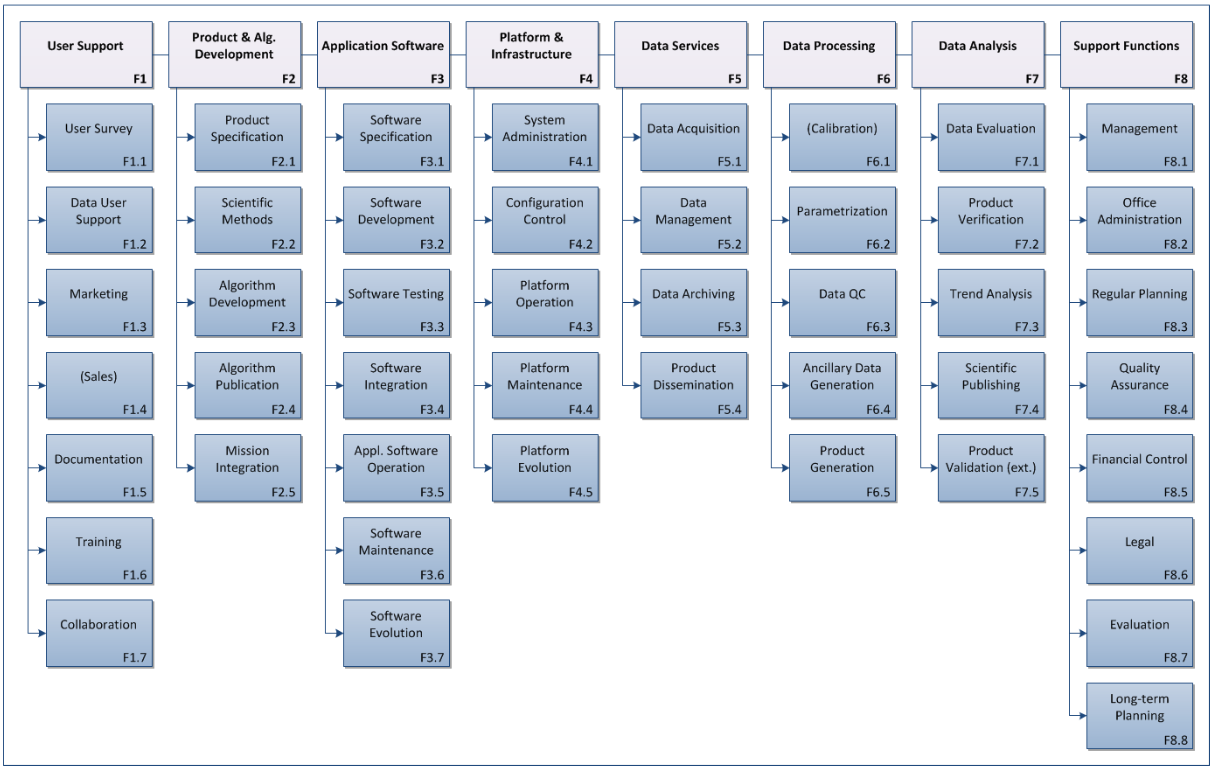


Figure 1: System Functions Hierarchy

Figure 1 shows the System Function Hierarchy, i.e. a decomposition of Functional Areas (light blue boxes in the top row) to individual System Functions (all other boxes). The System Functions are derived from the System Objectives listed in section 3.1, Table 3. A Functional Area is implicitly defined by the System Functions it aggregates. The Functional Areas are characterized in more detail in the subsequent paragraphs. Other architectural views refer to the System Functions Hierarchy and define System Elements composed of entities, structures, processes, interactions, etc. as needed in order to implement the System Functions identified in this section.

System Functions are grouped by the collectively shared entities they are operating on. The Functional Areas (or main System functions) specified in this way are:

- **(F1) User Support:** All types of interactions with Data Users including support for software users and developers and collaboration with external entities. This Functional Area provides an external system interface. It implies the use of certain Application Software (web portal, version control systems, etc.) and may overlap with other Functional Areas.
- **(F2) Product & Algorithm Development:** Science related activities, from the definition and specification of ECV Data Products to the development of Scientific Methods and Processing Algorithms.
- **(F3) Application Software:** All activities related to the development, operations and maintenance of System's Application Software.
- **(F4) Platform & Infrastructure:** Provisioning of the hardware, software and facilities, which are required to host the System's Application Software and to store (archive) its data. This includes the administration, operations, maintenance and development activities required to sustain the operational platform in the long-run.
- **(F5) Data Services:** Acquisition, management and dissemination of the System's data and data products. Data Services will make use of dedicated Application Software (i.e. data management software) to perform these System Functions.
- **(F6) Data Processing:** Production of the ECV Data Products. This is a core Functional Area of the System. It is implemented by specific Application Software (i.e. the ECV Processors) to perform its functions.
- **(F7) Data Analysis:** Scientific evaluation and analysis of both incoming and outgoing data and data products. This heavily draws upon the capabilities of scientists who use specific Application Software (i.e. interactive data viewing and analysis tools) to perform these functions.
- **(F8) Support Functions:** Functions in support of all other Functional Areas, typically coordinating or controlling them at the system level. Examples are management and administration functions.

5.2 Organization View

This section describes the Human Organization of the System in terms of an Organization View. The View establishes the System Team and its organizational structure. The System Organization described in this section is essential, specifically for the System Functions which are implemented by team members (e.g. development and data analysis), but generally for all System Functions specified in section 5.1, because all of them require a certain amount of human effort and control to be effectively performed.

The Soil Moisture ECVPS' organizational structure is shown in the Organizational Chart of Figure 2. The chart depicts the organizational units and lists the principal roles of the team members.

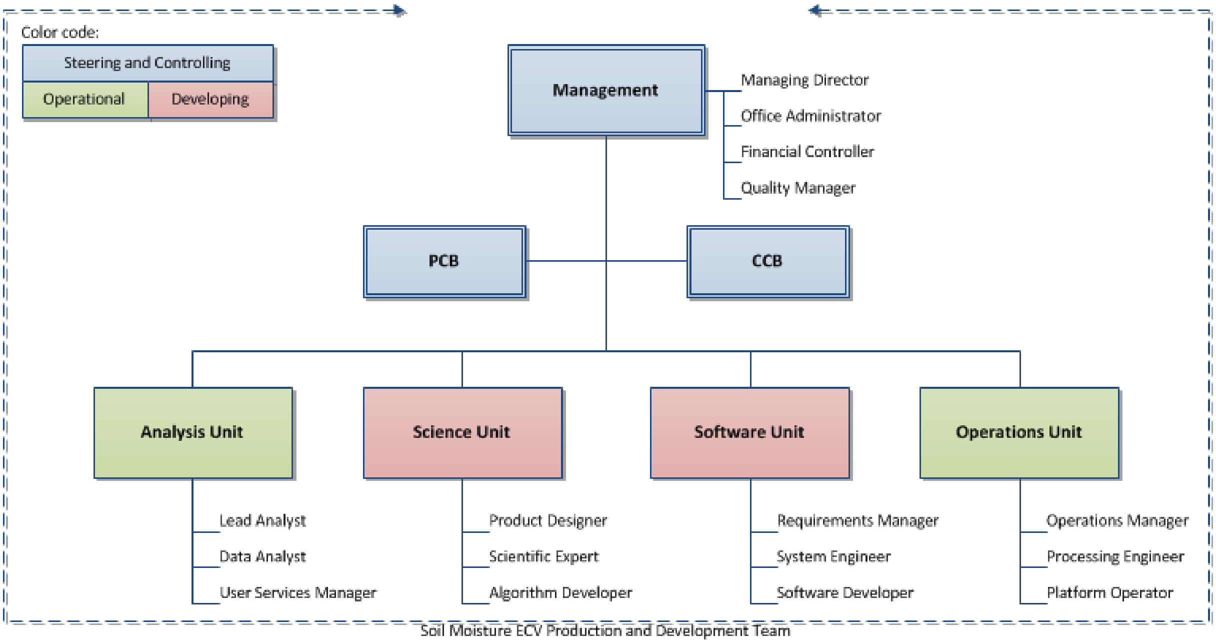


Figure 2: Organizational Chart of the Soil Moisture ECVPS

The System's organization is based on three principles: (a) separation of responsibilities for steering & control, operations and development, (b) dedicated teams for the geo-science and IT domains, and (c) inclusion of control boards to steer System processes.

The blue colored boxes in Figure 2 represent the entities with System-wide steering and controlling responsibilities: Management, assisted by the Product Control Board (PCB) and Configuration Control Board (CCB). While the PCB is in control of scientific evolution, the CCB controls the technical evolution of the system. Either board includes representatives from pertinent units and external stakeholders as appropriate.

The green colored boxes represent the teams with operational responsibilities: Analysis Unit and Operations Unit, while the red colored boxes represent the units with development responsibilities: Science Unit and Software Unit.

The Analysis and Science Units on the left of Figure 2 have a focus on the geo-science domain, while the Software and Operations Units on the right have their focus on applied information systems (software and IT).

The principal roles shown in Figure 2 are not necessarily meant to correspond to full time staff positions. The team size is indicated as 12 permanent staff in [SRD], section 6.5.1. In order to assign all team roles individual team members may either fill several roles or jointly share additional roles to their main roles.

5.3 Software View

This section specifies the Software Components among the System Elements in an architectural Software View. The Software View identifies software that is essential for the implementation of the System Functions defined in section 5.1.

The System's Software Components are grouped into three architectural layers:

- **Application Software:** The subset of software procured or developed in order to support or provide the System's operational functions. The users of Application Software are Data Users, Data Analysts, Scientific Experts and Product Designers. Examples of Application Software are web applications, data transfer software, the ECV Processor, and data viewing & analysis tools.
- **Platform Software:** Software required to provide the Software Environments (ENVs) for development and operation of Application Software. Among the Platform Software there are development tools, software frameworks, libraries, runtime environments, etc. The direct users of the Platform are those who develop, test, deploy and operate the Application Software. Algorithm Developers, Software Developers and Processing Engineers are in this group.
- **Infrastructure Software:** The bottom layer of software, constituted by virtualization software, Host Operating Systems (OSes), and any other software, which is required in order to operate software of the other two layers on the hardware infrastructure.

The individual diagrams and descriptions in the subsequent subsections specify in turn: (1st) the Application Software, (2nd) the Platform Software at a general level, and (3rd) the Platform composition (set-up) in terms of Software Environments, computational nodes and associated Platform Software.

The Infrastructure Software is specified together with the Infrastructure Hardware by the Infrastructure View in section 5.7.

5.3.1 Application Software

The content of the Application Software Layer is presented in Figure 3 in the form of a Software Configuration Hierarchy, partitioning the System's Application Software into: (A1) User Tools, (A2) Data Services, (A3) Automatic Processing, (A4) Interactive Analysis, and (A5) Utilities.

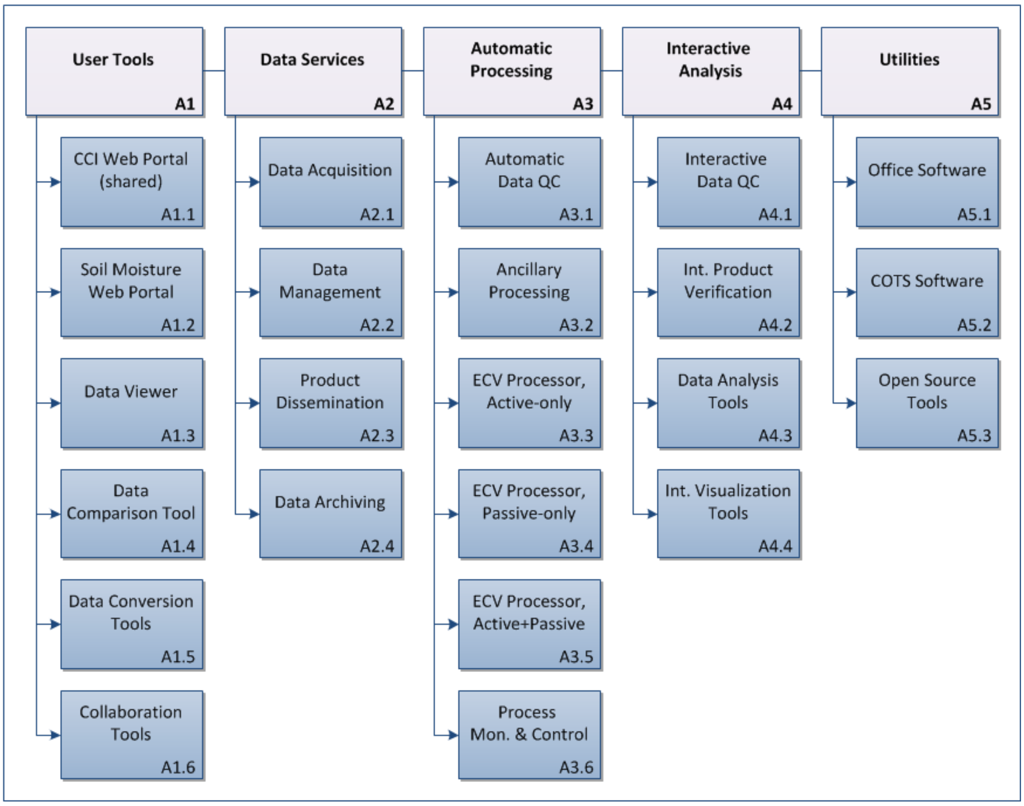


Figure 3: Software Configuration Hierarchy for Application Software

From left to right, the individual columns of Figure 3 are as follows:

- **(A1) User Tools:** Tools for use by the System's end users, which are primarily the Data Users but also internal users like Data Analysts, Algorithm Developers and Software Developers. The User Tools will be made available to the Science Community or the general public and will therefore be implemented as web applications.
- **(A2) Data Services:** Software Components dealing with System's data, from acquisition, handling of internal data storages and data flows, to dissemination to the Data Users and the long-term preservation in data archives.
- **(A3) Automatic Processing:** Software Components required to perform data processing in an automatic way including process monitoring and control. The Data QC software and the ECV Processors belong to this group, where the Data QC software will automatically run upon the acquisition of new input data.
- **(A4) Interactive Analysis:** GUI and command line tools to be used by Data Analysts for interactive review and analysis of L2 Data, Ancillary Data and the eventual ECV Data Products. Interactive Analysis tools will primarily be provided as desktop applications, though some could be implemented as web applications.

- **(A5) Utilities:** Software supporting the daily work, i.e. standard Office Software and other commercial or open-source software utilities. The specific set of Utilities can be adapted to users' personal preferences and may also change over time.

5.3.2 Platform Software

The Software Components used to implement the System's Platform of Software Environments are shown in Figure 4. Figure 4 shows low-level (i.e. close to infrastructure) software components at the bottom and high-level (i.e. close to application software) components at the top.

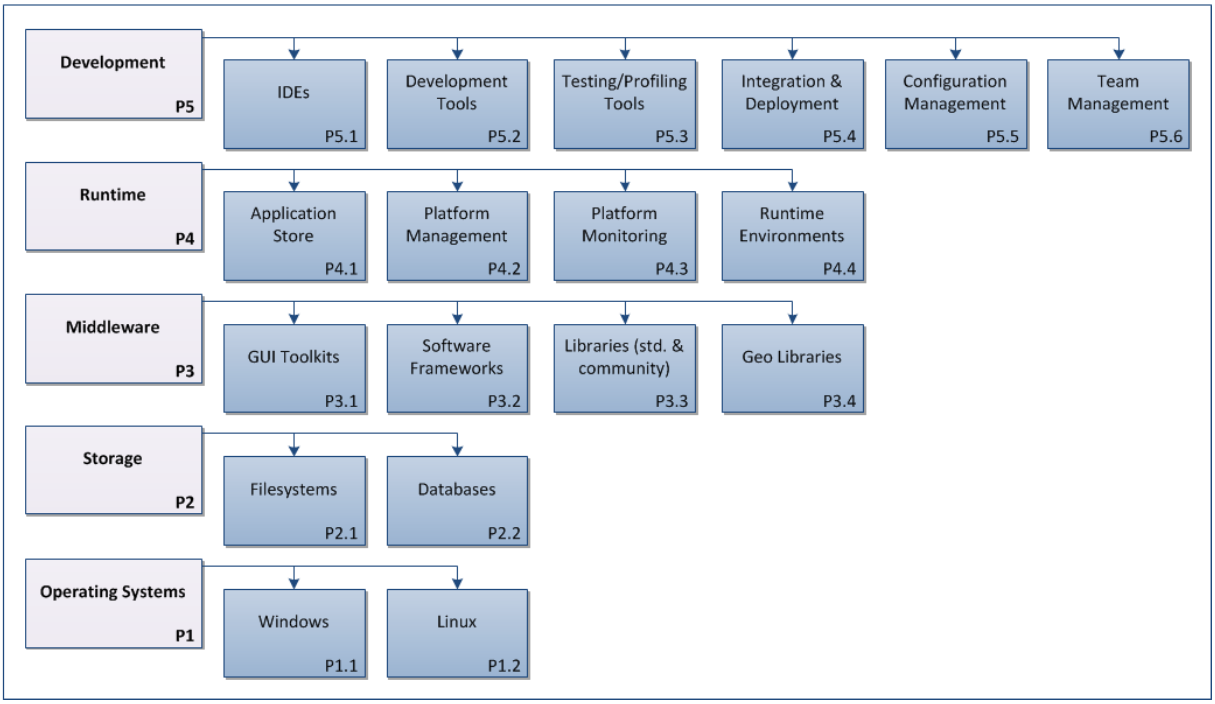


Figure 4: Software Configuration Hierarchy for Platform Services

From bottom to top, the individual levels of Figure 4 are:

- **(P1) Operating Systems:** Operating Systems for deployment on Virtual Machines (Guest OS for VMs).
- **(P2) Storage:** File system and database software used to implement the System's operational and archive data stores. These data stores provide access to the System's data for software in the Platform and Application layers. The data store consists of shared file systems and Relational Database Management Systems (RDBMS).
- **(P3) Middleware:** Software stacks used to implement the System's custom-built software. Matching to the applications' chosen programming languages and software architectures, it provides the appropriate GUI toolkits and software frameworks (e.g. for parallel computing, process chaining, elastic computing if the Platform is

deployed on a Cloud infrastructure, security, persistence). Middleware further includes standard and community libraries for software development and the Geo Libraries further specified in Figure 5.

- **(P4) Runtime:** Software for storing application binaries and for accounting, Platform management and monitoring services, i.e. all services that are needed to manage the Platform itself. This layer also includes the runtime environments (Process VMs) and web containers as required by the individual programming languages for the execution of the pertinent applications. E.g. in the case of Java these would be the Java Runtime Environment (JRE) and Servlet containers.
- **(P5) Development:** This top layer completes the Platform's software stack. It provides the tools required by software developers, testers and integrators in order to perform their work on the Platform. On a platform with layers P1-P4 Application Software can be run, on a platform with layers P1-P5, Application Software can be developed.

Figure 5 shows the decomposition of element Geo Libraries (P3.4) from Figure 4. The Geo Libraries are a tailored set of libraries selected or custom-built specifically for the Soil Moisture ECV domain. Geo Libraries are the building blocks for the Soil Moisture data processing software (and the ECV Processors). They are reusable and available to Algorithm Developers and Software Developers in order to avoid a redundant implementation of common processing functions, to facilitate the development of modular applications, which are maintainable and implement a commonly agreed standard, and thus to increase the overall efficiency of software development and evolution processes.

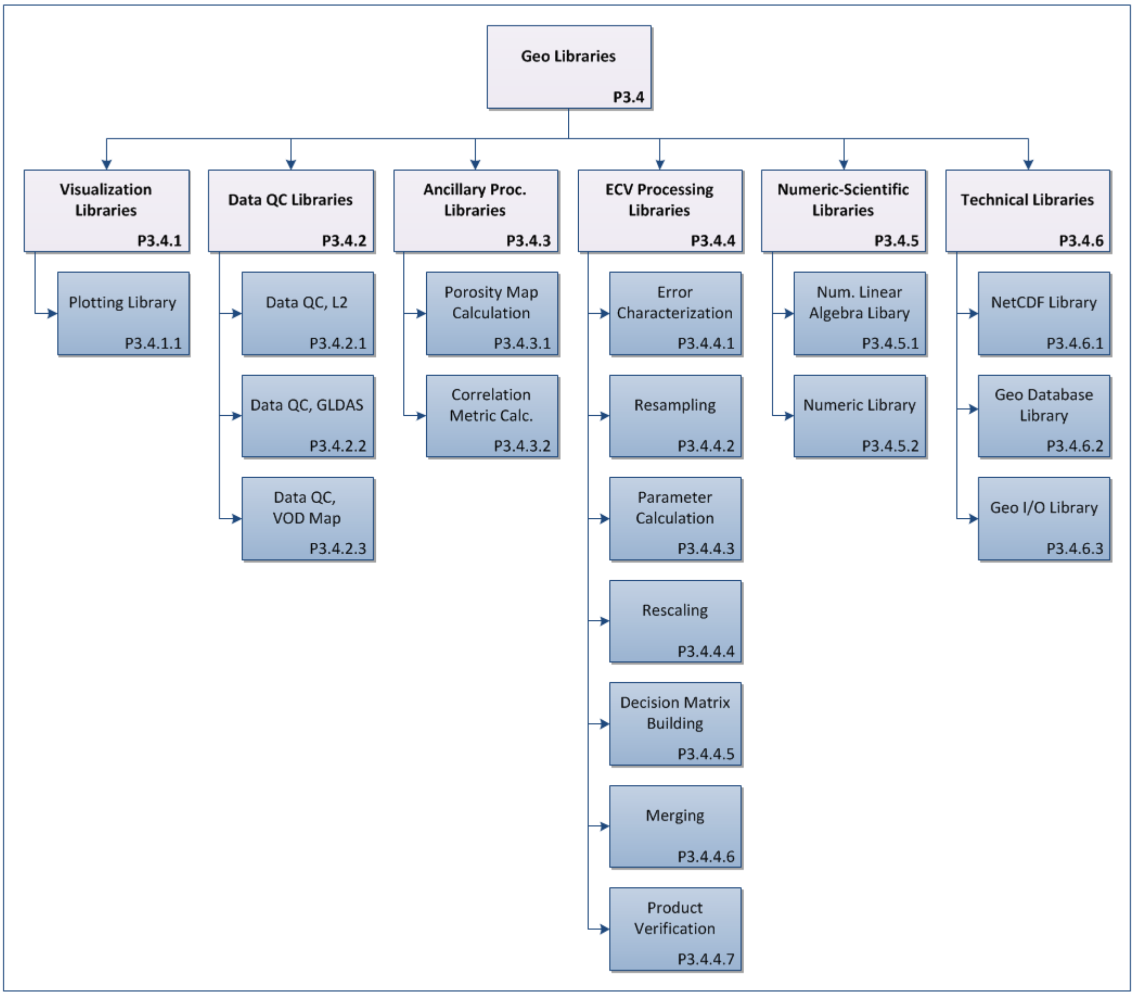


Figure 5: Software Configuration Hierarchy for Geo Libraries

5.3.3 Platform Composition

After specifying the Platform's software and services in the previous section this section describes the composition of the Platform in terms of Software Environments, architectural Platform layers and virtual nodes. Figure 6 provides a schematic view of the Platform composition.

In a physical architecture the nodes (grey boxes in Figure 6) would correspond to hardware (servers and storage devices), while ENVs may correspond to LANs. However, due to the use of virtualization all environments are implemented with the components of the Platform Software shown in Figure 4, on top of the Infrastructure Layer specified in section 5.7.

Individual Software Environments are defined by shell-like structures: At the center is an active software process, which may run within a process-specific VM (e.g. for IDL or Java processes) constituting the first shell. The next shell is defined by a System VM (i.e. the VM instances shown in Figure 6) running the process-specific VM. The System VM is instantiated from a VM template (shown as ENV cells and partitions in Figure 6). Finally, the outer shell is defined by the complete Platform. Each shell has its specific methods for controlling and securing the software processes and services in its interior and also provides services to access these processes. This shell-like structure is depicted in Figure 6 from the level of System VMs (grey boxes) via ENVs to the Platform.

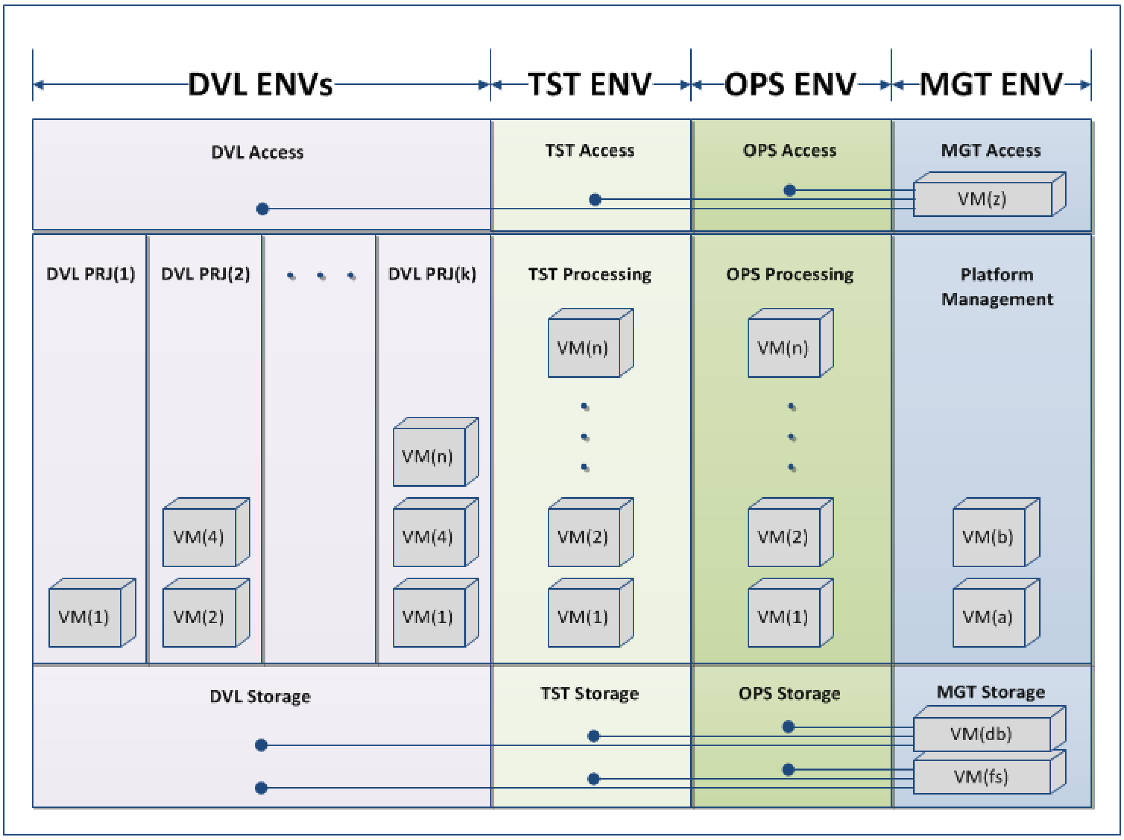


Figure 6: Platform Composition by Environments (ENVs) and Virtual Machines (VMs)

The following paragraphs describe the Platform's composition shown in Figure 6: A) The Platform's partitioning structure from left to right, B) its layering structure from top to bottom, and C) its set-up from virtual nodes and allocation of software and services to these nodes.

A) Platform Partitioning Structure: The Platform consists of four largely isolated Software Environments (ENVs), each one serving a dedicated purpose:

- **Development Environments (DVL ENVs):** A set of virtual environments, one per development project, PRJ(i), where each DVL ENV is customized to the project requirements. DVL ENVs are equipped with IDEs, compilers, development libraries, etc., and the runtime environments to execute & debug the developed software. A DVL ENV may be shared between several developers who work on the same project.
- **Test Environment (TST ENV):** Resembles the OPS ENV (see below), and in addition provides testing, profiling and integration tools. Fully isolated from OPS ENV to

enable testing and demonstration of software without repercussions on the operational services.

- **Operations Environment (OPS ENV):** Provides the runtime environments (processing VMs, web containers and other runtime services like databases, file systems, etc.) as required for the execution of the System's operational Application Software.
- **Platform Management Environment (MGT ENV):** Provides all services as required for administration and management of the Platform as a whole.

B) Platform Layering Structure: The Platform is stratified into three layers (from top to bottom):

- **Access Layer:** A layer for security and access control with firewalls, reverse proxies, etc., isolating public services exposed to the Internet from internal services.
- **Process Layer:** A layer for hosting the Application Software processes (in the DVL, TST and OPS partitions) and Platform management services (in the MGT partition). This layer provides the Platform's computational and processing power.
- **Storage Layer:** A layer for providing shared storage services to each partition in the form of databases and file stores. There are operational and archive instances of data stores: e.g. operational databases named dbd, dbt, dbo, dbm, and archive databases named dbad, dbat, dbao, dbam; as well as operational file systems at mount points /dvl, /tst, /ops, /mgt, and archive file systems at /arc/dvl, /arc/tst, /arc/ops, /arc/mgt.

C) Platform Set-up from Virtual Nodes and Allocation: The virtual nodes are the Platform's building blocks; they are Virtual Machines (VMs) simulating standard x86 hardware architectures. VMs are subject to the following conditions:

- Each VM is equipped with a Guest OS, suitable libraries and runs services providing an appropriate runtime environment for executing the desired software.
- Application and Platform Software listed in the Software Configuration Hierarchies of Figure 3 and Figure 4, respectively, are allocated to VM templates (the expressions like "VM(2)" in Figure 6 are identifiers for such templates).
- The allocation of OS and software to VM templates is performed dynamically during the System's operational use with the help of the Platform Management Software and the Application Store (elements P4.2 and P4.1 in Figure 4, running as services in VMs in the MGT ENV, i.e. VM(a) in Figure 6).
- A VM instance can be created from a VM template by the Platform Operator or other users with the permission to do so in accordance with the partitioning and layering

constraints as described above. After instantiation, the VM is allocated to a particular cell of the partitioned and layered grid in Figure 9.

- VMs (i.e. instances) can be started and stopped by the Platform Operator or other users having the relevant permission. (The grey boxes in Figure 6 symbolize individual VM instances, e.g. VM(2) is instantiated three times in this example, once in each of DVL PRJ(2), TST ENV and OPS ENV.)
- Only the Platform Operator is allowed to move VM instances between cells of the partitioned and layered grid.
- A VM instance generally provides its services within the allocated partition (ENV) but may also provide services to other partitions (ENVs) or to the Internet through the Access Layer and depending on individually assigned access rights. Some examples are shown by interface nodes in Figure 6, e.g. the interface nodes extending from VM(db) to the partitions on the left.
- The default access policy for the shared storage services VM(db) and VM(fs) is: An ENV is granted read-access to databases and file systems of all ENVs to its right in Figure 6 but has write-access only to its own database and file system. For security reasons an ENV may overrule the default policy and restrict read-access for others for certain of its data.
- Virtualization software (hypervisors), Host OSes and basic services for managing and instantiating VMs are provided by an underlying Infrastructure Layer (not shown in Figure 6, but further described in the Infrastructure View in section 5.7).

Use of VMs as building blocks increases both the Platform's robustness and its flexibility. Application Software can be isolated from each other, and VMs can be dedicated to specific projects and purposes, avoiding conflicts between application, library and OS versions. Due to the virtualization, the allocation of software to nodes can be dynamically adapted to changed requirements and can evolve together with the System.

The Platform's composition in terms of the four described partitions enables the isolation of activities for different purposes: Development activities are strictly separated from testing & demonstration activities, and both are isolated from System operations. A fourth isolated activity is platform management, while common services are consistently provided to all partitions and activities. This makes the software system very robust and reliable.

The model also greatly facilitates management, administration and confidence building among the Platform users: It provides an area where software behavior can be demonstrated and evaluated in a near-operational environment (TST ENV), and it allows setting up a structured software development process, as specified in the Process View in section 5.4.

The Platform model furthermore clarifies some of the organizational roles introduced in the Organization View (Figure 2 in section 5.2): The primary users of the DVL ENVs are the Software Developers, those of the TST and OPS ENVs the Processing Engineers, and those of the MGT ENV the Platform Operator and other users in a system management role.

5.4 Process View

The Process View is one of the two architectural views used in this document, which describe the System's dynamic behavior. It specifies the main processes to be implemented in the System. Individual processes may interact with the external System Environment specified by the Operational Scenarios. In contrast to the second dynamic view, the Use-Case View in section 5.5 describing the System's response to external actions, the Process View references elements from the static architectural views in sections 5.1 and 5.2.

The inputs to the Process View are the main Operational Scenarios identified in section 3.1. The Process View is constituted by the set of UML Activity diagrams [UML] and their corresponding descriptions.

5.4.1 Data Acquisition Process

Figure 7 shows the data acquisition process, i.e. the flow of activities triggered by a Data Provider publishing new L2 Data or Ancillary Data, which are required for the production of the Soil Moisture ECV Data Products. A guiding principle for this activity flow is that data acquisitions shall be made early and independently from the actual ECV Processing, as soon as new data have been published, for the following reasons:

- To perform an automatic Data QC check to confirm completeness and to reveal any data issues;
- To verify data with respect to its suitability for ECV Production in an interactive review performed by Data Analysts in addition to the automatic Data QC;
- To provide immediate feedback to Data Providers in case data issues were observed, giving them the chance to fix issues before the next scheduled ECV Processing;
- To avoid input data issues at the time when ECV Processing is scheduled to start;
- To have the data available in the System for other evaluation, analysis or development activities, that may be performed by Analysts or Scientists.

In Figure 7 newly available data are represented by the "DataProduct" UML object. This object is understood to represent a specific type of L2 Data or Ancillary Data, as specified in [SRD] section 5.1. It may represent a data product in its entirety, e.g. after reprocessing by the Data Provider, or it may represent a subset, e.g. in interval-based processing and transfer modes.

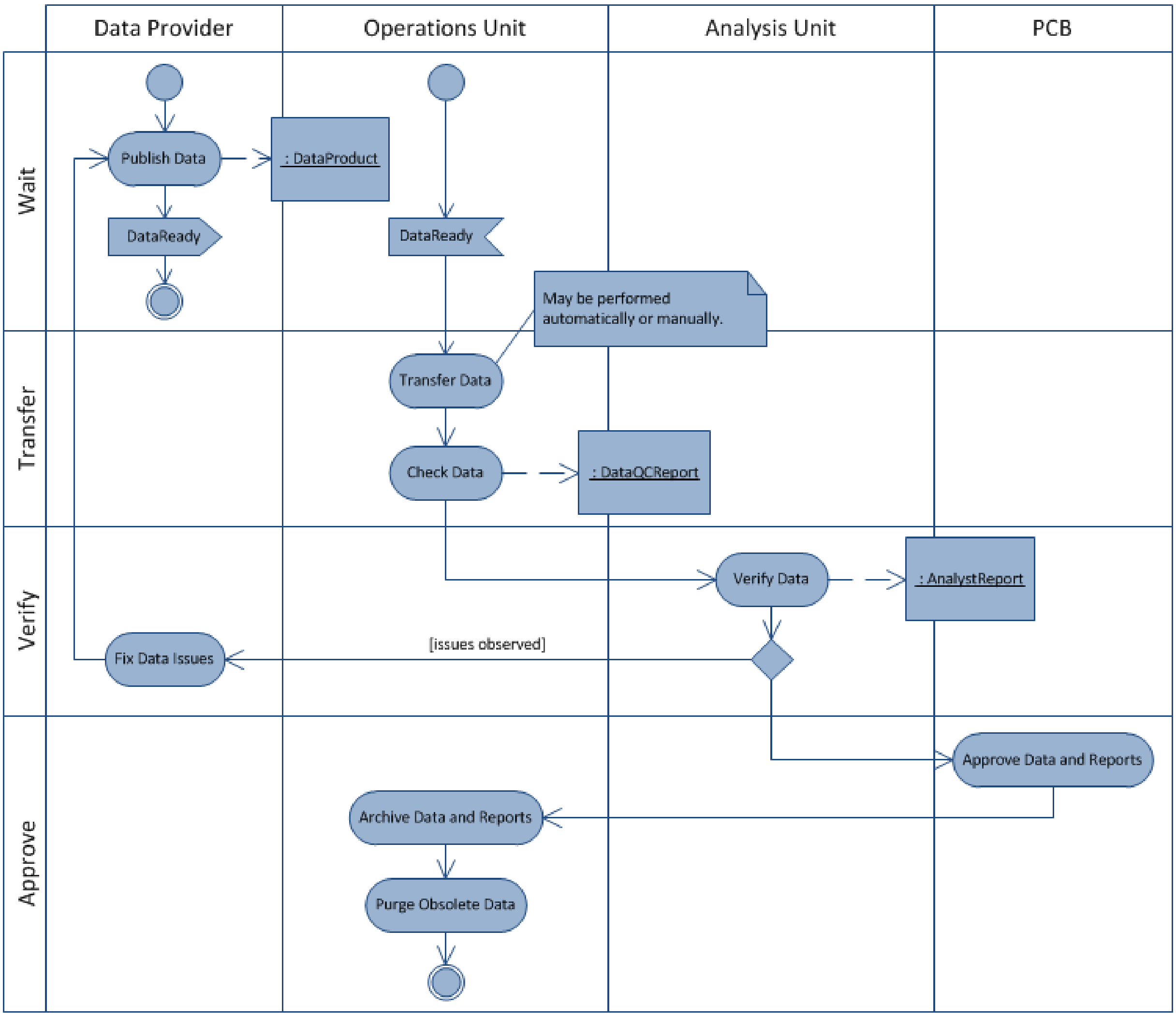


Figure 7: Activity Flow for acquisition of L2 Data and Ancillary Data

A Data Product may be transferred manually or automatically. E.g. certain L2 Data are planned to be transferred automatically once automated sender-receiver systems are in place on both sides (cf. [SRD] SR-0421), while until then L2 Data and in any case Ancillary Data will be transferred manually by Processing Engineers (cf. SR-0420). Automated transfers will be performed based on time intervals. In this case Automatic Interval Processing (cf. SR-0130) will start immediately after automatic Data QC in order to create a near real-time ECV Data Product as soon as possible and in a fully automated way. In the case of manual data transfers processing is manually scheduled by the Operations Unit, typically with a longer delay after the data have been received and verified.

The Analysis Unit verifies the received data and the results of automatic Data QC. Both the automatic and interactive verification results are recorded in the System. These records are represented by the DataQCReport and AnalystReport UML objects in Figure 7. Finally, the PCB approves the data together with both reports.

The process completes with registration and copy of data and verification records to the permanent Data Store and Data Archive and with the removal (purging) of any obsolete data or reports that may still exist in temporary stores.

5.4.2 ECV Production Process

Figure 8 shows the flow of activities for the production of a Soil Moisture ECV Data Product, be it for incremental extension of an existing product or full reprocessing. The original initiative to start the production may come from anywhere. However, for the purpose of a defined process the Analysis Unit formally initiates the production by issuing a Production Request (represented by the "ProdRequest" UML object in the diagram). The Production Request specifies the details of the Data Product to be produced and any complementary data required for the production run. The process subsequently progresses through preparation, generation and verification phases, and eventually completes with the publication of the new ECV Data Product, which may be either a new version of a previously published product, or an increment thereof, or the fully reprocessed data set, or even a new Data Product.

Once the Production Request is available the Operations Unit checks if all preconditions are met, that is, if the required input data are available in the System and have been verified and approved. The Production Request is then submitted to the PCB for approval. The PCB prioritizes Production Requests in order to manage data production with respect to Data Users' needs vs. the System's production capacity, and issues Work Orders to the Operations and Analysis Units to proceed with generation and verification.

The Operations Unit plans the production run. Then an assigned Processing Engineer configures an instance of the ECV Processor that meets the requirements of the Production Request. The full set of Configuration Parameters (shown as "ConfigPars" UML object in Figure 8) is later attached to the Data Product to allow for a repetition of the production run in the future.

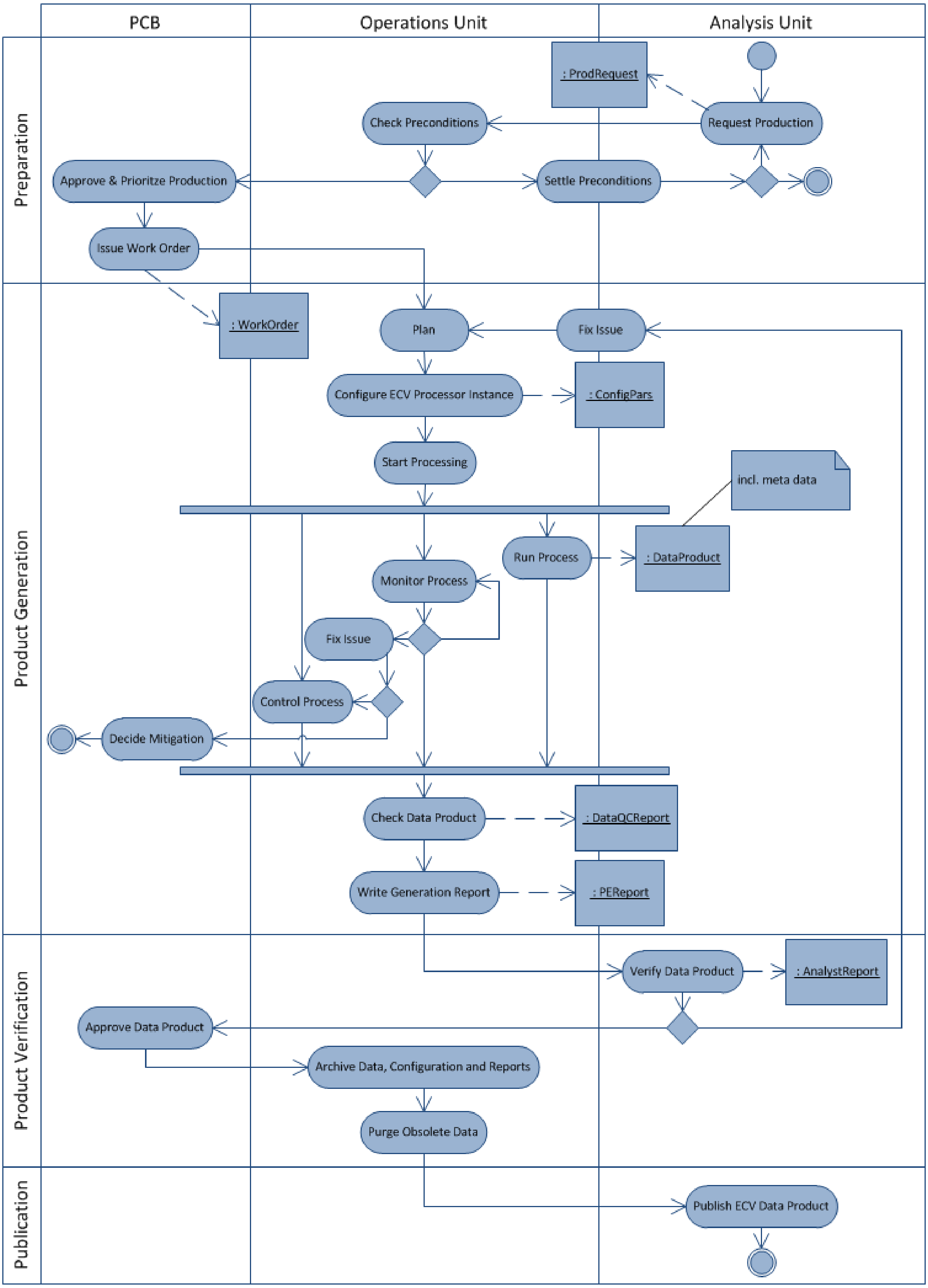


Figure 8: Activity Flow for the production of ECV Data Products

After starting the ECV Processing the Processing Engineer monitors the running processes and assures smooth processing by addressing technical issues which may be observed during execution time. Individual processing runs may take weeks to months to complete (cf. section 4.3, TM-014). Monitoring and issue fixing is supported by Process Monitoring & Control software (item A3.6 in Figure 3). The Operations Unit may request support from Analysis and Science Units in case issues related to the data or to Processing Algorithms are detected. In extreme cases the production run is aborted if relevant problems cannot be addressed within the production process. In such cases the PCB is involved to decide about next steps.

Upon successful completion of the processing run the newly generated ECV Data Product is available in the System's temporary Data Store. The Processing Engineer uses the Data QC software to check the new Data Product and to generate an automated Data QC Report. The product generation phase completes with the Processing Engineer compiling a report (PE Report) about the processing run including any events observed, issues detected and fixes or workarounds applied.

The new ECV Data Product together with the automated Data QC Report and the PE Report are made available to the Analysis Unit for interactive review and verification by Data Analysts, whose verification results are documented in an Analyst Report. At this stage Data Analysts and Processing Engineers may decide to fix observed issues or repeat parts of the production run to address remaining issues. Upon positive verification the assembled reports are submitted to the PCB for approval of the new product.

After approval the ECV Data Product, the applied configuration data, any intermediate data to be preserved, and the verification records are registered and copied to the permanent Data Store and Data Archive. Any obsolete data or reports that may still exist in temporary stores are purged.

The process completes with the publication of the new ECV Data Product on the Soil Moisture Web Portal making the new product available to the Data Users.

5.4.3 Software Change Process

Figure 9 shows the flow of activities associated with the resolution of reported software issues, whether they are to be addressed by the implementation of software changes in the Software Unit or by software installation or configuration changes in the Operations Unit.

The "Software User" column in Figure 9 represents both internal and external users. Data Users, Data Analysts, Processing Engineers and Algorithm Developers are prominent examples, but also Software Developers may be in the role of Software Users in the context of Figure 9.

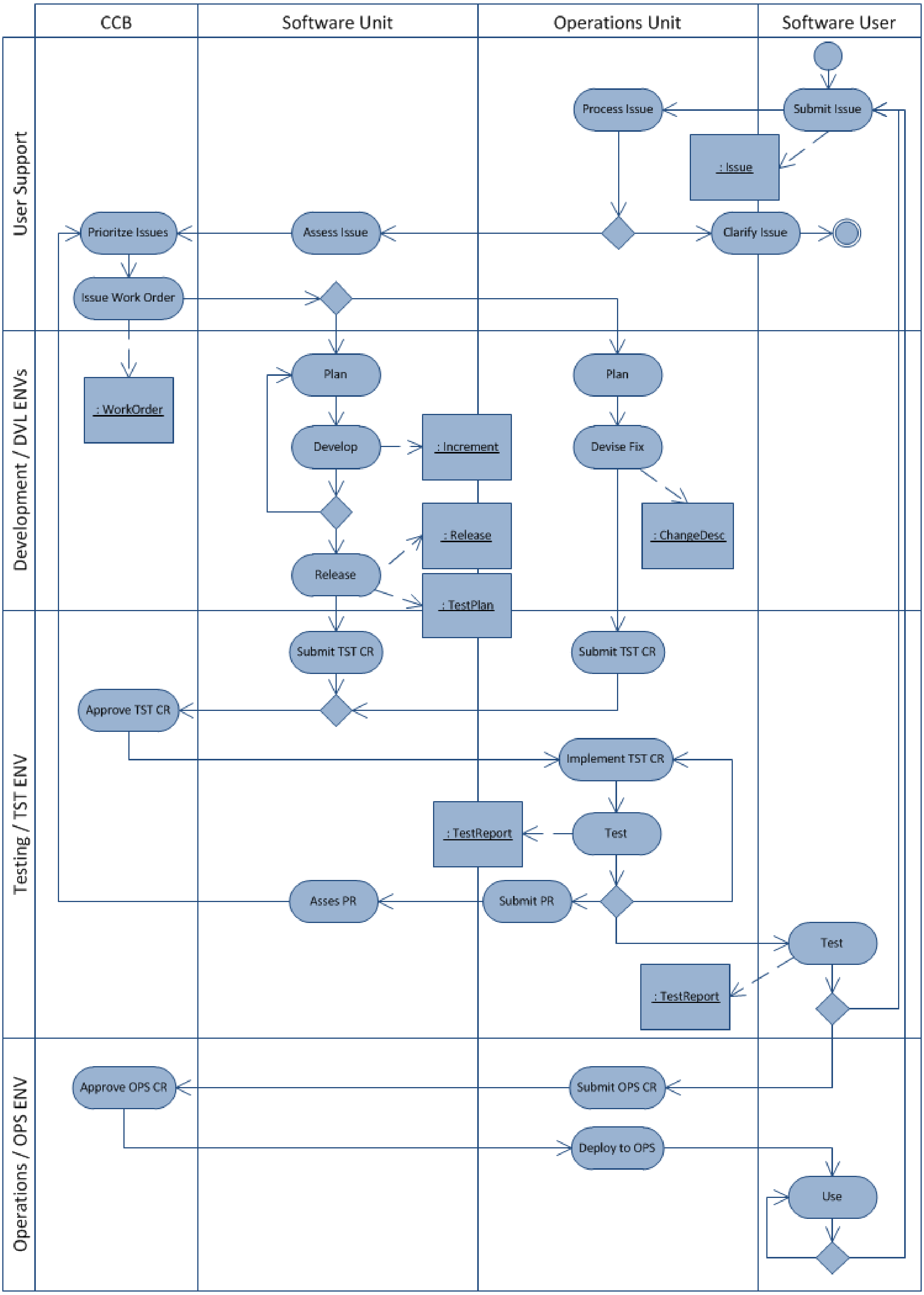


Figure 9: Activity flow for resolution of software issues, software development and testing

Software issues are represented by the "Issue" UML object in Figure 9. An Issue is here understood to represent any kind of software related request: a general problem report, a specific bug report, an enhancement request, a request for installation & configuration or a request for development of a new Software Application. An example may be an Algorithm Developer who wants to update the ECV Processor with new Processing Algorithms and submits an Issue with an appropriate enhancement request.

The software change process is supported by issue tracking software, which is used to manage all reported Issues and related artifacts like assessment reports, implementation proposals, test plans and test reports. The scope of reported Issues may widely vary from trivial change requests to complete developments projects. Issues are first received and assessed by the Operations Unit. Issues requiring changes to software applications and source code are forwarded to the Software Unit.

The CCB prioritizes reported Issues and issues Work Orders to the Software Unit if required. The CCB thus controls the development activities and balances them against available development resources.

The Software Unit plans and implements software changes within the development environments, i.e. the DVL ENVs shown in Figure 9. Software changes can also be implemented by external Software Developers. Dedicated virtualized development environments can be provided on DVL VMs together with the relevant development tools and test data regardless of the physical location of the assigned developers.

The installation of a Change Request (CR) is then proposed to the CCB, which controls the configuration of the TST and OPS ENVs. In cases where the Software Unit does not need to be involved the Operations Unit can directly submit a CR.

Approved changes are installed and tested by the Operations Unit on the TST ENV, typically based on release notes and test plans provided by the Software Unit or based on specific change descriptions. The Operations Unit compiles a Test Report with the results of the test, and in case of a positive test result, makes the Software changes available for user testing by the Software Users, still on the TST ENV. If the Software Users report no relevant Issues related to the tested change the Operations Unit submits an OPS CR to promote the changes to the OPS ENV.

The deployment to the OPS ENV is again done after approval by the CCB. Once deployed to the OPS ENV, the updated Software is ready for operational use. The Software Users may submit further Issues if new problems are observed or a need to change the software has been identified.

5.4.4 ECV Change Process

Figure 10 shows the flow of activities associated with changes to the ECV Data Products, which are triggered by issues reported by internal or external Data Users. Possible activities range from providing clarifications to the reporting Data Users, via new ECV Data productions with changed parameters, to full cycles of product development involving changes to product specification, scientific methods, processing algorithms and processing software. The Analysis Unit, Science Unit and PCB are the main actors of the ECV Change Process, which utilizes both the ECV Production Process (cf. section 5.4.2) and the Software Change Process (cf. section 5.4.3). The "Data User" column in Figure 10 represents both internal and external Data Users.

ECV Data Product issues are represented by the "Issue" UML object in Figure 10. An Issue is here understood to represent any kind of product related request: a problem report, a request for change of production parameters, a product enhancement request, a change in Data User requirements, or a request for development of a new Data Product.

The ECV change process is supported by issue tracking software, which is used to manage all reported Issues and related artifacts like product specifications, change proposals, evaluation plans and reports from Analysts and Data Users. The scope of Issues may widely vary from trivial clarification or change requests to requests for the development of new products.

Issues are first received and assessed by the Analysis Unit. The Analysis Unit may directly resolve Issues, e.g. by providing clarifications to the reporting users or by requesting an ECV Data production with appropriate (changed) parameters. In the latter case the ECV Production Process of section 5.4.2 will be invoked. Issues requiring changes of product specifications, scientific methods, processing algorithms and processing software are forwarded to the Science Unit, who assesses the received Issues and asks the PCP for prioritization.

The PCB prioritizes reported Issues and issues Work Orders to the Science Unit if required. The PCB thus controls the research and development activities with respect to the available resources.

The Science Unit implements the Work Orders, performing the appropriate planning, research and development activities. Updates of product specifications require approval by the PCB, as is the case for new product specifications. Algorithm development can be carried out on the DVL ENV. A first test product (shown as "TestProduct" UML object in Figure 10) is generated on the DVL ENV and submitted to the Analysis Unit for evaluation.

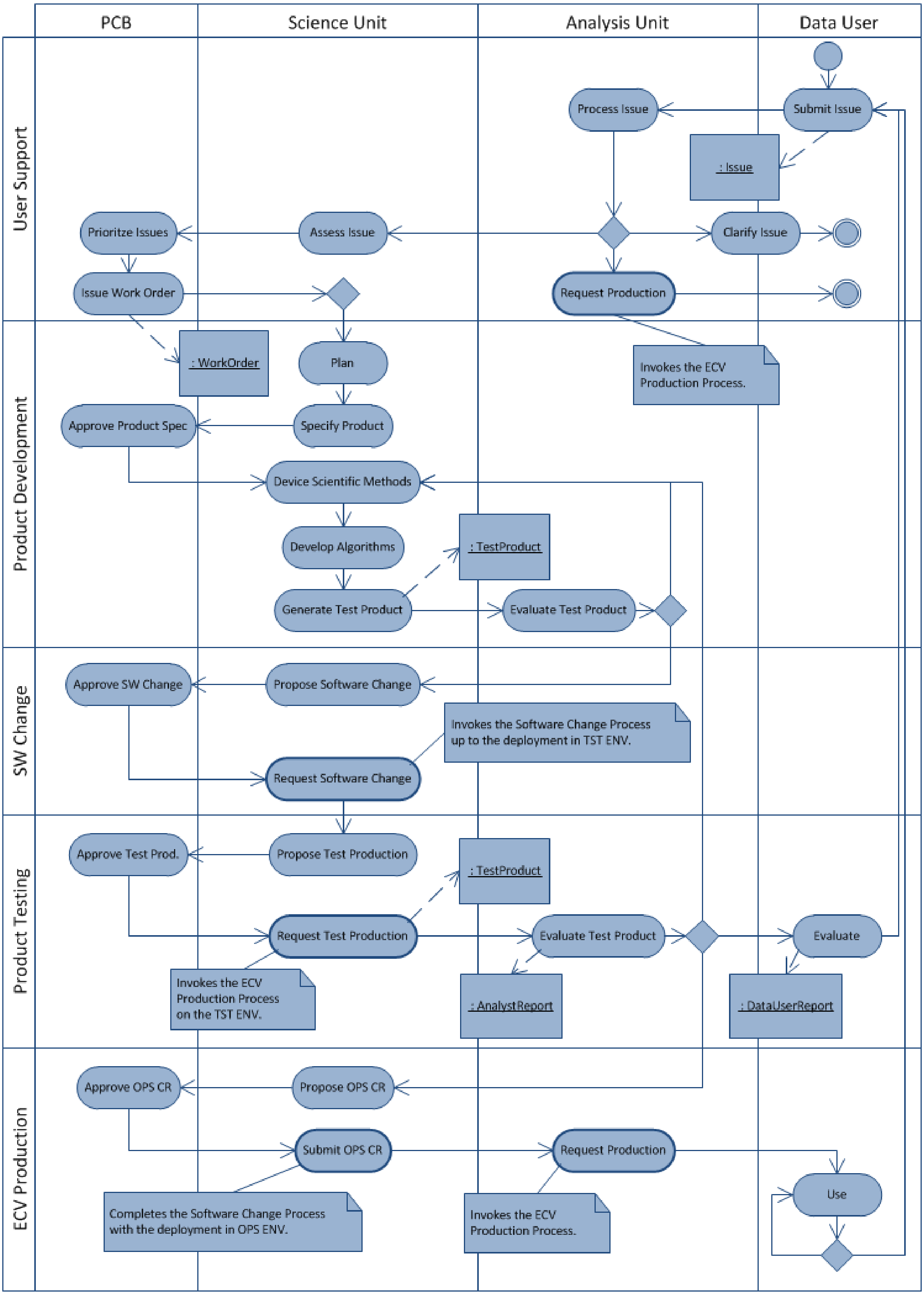


Figure 10: Activity Flow for resolution of ECV Data Product issues, and Product development

Required changes of processing software are then proposed to the PCB. Upon approval, the software change is implemented by invoking an extended Software Change Process – i.e. the process specified in section 5.4.3 is extended after deployment to the TST ENV by adding a cycle for generation and evaluation of a Test Product using the facilities of the TST ENV.

The generation of a Test Product is controlled by the PCB and performed on the TST ENV. The ECV Production Process of section 5.4.2 is invoked to carry out the generation of the Test Product on the TST ENV. The Test Product is evaluated by the Analysis Unit with feedback to the Science Unit. Evaluation of the Test Product may also be carried out by (external) Data Users if this is useful.

After successful product testing the PCB approves the deployment to the OPS ENV. Upon submission of the OPS Change Request (CR) the remaining steps of the Software Change Process are completed. The ECV Production Process is then performed on the OPS ENV to create the changed, improved or new ECV Data Product. Once the new ECV Data Product has been published Data Users may use it and submit further Issues, which will trigger the ECV Change Process again.

5.5 Use-Case View

This section specifies interactions of external Actors with the System in a Use-Case View. The View describes the System's response to externally initiated actions, while internally initiated actions and activities are specified by the Process View in section 5.4. The view is constituted by the set of Use Cases listed in Table 22:

Table 22: Main Use Cases of the Soil Moisture ECVPS

UC-#	Main Actor	Goal
UC-001	Data User	Download ECV Data Product
UC-002	Data User	View ECV Data Product
UC-003	Data User	Compare ECV Data Products

Each Use Case has a unique identifier and is described in the format shown in Table 23. Use Case descriptions typically do not refer to internal System Elements but treat the System as a single entity, which interacts with the Actors of the Use Case.

Table 23: Template for Use Cases (UCs)

Use Case Name: *A unique name given as a description of a User Goal*	
Unique ID	UC-#
Primary Actor	*Initiator of the interaction, typically the person having the User Goal.*
Preconditions	*What must be the case for the Use Case to start.*
Success Guarantees	*What will be the outcome of the Use Case.*
Main Success Scenario	*1. First step.* *2. Second step.* *3. …* *n. Last step.*
Extensions	*Alternative scenarios and outcomes amending or substituting specific steps.*
Source	*Reference to requirements, documents or persons with relevant input to the UC.*
Priority	*Priority value from MoSCoW schema as described in [SRD], section 3.7.*
Comments	*Any other comments for clarification and future consideration.*

The Soil Moisture ECVPS seen from outside as a single entity has only a small set of external User Goals, namely those identified in section 3.1 and listed in Table 22. There is a single external Actor, the Data User who approaches the Soil Moisture ECVPS in this way. Other external entities from Science Communities, Data Providers, Software Developers, Hosting Organization and Sponsors are integrated with the System's organization and processes (for example via their participation in the PCB and CCB boards); their interactions with the System are therefore described in the Process View in section 5.4.

Table 24: (UC-001) Download ECV Data Product

Use Case Name: Download ECV Data Product	
Unique ID	UC-001
Primary Actor	Data User
Preconditions	• User is registered at Soil Moisture Web Portal and duly authorized for data download. • ECV Data Product was published and is available in System's File Store.
Success Guarantees	ECV Data Product has been downloaded to User's local environment.
Main Success Scenario	1. User logs into Soil Moisture Web Portal. 2. Soil Moisture Web Portal checks credentials and admits User. 3. User navigates to the desired ECV Data Product and selects download link. 4. Soil Moisture Web Portal checks User's permissions and admits User. 5. Soil Moisture Web Portal sends an email to User's registered email address with download instructions and temporary login credentials for System's FTP download server. 6. User connects to FTP download server with their preferred FTP client. 7. User initiates FTP download and waits until download has completed. 8. User verifies in the local environment that the data was received and can be used for further analysis and visualization using local tools.
Extensions	2a. If User is not admitted, Soil Moisture Web Portal displays a self-registration page allowing the User to register and request permission for data download. 4a. If User is not admitted, Soil Moisture Web Portal displays an administration page allowing the User to request permission for data download.
Source	[SRD] SR-0210, SR-0460, SR-0470, SR-0480, SR-0490
Priority	MUST-HAVE
Comments	Soil Moisture Web Portal may be entered from the common CCI Web Portal. If requirement SR-0210 in [SRD] can be relaxed it may be recommendable to replace FTP by HTTP(s) for data download.

Table 25: (UC-002) View ECV Data Product

Use Case Name: View ECV Data Product	
Unique ID	UC-002
Primary Actor	Data User
Preconditions	• The ECV Data Product of interest was published and is available on System's File Store.
Success Guarantees	Subsets of ECV Data Product raster images and timeseries data have been viewed by User in a web browser.

Use Case Name: View ECV Data Product	
Unique ID	UC-002
Main Success Scenario	1. User navigates his web browser to Soil Moisture Web Portal. 2. User navigates to the desired ECV Data Product and selects the view link. 3. User's web browser displays the Data Viewer page showing the selected ECV Data Product at the last available data day. 4. User uses Data Viewer's inline zoom, pan and calendar tools to select the spatial-temporal area of interest. 5. User views ECV Data Product raster images overlayed on a geographic map and may select a geographic point of interest to display timeseries data for the selected point.
Extensions	4a. User may use Data Viewer's inline product selection tool to navigate to a different product or product version.
Source	[SRD] SR-0620
Priority	MUST-HAVE
Comments	Soil Moisture Web Portal may be entered from the common CCI Web Portal.

Table 26: (UC-003) Compare ECV Data Products

Use Case Name: Compare ECV Data Products	
Unique ID	UC-003
Primary Actor	Data User
Preconditions	• The ECV Data Products of interest were published and are available on System's File Store.
Success Guarantees	Spatial-temporally linked subsets of up to four ECV Data Products have been compared by User in a web browser.
Main Success Scenario	1. User performed UC-002, View ECV Data Product, and views the spatial-temporal subset of interest. 2. User uses Data Viewer's inline map utilities to display up to four maps in a side-by-side view and has the option to mutually link maps spatially and/or temporally. 3. Individually per displayed map, User uses Data Viewer's inline product selection tool to navigate to a product or product version of interest. 4. In one of the displayed maps, User uses Data Viewer's inline zoom, pan and calendar tools to select the spatial-temporal area of interest. 5. Data Viewer synchronizes and updates other displayed maps in correspondence with the selected link modes. 6. User compares ECV Data Product raster images overlaid on geographic maps side-by-side, and may select individually per displayed map a geographic point of interest to display and compare its timeseries data.
Extensions	
Source	[SRD] SR-0630
Priority	MUST-HAVE
Comments	Soil Moisture Web Portal may be entered from the common CCI Web Portal.

5.6 Information View

This section describes the System's data and information items in terms of an Information View by establishing the Information Model – which is a formal, conceptual and abstract representation of entity types that may include their properties and relationships, and the operations that can be performed on them. The Information Model provides a formalism for the description of the business domain without constraining how this description is mapped to an actual implementation in software. The content of this section was developed in collaboration with the Soil Moisture Data Cubes project [SMDC].

Because processing of Soil Moisture data from remote sensing Earth Observation (EO) satellites is the main System purpose, the scope of this section is widened to start with a survey of the specifically relevant fields: EO, remote sensing and Soil Moisture. The focus is on identifying the relevant terminology so that the features of the data can be adequately described. This section also reviews elements outside the System (i.e. context elements) – insofar as they play a role in the production and representation of Soil Moisture data.

The organization of this section is as follows:

- Section 5.6.1 provides the context description and survey of the relevant fields in the narrative form. This section will be most useful for readers without background in the fields of EO, remote sensing or Soil Moisture.
- Section 5.6.2 provides an external schema that lists and describes the key terms of each field. It may be used as a supplement to section 5.6.1, or autonomously as a reference of the System's business domain terminology.
- Section 5.6.3 provides the external schema of the Data Processing Levels. This schema is generally used in the EO field for the organization of data processing and the classification of EO data. It has been adapted to the System specifics. Examples were added to improve understandability.
- Section 5.6.4 provides the conceptual schema of the Soil Moisture Data Items. It shows the various types and representations of the business data acquired, stored, processed and emitted by the System, additional descriptive and administrative data items (metadata), and the relationships among them.
- Section 5.6.5 provides the conceptual schema for the Information Items generally present in an Information System and particularly required for System Monitoring & Control.

The Information Model formally defines the structure and semantics of the data and information items within the System and its environment in terms of schemata.

The schemata presented in sections 5.6.4 and 5.6.5 are conceptual, meaning they provide a single integrated definition of the System's data and information items which is a) not biased towards any specific application of data, b) independent of how the data is physically stored or accessed, and c) does not refer to a particular software implementation.

A schema may be specified at different abstraction levels (levels of detail). E.g. [IDEF1X] defines three levels: Entity Relationship (ER), Key Based (KB), and Fully Attributed (FA), where the ER level is the most abstract: It models the fundamental elements of the subject area – the entities and their relationships. It is usually broader in scope than the other levels. The KB level adds keys and the FA level adds all the attributes.

The schemata presented herein are at the ER level. For a description of the ER Model cf. [Chen76]. In order to distinguish different kinds of entity types the color code shown in Table 27 is used.

Table 27: Color Code used in ER Schemata for different Kinds of Entity Types

Color	Kind of Entity Type
Blue	Data Item of the Soil Moisture business domain.
Cyan	Status information about a Data Item.
Purple	Administrative data created by automatic processing.
Orange	Administrative data created by a human activity.
Red	Static / Descriptive Data Item.

The Information Model as specified herein must be clearly distinguished from a "Data Model" when the latter is understood in the sense of a mapping to an actual implementation in physical storage, database or software.

The data modeling language used is [UML], and the ER schemata are depicted in the form of UML class diagrams (with attributes and operations suppressed).

5.6.1 Context Description: Soil Moisture Data from Earth Observation Satellites

The objective of the System is to produce the most complete and most consistent global Soil Moisture ECV Data Products based on the measurements (observations) made by imaging microwave instruments (syn. sensors) flying onboard of earth observation satellites. Soil Moisture is one out of 13 geo-physical parameters recognized as Essential Climate Variable (ECV) under the framework of the ESA CCI initiative [Hollmann13], [Jeu12a] and [Jeu12b].

This section gives an overview of the different types of data and the methods, technologies and processes employed to generate the Soil Moisture Data Products. Because this section captures the complete process – from the measurement to the application/analysis – it constitutes an external and context view, with the System being a part.

5.6.1.1 Microwave Instruments onboard Earth Observation Satellites

There are two principal types of remote sensing, corresponding to the following types of microwave instruments: (a) scatterometers and radars which measure the radar backscattering coefficient σ^0 in physical units [dB] or [m^2/m^2], and (b) radiometers which measure the brightness temperature T_B in physical unit [K]. Instruments in group (a) are called *active* because they use their sown source of electromagnetic energy for the measurement, while the ones in group (b) are referred to as *passive* instruments because they measure energy that is reflected or emitted from the earth surface.

Figure 11 identifies the active (at the bottom) and the passive (at the top) microwave instruments that are used for the production of the Soil Moisture ECV Data Products, their hosting satellites, and their times of operation.

- AMI-WS and ASCAT are (active) C-band scatterometer instruments onboard the ERS and METOP satellites, respectively. Note that AMI-WS is still labeled “SCAT” in the picture, though the former name is nowadays preferred.
- SMMR, SSM/I, TMI, AMSR-E and Windsat are (passive) multi-frequency radiometer instruments onboard the Nimbus-7, DMSP, TRMM, Aqua and Coriolis satellites, respectively.

These instruments are characterized by their high suitability for Soil Moisture retrieval and their technological maturity.

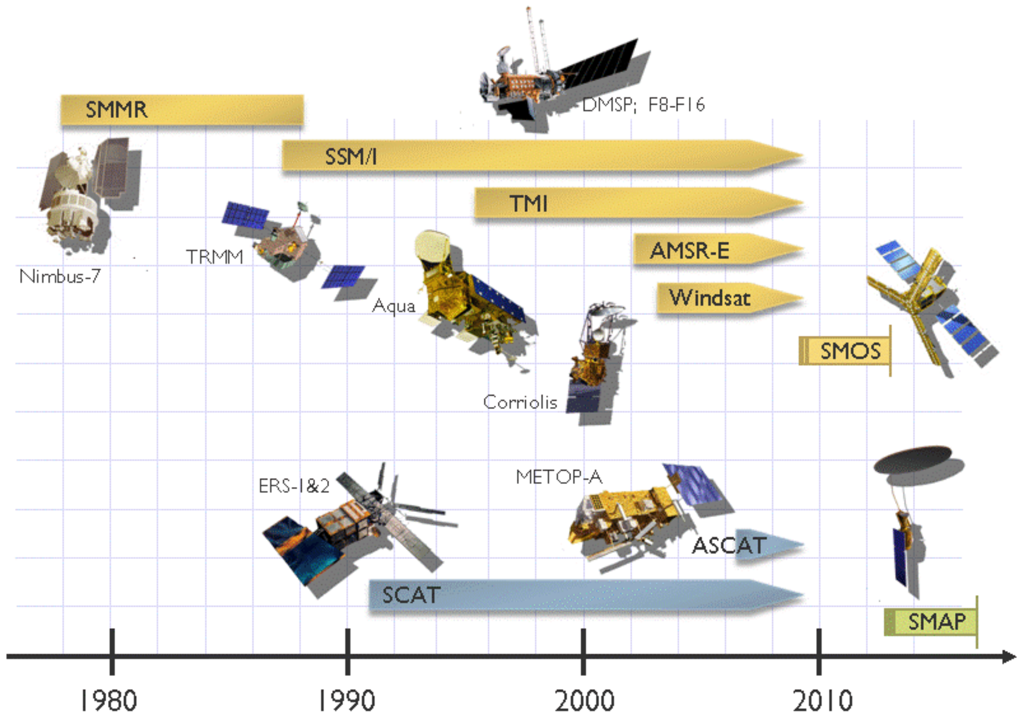


Figure 11: Microwave Instruments used for the generation of Soil Moisture ECV Data Products

The System is designed to allow the integration of new instruments during its operational use phase. Candidate microwave instruments for future use in the Soil Moisture retrieval are:

- MIRAS, a radiometer onboard the SMOS satellite observing in the L-band range and using aperture synthesis,
- A (passive) radiometer instrument observing in the L-band range and an (active) L-band SAR instrument onboard the SMAP satellite,
- A C-band SAR instrument onboard the Sentinel-1 satellite, and
- other Synthetic Aperture Radars (SARs) and radar altimeters.

These instruments are not considered in the initial operational phase due to their recentness or their unfavorable spatio-temporal coverage.

In a single overpass (one orbit around the earth) the satellite's instrument observes a wide swath of the land surface of a width somewhere between 500 to 1400 km. Satellite swath data is made of individual nodes (syn. pixels), each node being a measurement of either backscattering coefficient σ^0 (active) or brightness temperature T_B (passive) from an area (footprint) of 10-50 km diameter on the earth surface.

Figure 12 shows the geometry of the wide swath imaging measurements made by AMI-WS (SCAT) and ASCAT instruments onboard the ERS and METOP satellites. The AMI-WS scatterometer (left part of Figure 12) consists of three antennae producing three beams looking 45° forward, sideways (90°), and 45° backward with respect to the satellite's motion direction along the orbit. The measurements from each beam consist of 19 nodes spaced 25 km apart. As the satellite beams sweep along the earth surface yielding an approximately

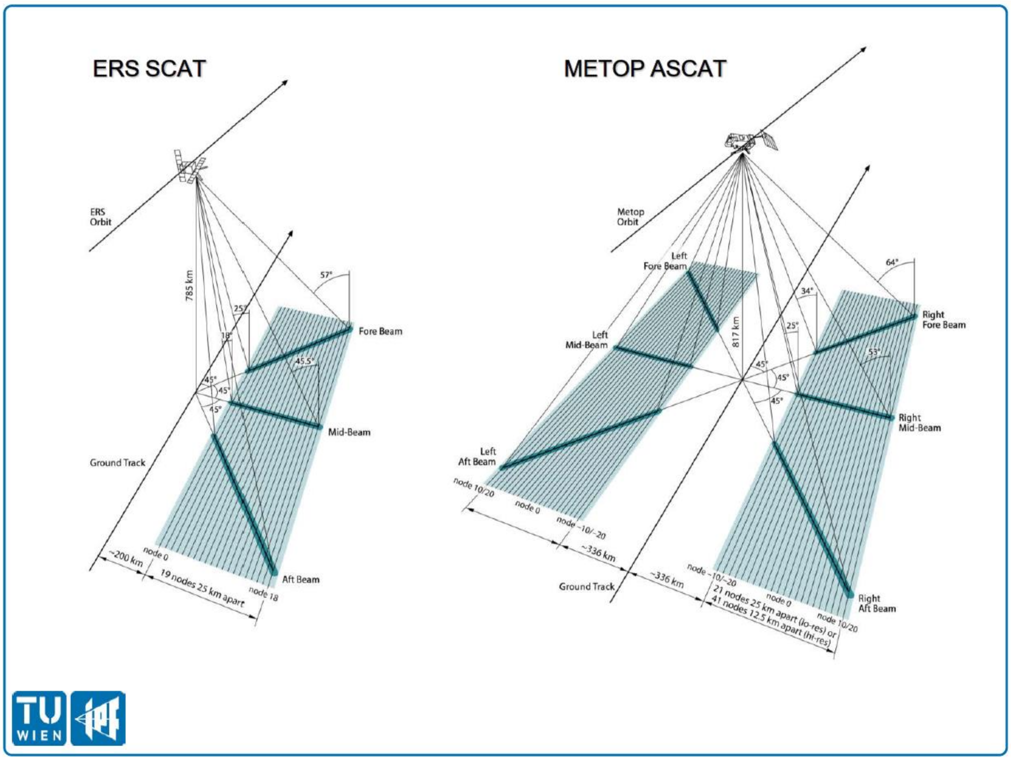


Figure 12: Beams and Wide Swaths of AMI-WS and ASCAT Instruments

500 km wide swath, each node produces its own σ^0 backscatter measurement, integrated over an area around 50 km in diameter. The three measurements originating from the three beams during the single satellite overpass are called triplets.

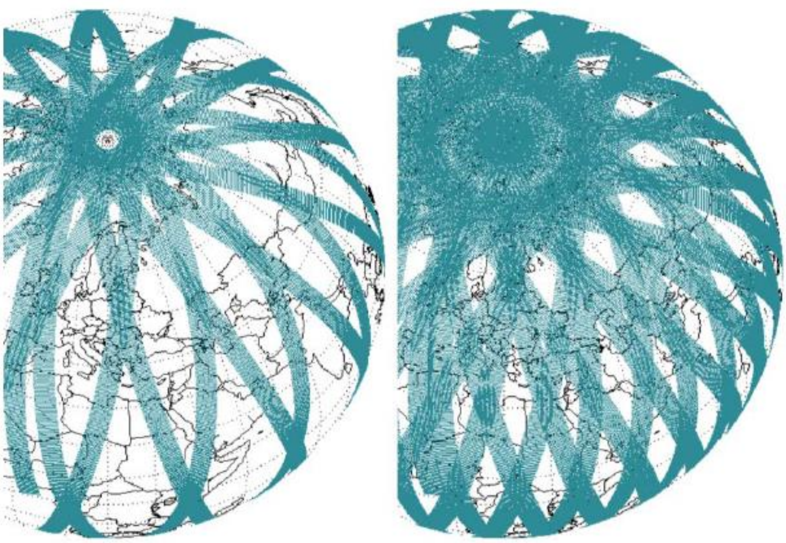


Figure 13: Wide Swaths and Daily Global Coverage of AMI-WS (left) and ASCAT (right)

The ASCAT scatterometer (right part of Figure 12) is similar but consists of three antennae on either side, producing six beams. The measurements from each beam consists of 21 nodes spaced 25 km apart (lo-res) or 41 notes spaced 12.5 km apart (hi-res). While the satellite beams sweep along the earth surface a wide swath is observed on either side.

Figure 13 shows the orbits made by a single ERS satellite (left) and a single METOP satellite (right) and the wide swaths captured by AMI-WS and ASCAT instruments, respectively, during one day of observation time. The wide swaths of ASCAT have near-global revisiting times of 2–3 days.

5.6.1.2 Data Processing Chain

The payload data received from a satellite passes through a chain of processing which has the following principal steps:

- Level 0: Reconstructing the raw instrument data from a satellite's payload data transmissions received at satellite ground stations, and archiving them at a data center.
- Level 1: Converting the raw instrument data of a single overpass swath to calibrated measurements (observations) of radar backscattering coefficient σ^0 in [dB] or [m^2/m^2] (in case of active instrument), or brightness temperature T_B in [K] (in case of passive instrument).
- Level 2: Retrieving the Surface Soil Moisture (SSM) geophysical parameter – in physical units [%] or [m^3/m^3] – from the calibrated measurements of a single overpass swath.
- Level 3: Generating the consistent and global Soil Moisture ECV Data Products by merging the SSM timeseries acquired from many satellites over many years – this is the objective of the System specified herein.
- Level 4: Analyzing the lower level data to determine the state or change of state (events, processes) of regional or global climate systems.

Below is a brief survey of the different processing levels tailored to Soil Moisture. There is also a fine-grained organization within the levels, in terms of sublevels or product categories, but these are suppressed in the first survey. A full specification of data processing levels is available in section 5.6.3.

Levels 0-1:

The steps up to Level 1, from the instrument to the calibrated measurement result, follow the principles of measurement as outlined in [JCGM-VIM] and [JCGM-GUM].

Note that a measurement (syn. observation) is not necessarily made with a single antenna, instrument, or even a single satellite. Crucial is that the individual measurement devices contributing to the measurement result are synchronized (with sufficient accuracy in space and time) to the spatial-temporal coherency characteristics of the observed source.

Therefore, the calibrated measurement may be summed from many antennas (e.g. in case of the SMOS MIRAS instrument, which uses aperture synthesis) or even summed from satellites in formation flying. In any case, because the observations made at different orbits do not satisfy the coherency criteria, Level 1 data are observations from a single overpass.

Levels 2-3:

The situation changes for the Levels 2 and 3, which are about the art of estimating values of quantities that cannot be directly measured with remote sensing instruments. The mathematical-physical basis for this process is that there is a functional relation between object parameters X and instrument observables Y. Symbolically, the relation may be written as Y = f(X) (forward model) or X = g(Y) (inversion). It is known that at any time object parameters X and observables Y have definite values. Once sufficient information is obtained from measurements of Y, estimates of X can be determined under certain conditions, iff there is sufficient knowledge of a suitable model or inversion. To indicate this radical difference, this process is no longer called measurement but “retrieval”.

In the case of the ECV Production System the observed object is an area of the top soil surface layer (of approx. 25x25 km size and < 2 cm thickness), and the retrieved geophysical parameter is Soil Moisture, expressed in the physical unit [%] or [m^3/m^3].

Depending if the measurement was made with an active or a passive microwave instrument, respectively, either a change detection approach or a forward modeling method is utilized. Both methods, like geophysical parameter retrieval in general, require use of additional ancillary data, which can be long-term reference data (historical data, climate data) or model parameters.

Levels 2 and 3 have in common that the retrieved parameters are local state variables: they characterize the state of a local and compact object in space (i.e. a local physical system), at a moment or period in time. They are thus distributions in space (over the land surface in case of Soil Moisture) and time. The difference between Level 2 and 3 is that the parameter retrieved by a process of Level 2 adheres to the locality condition of a single Level 1 measurement, while this condition is relaxed at Level 3. In practice, Level 3 is about merging Level 2 data from different observation times, orbits or instruments.

For full descriptions of the retrieval and merging methods cf. [ATBD] and [Wagner12].

Level 4:

At Level 4 analyses of data from Levels 3, 2 or 1 are performed in order to determine the state or the change of state of specific geo-physical systems of interest, in particular regional and global climate systems. Detection of change of state typically involves statistical analysis methods (e.g. trend analysis) and process model calculations (simulations) using lower level data as input.

Present at all Levels of the data processing chain (but omitted in the discussion so far) are processes to determine measures for the degree of definiteness of the resulting data. At Levels 0 and 1 an adequate measure is the measurement uncertainty as described e.g. in [JCGM-GUM], and at Levels 2 and 3 the error characterization as described in [CECR]. We do not know what measures are used at Level 4 but we note that results of analyses and simulations may be in the range of evidentiary to indicative.

5.6.1.3 Grid Types

The data acquired by a satellite instrument constitute individual measurements per swath node (actually one measurement result per beam, overpass and swath node) and are spatially arranged in the geometry of a *swath grid*, as illustrated in Figure 12. Swath grids are dynamical (with respect to the earth surface) because they are a projection made by the moving satellite. The Level 0 data are necessarily raster data from a swath grid. Typically, Level 1 and 2 data provided by ESA, EUMETSAT, NASA and JAXA are still in the geometry of swath grids, and annotated with time- and geo-referencing information to allow inference of time and geo-location of each swath node. Data in swath grid geometry are usually distributed as one raster per half-orbit i.e. one raster for the ascending orbit direction and a second raster for the descending orbit direction. The format of geo-referencing information varies; it may be given per raster as a whole or even as (latitude, longitude) coordinates per grid node.

Technically, it can be of advantage to map the swath grid nodes, onto a (spatial) *global grid*[6] which is uniformly and statically covering the earth surface. Cases where mapping to a global grid is already performed at Level 1 are Nimbus-7 SSMR, DMSP SSM/I and SMOS MIRAS, and at Level 2 METOP ASCAT. Note that a global grid maps the space dimensions only, while time-referencing annotation is still maintained for the individual measurements.

Figure 14 shows the equally spaced, equal area WARP 5 global grid used for the ASCAT Level 2 SSM data. The principle of construction (indicated in the upper part of Figure 14) is as follows: a) create equally spaced latitude small circles at distances *a*, and b) on each latitude circle create grid points at discrete longitudes, again with a fixed spacing of *a*. Due to the construction the global grid has areas with irregularities. The lower part of Figure 14 shows: a) the regularity at origin, b) the divergence at the North Pole, and c) the dislocation at the 180° meridian.

[6] The technical term is Discrete Global Grid (DGG), however, as we do not know how a non-discrete (i.e. continuous) grid would look like, we do not use this term herein.

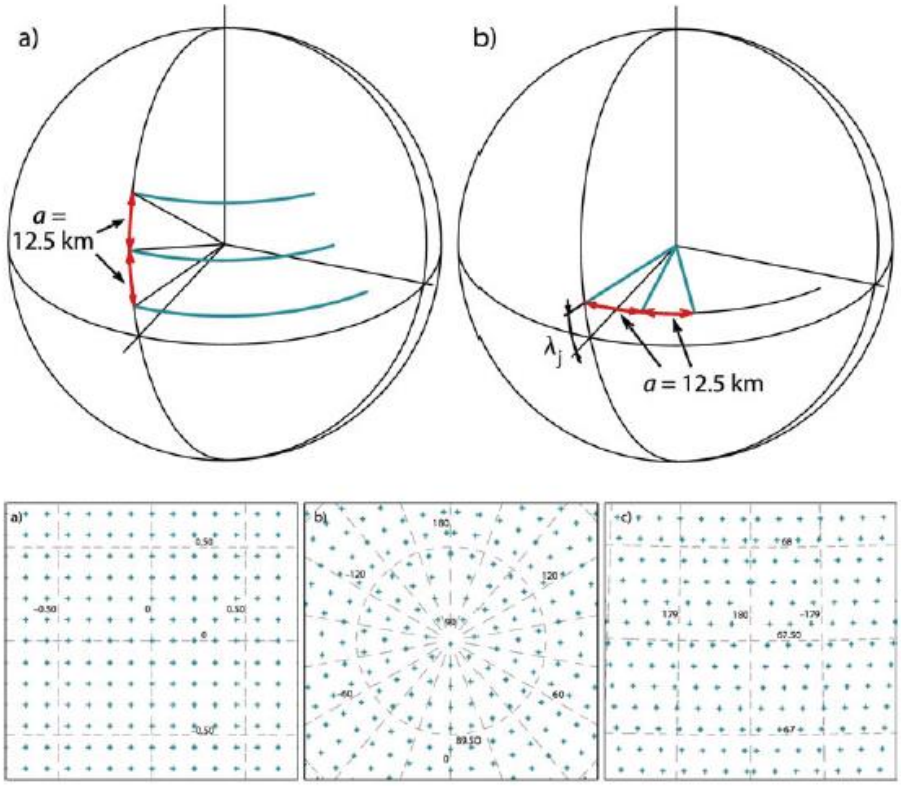
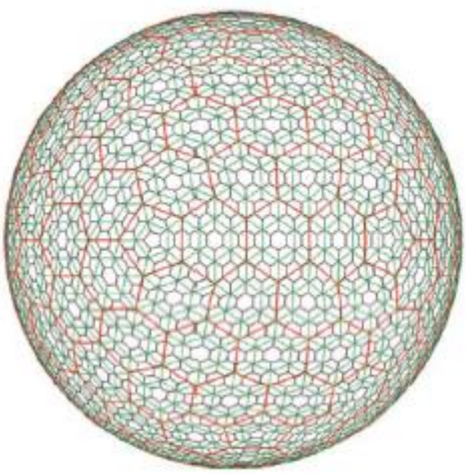


Figure 14: Construction of the WARP 5 Global Grid used for ASCAT L2 Data

Figure 15 a) shows the geodesic grid used for the SMOS Level 1 and Level 2 data, known as Icosahedron Snyder Equal Area (ISEA) Aperture 4 Hexagonal (ISEA4H) global grid. This grid is constructed by a subdivision method. Creating the grid involves subdividing the 20 equilateral triangles forming the faces of the regular icosahedron into more triangles, yielding 20 hexagons and 12 pentagons on the surface of the sphere (the so-called resolution 1 grid). Higher resolution grids are formed iteratively by tessellating the obtained shapes.

a)



b)

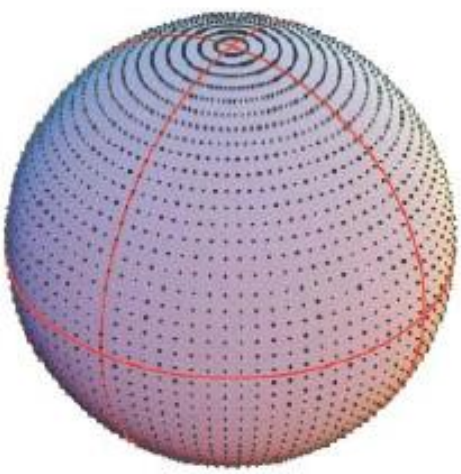


Figure 15: a) Geodesic ISEA4H grid of SMOS; b) regular latitude-longitude grid used for SM ECV

Figure 15 b) shows a regular latitude-longitude grid. Grid points are equally spaced in the latitude-longitude coordinate system, however, as can be readily seen in Figure 15, they are not equally spaced / equal area on the earth surface. This grid type is utilized for the Soil Moisture ECV Data Products.

Still for technical reason, it can be of advantage to proceed further and to map the value distributions onto a *global space-time grid*. This happens usually at Level 3 because it facilitates combining data from many orbits and satellites. Satellite data mapped to a space-time grid can be viewed in two ways: For a given time index we have a raster image that can be projected to and visualized on a geographic map; for a given spatial grid point we have a timeseries that can be visualized as a graph.

Table 28: Grid Systems used for Calibrated Measurement Data and SSM Geophysical Parameter Data

		Active		**Passive**					
		ERS AMI-WS	**METOP ASCAT**	**Nimb.7 SMMR**	**DMSP SSM/I**	**TRMM TMI**	**Aqua AMSR-E**	**Coriolis Windsat**	**SMOS MIRAS**
σ^0 / T_B	**L1**	swath	swath	EASE	EASE	swath	swath[7]	swath	ISEA4H
SSM	**L2**	WARP5	WARP5	lat-lon	lat-lon	swath	swath	swath	TBD

Table 28 lists the grid types used for the Level 1 calibrated measurements of radar backscattering coefficient σ^0 and brightness temperature T_B, and the SSM geophysical parameter data.

In the process of generating the Level 3 Soil Moisture ECV Data Products, the SSM geo-parameter data available in the different grid types are converted into a common space-time grid: a latitude-longitude grid (spatial resampling) and a reference time of 0:00 UTC (temporal resampling). This is the first step to produce a merged and harmonized ECV Data Product on a global and uniform space-time grid with a spatial resolution of 0.25 degree and a temporal resolution of one day.

A web-based tool to visualize and search for grid point information (e.g. geo-location, grid index, surface type) is available here: http://www.ipf.tuwien.ac.at/radar/dv/dgg/

For more information about global grid systems cf. [Kidd05] and [Zoltan09].

[7] Product name given in [DARD] is "AMSR-E/Aqua L2A Global Swath Spatially-Resampled Brightness Temperatures", but we count it as a L1 product. The product is in swath geometry although the name seems to indicate a global grid.

5.6.1.4 Data Flow and Soil Moisture Data Processing Modules

Figure 16 shows the Soil Moisture data processing modules of Level 2 and Level 3, the latter constituting the Soil Moisture ECV Production System (ECVPS, the System). Level 1 input data are shown on the left side, the produced Level 3 ECV Data Products on the right. The arrows indicate the data flow between processing modules.

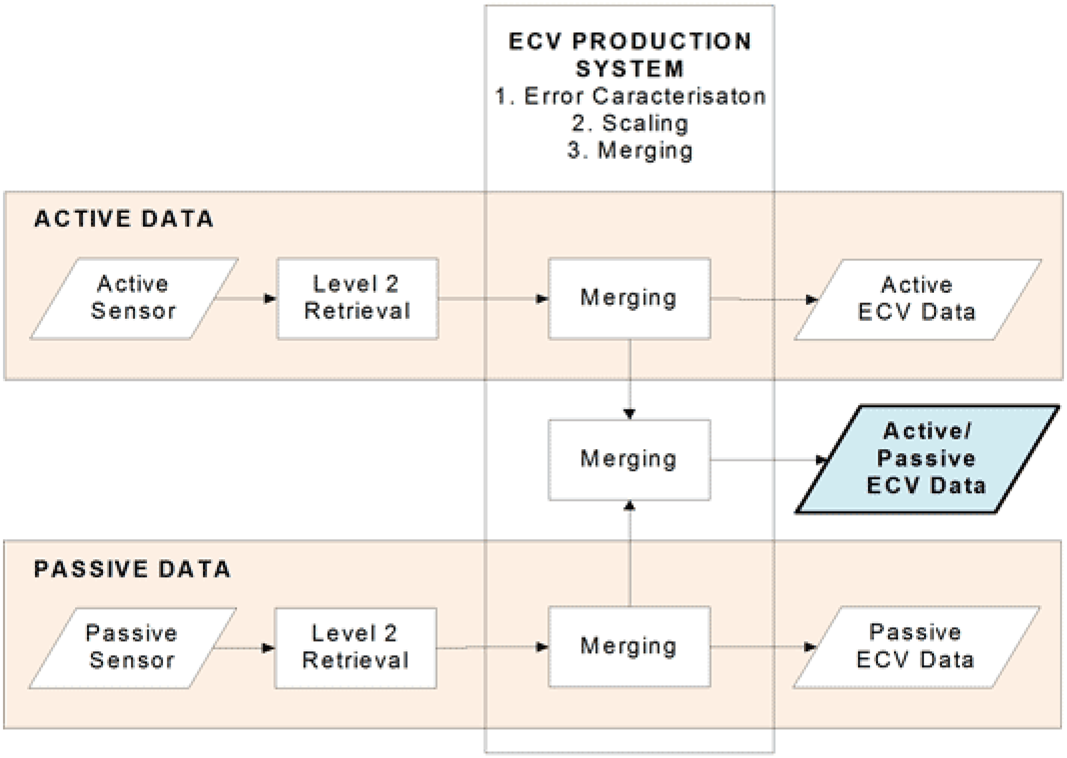


Figure 16: Principle Data Flow for the generation of Soil Moisture ECV Data Products

The overall setup of systems follows a modular design to make best use of existing European and international services, and to ensure that new satellites and instruments such as SMOS, SMAP and Sentinel-1 can be integrated.

Level 1 calibrated measurements of (active) radar backscattering coefficient and (passive) brightness temperature, the outputs from the Level 1 processing step (cf. section 5.6.1.2), are shown in Figure 16 as items labeled "Active Sensor" and "Passive Sensor", respectively, and are available from the major providers of satellite data ESA, EUMETSAT, NASA and JAXA. A complete listing of the required Level 1 data is available in [DARD].

The conversion of the Level 1 calibrated measurements to the Level 2 Surface Soil Moisture (SSM) geo-parameter is performed individually per instrument and labeled "Level 2 Retrieval" in Figure 16:

- SSM parameter retrieval from the calibrated measurements of each (active) microwave scatterometer instrument is performed with the WARP change detection algorithm of TUW.

- SSM parameter retrieval from the calibrated measurements of each (passive) microwave radiometer instrument is performed with the LPRM algorithm of VUA and NASA.

SSM parameter retrieval is implemented using the established Level 2 services of EUMETSAT, NASA, and JAXA that they provide for the different satellites and instruments. Descriptions of the processes can be found in [DPM] sections 4.1 and 4.2 for the active and passive cases, respectively.

The Soil Moisture ECVPS performs the final processing step, accomplished by three major modules according to Figure 16:

- Active Merging[8]: (1.) The full Level 2 SSM data set from each active instrument is resampled to a global space-time grid, (2.) each timeseries at each grid point is rescaled to the reference climatology of ASCAT, and (3.) all timeseries at each grid point are merged to a homogenized “Merged Active-only ECV Data Product”.
- Passive Merging: In principle, follows the same pattern of the Active Merging, but has some extra complications (pre-processing to convert swath data into timeseries, decomposition of the AMSR-E and SSM/I timeseries into their seasonality and anomalies, incremental merging), and the rescaling is made to the reference climatology of AMSR-E, yielding a homogenized “Merged Passive-only ECV Data Product”.
- Active-Passive Merging: (1.) The timeseries of the Active-only and Passive-only ECV Data Products are rescaled at each grid point to the reference climatology of “GLDAS Noah Land Surface Model L4 3 Hourly 0.25 x 0.25 degree Soil Moisture Estimate”, and (2.) merged to a homogenized “Merged Active-Passive ECV Data Product”.

For detailed listings and descriptions of the required Level 1 and Level 2 data cf. [DARD], for Level 2 see also [SRD] section 5 and [IODD]. The internal data flows of the Level 3 ECV generation are specified in [IODD]. The descriptions of the merging algorithms are in [ATBD] and [Wagner12]. Full process descriptions of the Level 2 active and passive parameter retrieval modules and the Level 3 ECV generation modules are available in [DPM].

A prototype system implementing the three merging modules of the Level 3 ECV generation was developed under the ESA CCI Phase 1 Soil Moisture project. For descriptions of the Soil Moisture Prototype ECV Production System (SM PECVPS) cf. [SPD] and [SVR].

---

[8] Note that terms ‘merging‘, ‘blending‘ and ‘fusion’ appear in the literature, sometimes synonymously and sometimes with diversified meanings. We do not make a distinction herein and just call it ‘merging’.

5.6.1.5 Soil Moisture Data Products

As an example of a Soil Moisture product, Figure 17 shows the Soil Moisture climatology for August derived from combing measurements of six satellite instruments over the period 1979 to 2010.

Surface Soil Moisture (SSM) parameter data from ERS AMI-WS and METOP ASCAT are an example of a Level 2 data product. It is retrieved from calibrated measurements of an active microwave instrument using the change detection method of TUW. The method relies upon the multi-incidence observation capabilities of the ERS and METOP scatterometers to model the effects of vegetation phenology. The SSM values are scaled between 0 and 1, representing zero Soil Moisture and saturation respectively. Parameter retrieval is not possible over tropical forest which affects about 6.5% of the land surface area.

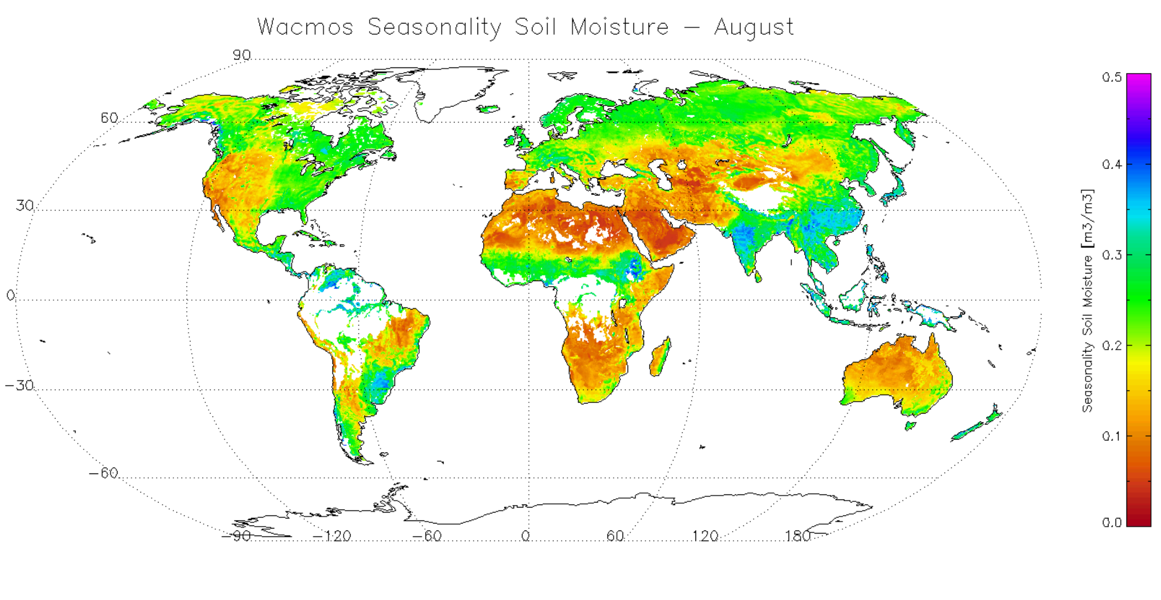


Figure 17: Soil Moisture Climatology for the Month August

Soil Water Index (SWI) parameter data are an example of a Level 3 data product. SWI is a measure of the profile Soil Moisture content obtained by filtering the SSM timeseries with an exponential function.

Other typical examples of Level 3 products are the estimates of the Soil Moisture content for different soil layers and different temporal and spatial sampling characteristics, tailored to the needs of specific user groups.

SSM and SWI from METOP ASCAT are mapped onto a global grid and can be viewed online here: http://www.ipf.tuwien.ac.at/radar/dv/ascat/

The SM ECV Data Products that will be produced by the System are further examples of Level 3 products. Here, the objective is to generate complete and consistent global-continuous Soil Moisture products for use in climate modeling. These products are combing

Soil Moisture data that was retrieved by utilizing many earth observation satellites and constitute a merged and harmonized data product on a global and uniform space-time grid with a spatial resolution of 0.25 degree and a temporal resolution of one day.

The theoretical and algorithmic base of the Soil Moisture ECV Data Products is described in [ATBD], and reported by [Wagner12]. An overview of all known errors of the Soil Moisture Data Products is provided in [CECR]. The product specification is in [PSD]; the listing and description of the used input data is in [DARD].

The SM ECV Data Products can be viewed online here:
http://www.ipf.tuwien.ac.at/radar/dv/wacmos+cci/

At the time of writing available are: Surface Soil Moisture (SSM) Monthly Mean in [m^3/m^3], and the Surface Soil Moisture Monthly Anomaly in [m^3/m^3], merged from the Level 2 data retrieved from measurements made by the AMSR-E, SSM/I, AMI-WS and ASCAT instruments.

Soil Moisture has an important role within the environment and climate system. It influences hydrological and agricultural processes, runoff generation, drought development and many other processes. It also impacts on the climate system through atmospheric feedbacks. Soil Moisture is a source of water for evapotranspiration over the continents, and is involved in both the water and the energy cycles. Soil Moisture was recognized as an Essential Climate Variable (ECV) in 2010.

5.6.2 External Schema: Terminology of Remote Sensing and Earth Observation

Table 29 lists the prime business domain terms from the scientific and technological domains of remote sensing, satellite-based Earth Observation and Soil Moisture retrieval. These terms emerged from an analysis of the text in section 5.6.1 and the referenced documents.

Table 29: Business Domain Terminology: Remote Sensing, Earth Observation, Soil Moisture

Context	Term (Acronym)	Description	Synonyms
1.	**Remote sensing**	Is the acquisition of information about an object or phenomenon without making physical contact with the object. There are two main types of remote sensing: passive remote sensing and active remote sensing, corresponding to active and passive instruments, respectively (see below).	
1.1.	Sensed object	The object subject of a remote sensing activity.	Sensed target Sensed system
1.2.	Instrument	In the given context: An imaging microwave instrument flying onboard of an earth observation satellite. The instrument is a payload of the satellite. In a wide sense: Any device that can be used to perform a measurement.	Sensor Measurement device
1.2.1.	Active instrument	An instrument that uses its own source of electromagnetic energy for the measurement.	
1.2.1.1.	Scatterometer	An active microwave instrument, measuring the radar backscattering coefficient $σ^0$ in physical units [dB] or [m^2/m^2].	Radar
1.2.2.	Passive instrument	An instrument that measures energy that is reflected or emitted from the sensed object.	
1.2.2.1.	Radiometer	A passive microwave instrument, measuring the brightness temperature T_B in physical unit [K].	
1.3.	Beam	The direction sensed by single antennae; used to express the fact that a satellite instrument may have several antennas each with its own field of view (FOV).	
2.	**Satellite**	In the given context always means Earth Observation (EO) satellite.	
2.1.	Orbit	The path of a satellite on its way around the earth.	
2.1.1.	Single overpass Single orbit	The term is used to stress the fact that a satellite instrument scans only once across the earth surface and makes a single or a tuple of measurements per swath node. Implies that individual beams contributing to a measurement result are synchronized to the sensed object.	Single scan

Context	Term (Acronym)	Description	Synonyms
2.1.2.	Revisiting orbit	The term is used to stress the fact that measurements of a satellite instrument were not made in a single overpass.	Overlapping orbits Multi-pass orbits
2.2.	Swath	The area of the earth surface along the ground track of a satellite that is sensed by the instrument in a single overpass; has a definite width but indefinite begin an end; made of an explicit number of nodes in the width-dimension that repeat in regular intervals along the direction of the ground track; each node is the sensed object of a single or a tuple of measurements.	
2.3.	L<n> Level Processing Level	Processing levels L0 to L4 are an organization of the chain of data processing performed on the payload data from EO satellites. Processing levels are described in detail in section 5.6.3.	
2.3.1.	Merging	Process of combining the measurements or retrievals from revisiting orbits of a satellite or from many satellites with the aim to reach an improved coverage of the observed domain in space or time. Terms 'merging', 'blending' and 'fusion' appear in the literature sometimes synonymously and sometimes with diversified meanings.	Blending Fusion
3.	**Grid**	Narrow: A set of nodes, lines or areas in a two-dimensional space that are arranged in a repeated pattern; may result from a construction principle or an actual process that repeats in regular intervals. Wide: Extended to n-dimensional space, and to space-time.	
3.1.	Node	One point (or area) out of a set of similar points (areas) that together form a grid.	
3.2.	Grid type	Classification of grids by construction principle.	
3.2.1.	Swath grid	The grid originating from a satellite instrument when it senses the individual areas of the earth surface in a repetitive pattern during a single overpass; each grid node represents a single or a tuple of measurements. A swath grid is the result of the projection made by the instrument's beams; images in swath grid are in the perspective of the instrument's point of view (POV). Nodes of revisiting orbits do not coincide, thus swath grids are dynamic and not statically linked to the earth surface.	Data in orbit geometry. Data in instrument projection. Image in the perspective of the instrument's point of view (POV).

Context	Term (Acronym)	Description	Synonyms
3.2.2.	Global grid	A spatial grid that is uniformly and statically covering the earth surface. Some global grids use a geoid model (e.g. WGS84 reference ellipsoid) to account for the earth's oblateness. Swath grids can be mapped onto global rids by means of a spatial resampling.	Discrete Global Grid (DGG) (Global) grid system
3.2.3.	Global space-time grid	The Cartesian product of a global grid and a timeseries index.	
3.3.	Geo-referencing	Linkage of grid nodes to points of the earth surface; expressed in latitude and longitude coordinates. Geo-referencing parameters (e.g. satellite ephemeris) allow linking measurement results with the remotely sensed objects on the land surface.	
3.4.	Time referencing	Linkage of a timeseries index (e.g. swath nodes) to UTC time.	
4.	**Measurement**	The process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity.	Observation
4.1.	Quantity	A property of a phenomenon, body, or substance, where the property can take on a value that can be expressed as a number and a reference.	Observable Variable
4.1.1.	Quantity value	A pair of number and reference, together expressing the value of a quantity.	
4.1.2.	Reference	A reference can be a measurement unit, a measurement procedure, a reference material, or a combination of such.	Scale
4.2.	Measurand	A quantity that is intended to be measured.	
4.3.	Calibration	An operation performed for an instrument that establishes a relation between measurement standards and the quantity values and uncertainties indicated by the instrument. In case of L1 data, radiometric and geometric calibration coefficients may be either attached (L1a) or applied (L1b).	
5.	**Distribution**	A function that assigns values to the points of a space, time or space-time domain; the assignment may be continuous or discrete, e.g. to the points of a grid.	
5.1.	Domain	The range in space, time or space-time that is designated for value assignment.	
5.2.	Coverage	The fraction of the domain where values have been actually assigned.	
5.3.	Raster	A two-dimensional array; every element is accessible by the pair of row and column index, is called a pixel, and contains a single or a tuple of values.	Image Raster image
5.3.1.	Pixel	An element of a raster.	
5.4.	Timeseries	A one-dimensional vector; every element is accessible by an index and contains a single or a tuple of values; the index represents points or periods in time that are sampled in equal (or near-equal) intervals.	

Context	Term (Acronym)	Description	Synonyms
6.	**Physical system**	A part of nature chosen for analysis. The part outside the system is its environment. Effects of the environment on the system are taken into account by an abstraction, and vice-versa. The cut between system and environment is a free choice, generally made to simplify analysis. An isolated system is one which has negligible interaction with its environment.	Geo-physical system
6.1.	Parameter	A measurable quantity suited to express the state of a physical system. In the given context synonymous for “geophysical parameter”.	Sate variable Geophysical parameter Environmental Variable
6.2.	Local parameter	Characterizes the state of a local and compact object in space (i.e. a local physical system). A Local parameter may be measureable with remote sensing instruments, while others may not be measurable with such instruments, but can still be inferred from direct measurements and additional knowledge about the sensed system (retrieval by an inversion or modelling approach).	Local state variable
6.3.	Model	A relation between parameters of a physical system, expressed in a mathematical or computable manner. E.g. as an equation system, Y = f(X), where X and Y being tuples of system parameters; or as a computer simulation.	Mathematical model Simulation
7.	**Retrieval**	The process of obtaining parameter values of a physical system from remote sensing measurements, although the parameters cannot be directly measured (observed) with the used remote sensing instruments. A retrieval is made on a mathematical-physical basis as follows: (1) It is known that at any time quantities X and quantities Y have definite values; (2) it is known that there exists a functional relation between system parameters X and instrument observables Y; (3) there is partial knowledge about the functional relation, e.g. in the form of a forward model Y = f(X) or an inversion X = g(Y); and (4) under certain conditions the knowledge of the functional relation is sufficient to determine X from the measurements of Y.	Parameter retrieval
7.1.	Retrieval algorithm	The algorithm (or approach, or method) used to retrieve a parameter value from remote sensing measurements.	Retrieval approach or method
7.1.1.	Surface Soil Moisture (SSM)	A local geophysical parameter expressing the content of water in the top soil surface layer (of < 2 cm thickness), in physical units [%] or [m^3/m^3]. Typical SSM data products from global satellite observations have a spatial resolution of 25x25 km. Regional products downscaled with use of SAR data are available in a resolution of 1x1 km.	

Context	Term (Acronym)	Description	Synonyms
7.1.1.1.	TUW change detection algorithm	The algorithm used to retrieve the SSM parameter from calibrated measurements of the radar backscattering coefficient σ^0 made by an (active) scatterometer microwave instrument. Uses climate data timeseries of seasonally varying dry and wet reference values. Developed at TUW; implemented by the WARP processing module.	
7.1.1.2.	Land Parameter Retrieval Model (LPRM)	An iterative forward modelling approach used to retrieve the SSM parameter from calibrated measurements of the brightness temperature T_B made by a (passive) radiometer microwave instrument. Uses a global database of physical soil properties. Developed by NASA and VUA.	
7.1.2.	Soil Water Index (SWI)	A measure of the profile Soil Moisture content of soil surface layers within 2-100 cm obtained by using an infiltration model (filtering the SSM timeseries with an exponential function).	
8.	**Climate**	Climate is a measure of the average pattern of variation of meteorological variables (e.g. temperature, humidity, atmospheric pressure, wind, precipitation, atmospheric particle count, etc.) in a given region over long periods of time. A region's climate is generated by the climate system, which has five components: atmosphere, hydrosphere, cryosphere, land surface, and biosphere.	
8.1.	Essential Climate Variable (ECV)	Essential Climate Variables are geophysical parameters that are required to support the work of the United Nations Framework Convention on Climate Change (UNFCCC), and that are technically and economically feasible for systematic observation. An ECV Data Product combines observations from multiple remote-sensing instruments into a space-time grid; is complete and consistent, with a global and continuing coverage; and is intended for use in climate modeling. Synonyms are Super-collated (L3S) Data and TCDR.	L3S Thematic Climate Data Record (TCDR)
8.2.	Climatology	Average of a geophysical parameter for a given region; obtained from long-term (many year) observation.	
8.3.	Seasonality	Seasonal variation of a parameter's climatology, e.g. variation over the months of a year.	

5.6.3 External Schema: Data Processing Levels

Table 30 provides a full specification of the data processing levels, sublevels and product categories in terms of the data and data product types that emerge at each level. The organization of data and processing in terms of “levels” appeared first in [NASA-EOS86] and was then adopted for other systems, as well as for the CCI in [DSWG].

Note that processing levels are not strictly consecutive in the sense that input and output data of a specific processor do not necessarily have to be from immediately adjacent levels. As an example, L4 may be derived directly from L1 or L2 data. The most relevant categories used by the ECV Production System (L1b, L2b, L3S) are marked with darker colors in Table 30.

Table 30: Description of Data Processing Levels

<table>
<tr><th>Level</th><th colspan="3">Description / Subtypes</th><th></th></tr>
<tr><td rowspan="2">L0</td><td colspan="3">Raw Instrument Data</td><td rowspan="6">MEASUREMENTS</td></tr>
<tr><td colspan="3">Raw Instrument Data are raster data from a remote sensing instrument (syn. sensor) flying onboard an earth observation satellite. They are in the geometry of the instrument’s swath grid, in the chronological order of the acquisition time, and in the full resolution of the instrument.
Communication artifacts (synchronization frames, headers, duplicate data, etc.) were removed.</td></tr>
<tr><td rowspan="4">L1</td><td colspan="3">Calibrated Instrument Measurements</td></tr>
<tr><td colspan="3">Calibrated Instrument Measurements are L0 data with: explicit time reference, radiometric and geometric calibration coefficients attached (L1a) or applied (L1b), and annotations of geo-referencing parameters (e.g. satellite ephemeris) and possibly other ancillary information.</td></tr>
<tr><td>L1a</td><td>L1b</td><td>L1c</td></tr>
<tr><td>Calibration attached, but not applied.
In full resolution.</td><td>Calibration applied, i.e. the measurement values are given in the physical unit of the instrument.
Not necessarily in full resolution.
Examples: Microwave scatterometer data, measuring radar backscattering coefficient σ^0 in [dB]; or microwave radiometer data, measuring brightness temperature T_B in [K]</td><td>L1b further processed, e.g. by applying corrections, or by mapping onto a global grid.
Example: SMOS MIRAS microwave radiometer, measuring brightness temperature T_B in [K]</td></tr>
</table>

<table>
<tr><td rowspan="4">L2</td><td colspan="3">Parameter Distribution Retrievals from Single Overpass</td><td rowspan="8">RETRIEVALS</td></tr>
<tr><td colspan="3">Distribution of a local parameter (local state variable) of a (geo-)physical system that cannot be directly measured (as in L1) but can still be retrieved from L1 data when using additional knowledge about the system; keeps the localization (in space and time) of the sensed objects of L1 and retrieves from a single overpass only.</td></tr>
<tr><td>L2a</td><td>L2b</td><td>L2c</td></tr>
<tr><td>From a single beam.
In same resolution and swath grid as the L1 source data.
Not interpolated.</td><td>Summed (averaged) over beams, and/or mapped onto a global grid.
Example: ASCAT SSM in [%], averaging fore, mid and aft beams and representing the Soil Moisture content within a thin soil surface layer (< 2 cm) at the time of overpass of the satellite</td><td>L2a or L2b further processed, e.g. by adding confidence flags after checking for cross errors, consistency and timeliness.
Named L2P in [DSWG].</td></tr>
<tr><td rowspan="4">L3</td><td colspan="3">Parameter Distribution Retrievals</td></tr>
<tr><td colspan="3">Distribution of a local parameter (local state variable) of a (geo-)physical system as in L2 but with localization and/or single overpass requirements relaxed to aim at any one or all of the following: (a) extending the sensed object in space (e.g. to reach lower soil surface layers), (b) increasing the observation time (e.g. to get daily, monthly, etc. means up to the parameter's climatology), or (c) complete and consistent coverage of the sensed space-time domain (e.g. combining multiple measurements from many orbits or satellites, interpolating missing points).
Typically mapped onto global and uniform space or space-time grids.</td></tr>
<tr><td>L3U (Uncollated)</td><td>L3C (Collated)</td><td>L3S (Super-collated)</td></tr>
<tr><td>From a single overpass (i.e. not combining observations from revisiting orbits).</td><td>Combining retrievals from revisiting orbits (of a single satellite).
Example: SWI</td><td>Combining retrievals from multiple orbits (of several satellites).
Example: SM ECV[9,10]</td></tr>
<tr><td rowspan="2">L4</td><td colspan="3">State and Change of State Analyses and Simulations</td><td rowspan="2">ANALYSES AND SIMULATIONS</td></tr>
<tr><td colspan="3">Reports on the state of (geo-)physical systems at certain times, derived from analyses of L3,2,1 data, typically about specific regional or global systems of interest, i.e. systems whose boundaries do not follow the (spatial-temporal) regularity patterns of the L2,3 distributions.
Reports about change of state of such systems, where change of state may be either conceived as events (instantaneous, short time) or processes (durative, long-time). Detection of change of state typically involves statistical analysis methods (e.g. trend analysis) or process model calculations (simulations) using L3,2,1 data as input.
Results of analyses or simulations may have a quality in the range from evidentiary to indicative.
Examples: Drought conditions, floods, precipitation feedback[11], crop yield, weather and climate models, global climate modelling[12,13,14], trend analysis[15]</td></tr>
</table>

[9] http://www.esa-soilmoisture-cci.org/node/127
[10] http://www.esa-soilmoisture-cci.org/node/1
[11] http://www.esa-soilmoisture-cci.org/node/148
[12] http://www.esa-soilmoisture-cci.org/node/128

We note that mapping to a global grid is often used as defining criteria for L3 data, but considering that global grids are already used for L1 and L2 data (cf. section 5.6.1.3 above) we do not adopt this definition herein. It has no physical relevance and its technical nature is doubtful: The original goal was that spatial grids should boost merging, comparison and interchange of Level 3 data, however, considering that there is no ideal grid for all purposes, and that many different grid systems have been developed (with pros and cons each), are in use nowadays, and are still searched for, it is clear that this goal cannot be reached.

[DSWG] also introduces a fifth level called 'Indicator'. The description in [DSWG] is not clear as it does not specify *what* is being indicated. Furthermore, we see indication as a degree of the definiteness of an analysis or simulation result (e.g. the result may be evidentiary or indicative), and as such the measure does not introduce a new level – like measurement uncertainty and error characterization do not introduce new levels but are part of L0-L3.

A possible fifth level may be constituted by human actions, or recommendations or guidance to such actions in response to the L4 analyses and simulations. We mention it as a proposal for L5 which could be added to the Table 30 above.

5.6.4 Conceptual Schema of Soil Moisture Data Items

This subsection specifies the Information Model for the Soil Moisture Data Items in the form of the conceptual ER schema shown in Figure 18. The schema covers all types of Soil Moisture data, from the L2 Data and Ancillary Data which are input data to the System, to intermediate data that are created during the processing, to the final ECV Data Products and associated Validation Data. Besides the mentioned entity types from the business domain, the schema also covers additional descriptive and administrative entity types (metadata) that emerge as a corollary of the management of the former entity types.

The diagram shown in Figure 18 has four distinctive parts: handles to the data (upper right), the file representation (upper left), a universal schema of the data (lower left), and the different subclasses of DataProduct (lower right). The color code used in Figure 18 is described in Table 27. Cardinalities are shown for emphasis only, meaning that an absent cardinality does not imply any specific cardinality of one or many.

Every Soil Moisture data set (or data product) is registered as a subclass of the DataProduct entity type (shown at upper right), therefore, entities of this type act as handles to the data sets. The hierarchy of DataProduct subclasses (i.e. subtypes) as known at the time of writing is shown on the lower right.

---

[13] http://www.esa-soilmoisture-cci.org/node/164
[14] http://www.esa-soilmoisture-cci.org/node/171
[15] http://www.esa-soilmoisture-cci.org/node/132

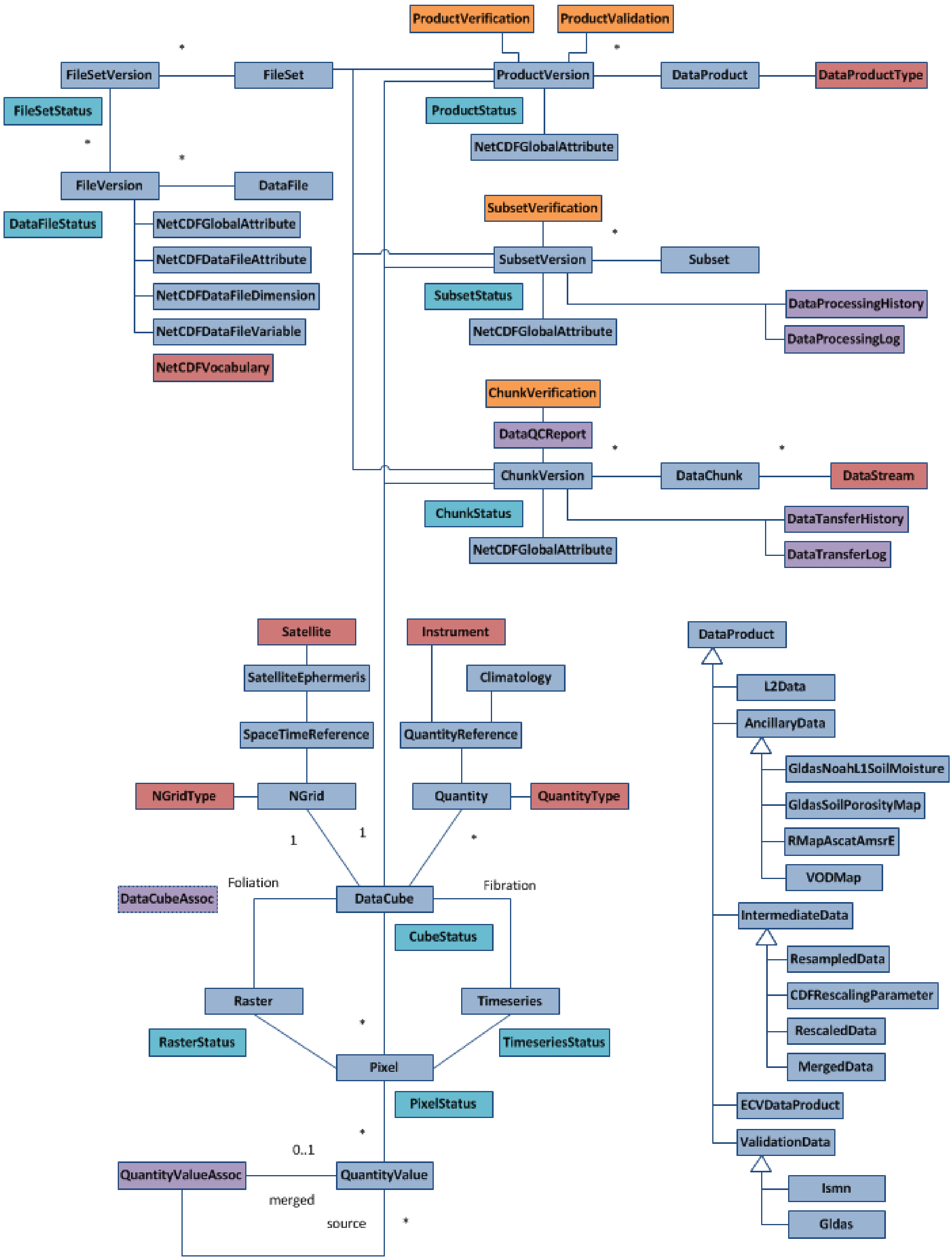


Figure 18: Conceptual Schema of Soil Moisture Data Items

Besides the DataProduct handle, there are two further handles to the data: DataChunk and Subset (both shown at upper right).

DataChunk is used in the context of automated data acquisition (transfers) by modelling individually transferred data as “chunks” and providing means for recording status and administrative information per DataChunk. An automated Data QC Report and its verification by an Operator or Analyst are examples for such administrative information.

Subset is a generic entity type that can act as a handle to any subset of Soil Moisture data in the System. Typical subsets required for automated data processing are rasters and timeseries. E.g. a stack of rasters per grid node set and data day is input to the initial spatial-temporal resampling, whereas all further data processing operates on a single timeseries per grid node. The Subset entity type allows organizing data processing and recording of process information like status, status history and the processing logs. Administrative information contained in DataProcessingHistory are the algorithm and software version used, and DataProcessingLog collects the log messages written by the data processor. Besides organizing the data processing, Subset also acts as a means for organizing the data verification activities carried out by Analysts and any use of the Soil Moisture data in an analysis by internal or external Data Users.

The universal representation of the Soil Moisture data is shown in the lower left of Figure 18: At the center is the DataCube entity type with references to a space-time grid (NGrid) and to a set of quantities (Quantity), both together defining the DataCube’s dimension attributes. Subsets of the DataCube at a fixed time are (2-dimensional) rasters, while subsets at a fixed point in space are timeseries, where each Pixel carries the set of quantity values (QuantityValue). QuantityValues constitute the DataCube’s measure attributes. The concept is generic so that it can apply to any type of grid-based value distributions in space and time, and it allows to link such distributions to the earth surface (geo-referencing), to the time (time-referencing), and to quantity references (i.e. physical units and reference scales of an Instrument or a Climatology).

Considering that merging (blending, fusing) Soil Moisture data is one of the main data operations carried out by the System, the conceptual schema allows for data associations at the level of quantity values (QuantityValueAssoc), which can be aggregated to the level of a DataCube entity (shown as derived entity type DataCubeAssoc in Figure 18).

For more information about the different classes of Soil Moisture data shown in the lower right of Figure 18 please cf. to [SRD] section 5 and to the references contained therein, as well as [IODD] and [PSD].

5.6.5 *Conceptual Schema of System Monitoring & Control Information Items*

This subsection specifies the Information Model for the Information Items of System Monitoring & Control in the form of the conceptual ER schema shown in Figure 19.

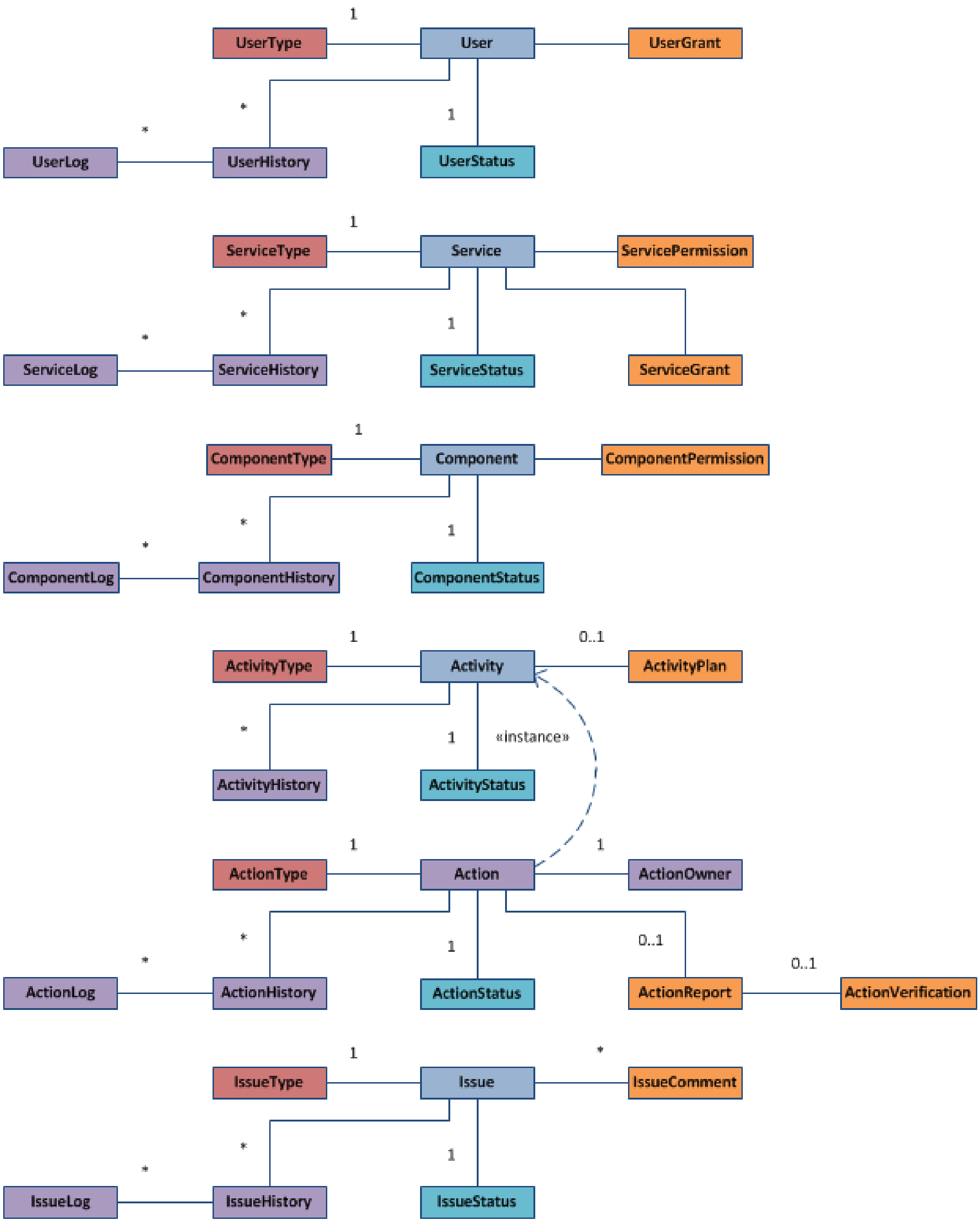


Figure 19: Conceptual Schema of System Monitoring & Control Information Items

The ER model shown in Figure 19 has two distinct parts: The upper part encompassing entity types User, Service and Component models System Elements for the purpose of monitoring and controlling them. These entities have a permanent nature, i.e. they are expected to exist for months, years or even beyond the System's life time. The lower part encompassing entity types Activity, Action and Issue models elements of a transient nature, i.e. these entities are expected to seize to exist after a certain time. The color code used in Figure 19 is described in Table 27. Cardinalities are shown only where relevant.

The User entity type includes two relevant aspects: On the one hand rights are granted to users to use services (UserGrant), and on the other hand user actions are logged for visitor monitoring purposes as required by SR-0362, SR-0363 and SR-0364.

Any System Element that is usable can be a candidate for registration as a Service entity for the purpose of monitoring and controlling its use. Generally, a Service entity is a front end, interface or entry point to the specific services offered by an underlying element registered as a Component. The Component entity type subsumes hardware components (e.g. server, storage, network devices) as well as software components (e.g. processing modules).

The Activity entity type allows control and monitoring of system behavior in a customizable, updatable and traceable way. It subsumes both human and machine activities, where the latter are invocations of those services that have been registered as Service entities. An Activity may be carried out at any time, though the concept allows for planning Activities in advance by an ActivityPlan. An Activity is instantiated when it is carried out. An instantiated Activity is modeled as an Action. An Action may be instantaneous but can also last for some time until it completes (e.g. a process). The Activity concept (and likewise the Action concept) follows the composite design pattern with the Workflow entity type (shown in Figure 20 but hidden in Figure 19) understood to be an Activity with a substructure that organizes other Activities into sub-Activities.

The Issue entity type allows recording and managing information items like problem reports, bug reports or change requests. It further allows for association of information to activities and actions (not shown in the diagram), and for the distinction between current and past information, with current information directly accessible while past information is kept retrievable.

Each of the main entity types described above has associated status (*Status), status history (*History) and log (*Log) entity types. In this way process information like status monitoring data, Operator interaction logs, log files from data processing, or user logs and user statistics can be captured, stored and analyzed.

Figure 20 shows details of the relationships between the Activity, Action and Component entity types for the example of a software component (SWApp). The diagram has three main parts: On the upper left the Activity hierarchy with the three subtypes HumanActivity, Workflow and SWAppActivity, on the lower left SWApp as a subtype of Component together with the software configuration model, and on the right side the SWAppProcess (subtype of

Action) – wrapping the actual computer process (ComputerProcess) of the executing software application (SWAppInstance).

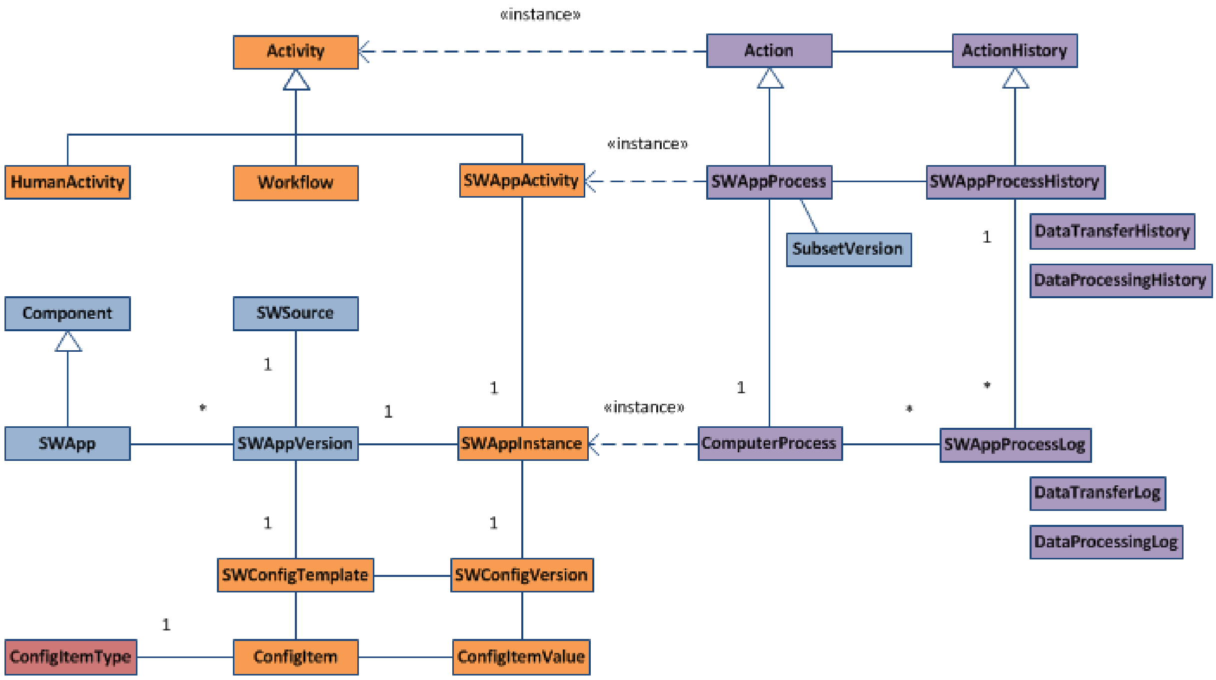


Figure 20: Conceptual Schema for Activity and Process Management

The schema allows for versioning of the software, maintaining different configurations of software applications (SWAppInstance) and versioning of the configurations (SWConfigVersion).

A SWAppActivity is thus an Activity that is prepared to execute a specific version of a software application (SWAppVersion) with a specific version of a configuration (SWConfigVersion). The executing SWAppActivity is represented as an Action of type SWAppProcess. A SWAppProcess records execution times and the command line arguments to SWAppProcessHistory. Log messages from the ComputerProcess are recorded in the SWAppProcessLog.

Typical software applications which fall under this regime are those for automated data transfer and data processing as indicated by the corresponding *History and *Log entity types.

The Information Model as specified by the schemata of Figure 18, Figure 19 and Figure 20 allows for full product tracing as required in SR-0931 and SR-0932. It supports recording of all processing steps and versions used for a particular output Data Product, where the following Information Items are versioned: a) processing algorithms and software applications, b) processing parameters (i.e. configuration), and c) data.

5.7 Infrastructure View

This section specifies the Hardware and Software Infrastructure required to run the Application and Platform Software, as specified in the Software View in section 5.3. Three basic options to establish the System's Infrastructure Layer are discussed:

A. Dedicated hardware in conjunction with Linux-based Host OS and virtualization software (hypervisor).

B. Deployment on a Cloud Computing Infrastructure.

C. A heterogeneous infrastructure combining both options A and B.

Option A can be implemented with a hardware configuration as specified in [SRD] section 6.5.3. This configuration permits the complete deployment of Platform Software and Application Software on a dedicated, hardware-based and isolated computing infrastructure. The performance observed for the Soil Moisture Prototype ECVPS (SM PECVPS) in the ESA CCI Phase 1 Soil Moisture project confirms that the specified hardware configuration will meet the minimum performance requirement [SRD] SR-0370 (cf. also the key characteristics in [SRD] section 4.3). However, this performance requirement may soon be overtaken by events with data coming in from new satellite instruments. In order to speed up ECV processing additional computing power for the OPS LAN in conjunction with software and methods for parallel computing and scaling would thus be required, cf. the discussion in section 4.3, TM-014.

Option B uses a Cloud Infrastructure Service (IaaS) offered by a Commercial (Public) Cloud Provider, or a similar Private Cloud Infrastructure maintained by the Hosting Organization. This is possible because the System's software architecture, as specified in the Software View of section 5.3, is a) based on virtualization technology and b) designed to be fit for a Cloud Computing Architecture, as described by the UCSB-IBM Cloud Computing Classification Model [Ahson11]. The three subsets of Cloud services map to the System's software as follows:

- The Soil Moisture ECV Application Software, as specified in section 5.3.1, could be provided as Cloud Software as a Service (SaaS).
- The Soil Moisture ECV Platform of Software Environments, as specified in sections 5.3.2 and 5.3.3, could be provided as Cloud Platform as a Service (PaaS).
- The former two elements are to be deployed on top of a Cloud Infrastructure as a Service (IaaS) offered by a Public Cloud Provider or a Private Cloud Infrastructure.

In order to accommodate the System's base workload computational, storage and communication resources acquired from a Cloud IaaS Provider will need to correspond approximately to the hardware configuration, as specified in [SRD] section 6.5.3. The Cloud setup has the important advantage that resources can scale up on demand, but they can only be efficiently used for processing if the ECV Processor is designed to use the Cloud's elastic compute feature, i.e. processing can be parallelized (cf. the discussion in section 4.3, TM-014).

Option B is currently being prototyped and evaluated under ESA's EP4SM project. Figure 21 copied from [EP4SM-SAD] shows the architecture of a suitable Cloud Infrastructure as required for deployment of the Soil Moisture ECVPS.

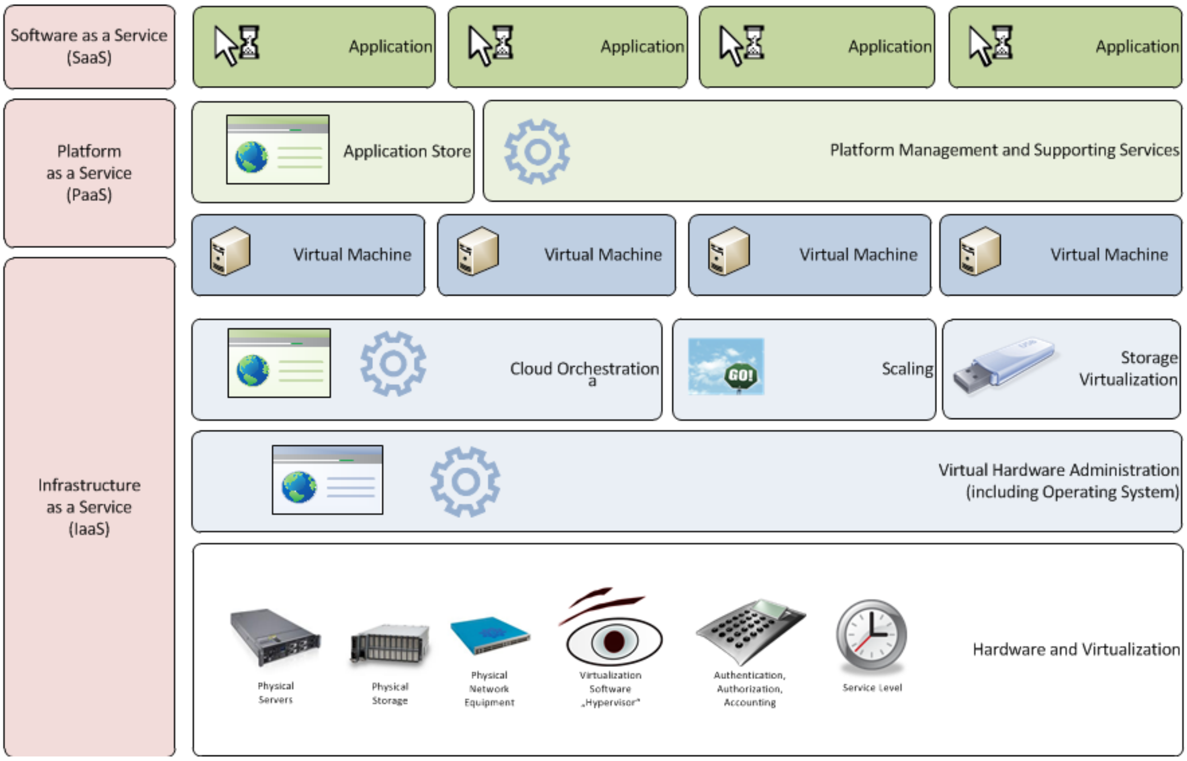


Figure 21: Cloud infrastructure for deployment of the Soil Moisture ECVPS

The Cloud architecture shown in Figure 21 contains the three layers SaaS, PaaS and IaaS as follows:

The SaaS layer provides the System's software applications to all application users. For example external Data Users will perform their Use Cases through a web portal. Included are also all internally used applications, e.g. for data acquisition, processing and analysis. They can be provided as SaaS if they are designed to be called via a web service API.

The PaaS layer provides the capabilities, which are required to develop, test and operate the System's software. It is controlled and monitored by software for Application Store and Platform Management and Supporting Services. The PaaS layer provides VM templates for runtime, development and testing environments with the compiler, development and profiling tools preconfigured for developers. The primary users of this layer are the Software

Developers and Processing Engineers who implement system changes. Instantiated development, testing and operations environments are isolated from each other and mapped to application groups in individual Virtual Machines.

The IaaS layer provides the System's computational, storage and networking resources and includes host Operating Systems, the virtualization software stack (hypervisors) and Cloud Orchestration Software. This layer can be implemented either via an IaaS Cloud Provider or via a private Cloud Infrastructure of the Hosting Organization. It is deployed on top of a Hardware Infrastructure consisting of physical servers, storage and networking equipment.

Option C, deployment on a heterogeneous Cloud, combines the local Private Cloud computing infrastructure with an externally provided commercial Cloud Infrastructure. In this scenario the base system can be operated on the local infrastructure, while peak demands can be addressed by temporarily increased external Cloud resources. As already discussed for Option B, the scaling potential of a Cloud infrastructure can only be efficiently used if the ECV Processor is designed for process parallelization. Temporarily increased Cloud resources may further be useful to host collaborative projects or activities within the Science Community or with external Software Developers. In such cases additional DVL and TST ENVs could be deployed in the extended Cloud infrastructure.

5.8 Life Cycle Sustainment Activities

System sustainment has financial, organizational and technical aspects and concerns sustainment of system operations, system maintenance and system evolution, potentially including system replacement. The discussion in this section is organized along the latter three categories. Financial sustainment essentially means to assure funding for the organizational and technical aspects of system sustainment and is not discussed further here. The organizational and technical aspects themselves are discussed in the following subsections for each of the three categories.

5.8.1 Sustainment of System Operations

Sustainment of operations entails all activities and measures taken to keep the system in operating order and to minimize or avoid system outages.

Organizational Sustainment: The Operator and Analysis Units discussed in section 5.2 operate the ECV Production System. The required staff needs to be made available, operational procedures need to be developed, documented and maintained, and know how needs to be retained by documenting working knowledge, by systematically transferring knowledge to new staff and by training the existing staff.

Technical Sustainment: All system components, which are required for operations need to be kept in working order. This extends in particular to the hardware, software and office infrastructure. Broken components need to be replaced, and service level agreements need to be in place for essential parts of the infrastructure (e.g. network connectivity, hardware, recommended infrastructure software updates, e.g. security updates, etc.). This aspect becomes even more vital if large parts of the infrastructure are outsourced, e.g. cloud infrastructure provisioning services. Then the service contracts themselves need to be sustained.

Application software issues (problem reports from users, issues detected by operators, bugs observed with product creation, etc.) need to be resolved by the developers of the software unit. This is one of the cases covered in the software change process described in section 5.4.3.

5.8.2 Sustainment of System Maintenance

System maintenance activities support operations by taking measures to reduce the risk that operational problems may occur in the future or by mitigating the impact of potential problems. Sustainment of system maintenance means to assure that the system remains well maintained over time.

Organizational Sustainment: The Software Unit discussed in section 5.2 maintains the system's software. The Science Unit has a support role in particular when maintenance activities are required for processing algorithms. The System Engineers in collaboration with the Operations Manager have the responsibility for infrastructure maintenance activities.

Management needs to assure that sufficient and continuous resources of the organization are allocated to maintenance activities, and that system maintenance does not fall through the cracks because of persistently higher operational priorities.

Technical Sustainment: Beyond addressing reported software issues for operations systematic software maintenance includes adapting the application software for infrastructure upgrades (adaptive maintenance), systematically reviewing and improving the existing software to eliminate potential problems even before they hit (preventive maintenance) and improving the maintainability of the system itself by updating documentation, systematically implement standards were deviations are found, eliminating existing dependencies where they are not actually required or improving the readability of existing code (perfective maintenance).

Maintaining the system infrastructure is not fundamentally different to sustaining its operation, but the future oriented aspect is more emphasized. Thus, typical infrastructure maintenance activities are planning upgrades and replacements of existing components ahead of time, coordination of these plans with all affected stakeholders (users, operators, developers, scientists, management), and evaluation of the need for new service level agreements as well as actually putting them in place.

5.8.3 Sustainment of System Evolution

While sustainment of operations and maintenance are aimed at assuring that an existing system can be operated well now and in the future, sustaining system evolution means to additionally assure that the system can be modified and further developed as its requirements may change.

This entails providing resources for system evolution, keeping the system changeable (i.e. apply current changes in a way not to obstruct future changes) and taking long term life cycle planning into account (i.e. how long shall the system be operated and how far shall it evolve before it is replaced; what is required at which time to get a timely replacement).

Organizational Sustainment: The Science and Software Units of the organization have the roles to evolve the underlying scientific methods, the algorithms used for data processing, the data products, the application software and the IT infrastructure. The Configuration and Product Control Boards (CCB and PCB) have the roles to approve proposed changes within the respective change processes.

Management not only has to allocate the resources for system evolution but also needs to assure long term planning and that appropriate system change processes and system documentation are in place and actually used. In particular the Quality Manager has the role to develop documentation standards, to review and adapt change processes and to assure that they are implemented.

Technical Sustainment: Sections 5.4.3 and 5.4.4 describe initial change processes for the system software and for the ECV products. Configuration and Product Control Boards (CCB and PCB) approve individual change proposals, which are then implemented and systematically validated before the changes are deployed in operations.

Maintainability and future changeability need to be taken into account when individual system components (e.g. software applications) or the system architecture are changed. Changes need to be systematically documented to retain the ability to track the system's change history, and existing documentation needs to be updated when the underlying information changes. This requires a systematic change process, which itself needs to be maintained and managed. These latter activities are an essential responsibility of the Quality Manager.

Long term life cycle planning starts when the system first goes into operation or even earlier when it is still developed. A suitable instrument for this purpose is a sustainment plan, which is an evolving document to be maintained and periodically updated by the Quality Manager in collaboration with management, scientific and technical staff to add information as it becomes available.

The initial focus will be on specifying a minimum operational system lifetime and planning all relevant measures to ensure sustained operations, maintenance and evolution capabilities for the specified period. Further into the system lifecycle the need to plan ahead for system decommissioning and replacement will become more relevant. This type of planning needs in any case to be started soon enough for a replacement system to be developed in time before the existing system will be decommissioned.

6 Traceability Matrix

Table 31 shows the traceability matrix, which allocates the System Requirements as specified in [SRD] to the System Functions as defined by section 5.1, Figure 1.

Table 31: Allocation of System Requirements to System Functions

			F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
SR-0010-FUN-GEN	Merged Active-Passive ECV Data Prod.	MUST																	c					c				c				c		c	M													
SR-0020-FUN-GEN	Merged Active-only ECV Data Prod.	COULD																	c					c				c				c		c	M													
SR-0030-FUN-GEN	Merged Passive-only ECV Data Prod.	COULD																	c					c				c				c		c	M													
SR-0040-FUN-ACQU	L2 Data Acquisition	MUST																	c					c			M	c																				
SR-0050-FUN-ACQU	Ancillary Data Acquisition	MUST																	c					c			M	c																				
SR-0070-FUN-PRE	Input Data QC - Data QC Report	MUST																	c					c			c	c					M															
SR-0080-FUN-PRE	Input Data QC - Missing File Check	MUST																															M															
SR-0090-FUN-PRE	Input Data QC - Outliers	MUST																															M															
SR-0100-FUN-PRE	Input Data QC - Unprocessed Grid Points	MUST																															M															
SR-0110-FUN-PRE	Input Data QC - Processed File Listing	MUST																															M															
SR-0560-FUN-PROC	Processor Module/s	MUST									c	c	c		X	M	X	X	M	c	c																											
SR-0120-FUN-PROC	Operator Initiated Reprocessing	MUST																	c					c				c							M													
SR-0130-FUN-PROC	Automatic Interval Processing	COULD																	c					c			X	c					X		M													
SR-0131-FUN-PROC	On-Demand Processing	WONT		M															c					c				c		c					X													
SR-0140-FUN-PROC	Interm. Data Storage - Resampling	MUST																										M							c													
SR-0150-FUN-PROC	Interm. Data Reuse - Resampling	MUST																										c							M													
SR-0160-FUN-PROC	Interm. Data Storage - CDF Rescaling Par.	MUST																										M							c													
SR-0170-FUN-PROC	ECV Data Reuse	MUST																										c							M													
SR-0175-FUN-PROC	Intermediate Data Storage - Rescaling	MUST																										M							c													
SR-0176-FUN-PROC	Intermediate Data Storage - Merging	MUST																										M							c													
SR-0180-FUN-PROC	Data Association Storage	MUST																					c				c	M			c	c	X	c	X													

			F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
SR-0181-FUN-PROC	Processing History Storage	MUST											c										c					M					X		X		c			c								
SR-0190-FUN-POST	ECV Data Prod./s Verification	MUST		X							X																	c									M				c							
SR-0200-FUN-POST	Trend Analyses	MUST									X																									X		M										
SR-0210-FUN-ACCS	ECV Data Prod./s Dissemination	MUST		M																								c		X																		
SR-0220-FUN-ACCS	ECV Data Prod./s Subsetting	SHOULD		M																								c		X																		
SR-0230-FUN-ACCS	Data Format Conversion	WONT																												M																		
SR-0240-FUN-ACCS	Data Units Conversion	MUST																												M																		
SR-0250-FUN-ARCH	On-line Data Arch. after Input Data QC	MUST																										c	M				c															
SR-0260-FUN-ARCH	On-line Data Archiving after Processing	MUST																										c	M						c													
SR-0261-FUN-ARCH	On-line Data Archiving - Restoring	MUST																						c				X	M																			
SR-0262-FUN-ARCH	Purging from On-line Data Archive	MUST																											M																			
SR-0263-FUN-ARCH	Configuration of Data Expiration Time	MUST																						c					M																			
SR-0270-FUN-ARCH	Off-line Data Archiving - Full Backup	MUST																						c					M																			
SR-0280-FUN-ARCH	Off-line Data Archiving - Incr. Backup	MUST																						c					M																			
SR-0290-FUN-ARCH	Off-line Data Archiving - Restoring	MUST																						c					M																			
SR-1210-FUN-MON	System Monitoring - Components	SHOULD		c					c										M			M	c	M			c	X	c	c			c		c							c	c					c
SR-1220-FUN-MON	System Monitoring - Activities	SHOULD	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	M	c	c	M	c	M	c	c	c	X	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
SR-1230-FUN-MON	Status Access Control	SHOULD																	M			M		M				M															M					
SR-1240-FUN-MON	Status Monitoring Database	COULD																	X			X		X				M															X					
SR-1250-FUN-MON	Status Planning	SHOULD																										X															M					
SR-1260-FUN-MON	Status History	SHOULD																										X															M					
SR-1270-FUN-MON	Status Notification	SHOULD																	M			M		M				X															M					
SR-1280-FUN-MON	System Control - Components	SHOULD		c					c										M			M	c	M			c	c	c	c			c		c													
SR-1290-FUN-MON	System Control - Activities	SHOULD		c					c										M			M	c	M			c	c	c	c			c		c								M					
SR-1300-FUN-MON	Access Control for Sys. Mon. & Ctrl.	SHOULD																	M			M		M				M															M					
SR-1310-FUN-MON	Status Monitoring Database	COULD																	X			X		X				M															X					
SR-0300-FUN-MON	Logging Program Execution	MUST																	M																													
SR-0310-FUN-MON	Logging Program Termination	MUST																	M																													
SR-0320-FUN-MON	Logging Automated Data Transfer	MUST																									M																					
SR-0330-FUN-MON	Logging Grid Point Processing	MUST																																	M													
SR-0340-FUN-MON	Logging Missing Data	MUST																																	M													

			F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
SR-0350-FUN-MON	Logging Processing Error	MUST																																	M													
SR-0360-FUN-MON	General Processing Log Information	MUST																																	M													
SR-0362-FUN-MON	Visitor Monitoring - Downloads	MUST		M																								c		X																		
SR-0363-FUN-MON	Visitor Monitoring - Online Viewing	MUST		M																								c																				
SR-0364-FUN-MON	Visitor Monitoring - Usage Statistics	MUST	c	M																								c																				
SR-0365-FUN-RCVR	Error Notification	MUST																	M								X			X			X		X													
SR-0366-FUN-RCVR	Processing Checkpoints	MUST																																	M													
SR-0367-FUN-RCVR	Processing Interruption	MUST																	c																M													
SR-0368-FUN-RCVR	Processing Recovery	MUST																	c																M													
SR-0370-PRF	Automatic Reprocessing Performance	MUST																																	M													
SR-0375-PRF	Automatic Data Acquisition Perf.	SHOULD																									M																					
SR-0376-PRF	Automatic Interval Processing Perf.	COULD																																	M													
SR-0380-INF-FRMT	Data Standards: Minimum Format Reqs.	MUST								M																									X													
SR-0390-INF-FRMT	Data Standards: Common Vocabulary	MUST								M																									X													
SR-0400-INF-FRMT	Data Standards: Gridded Data Coord.	MUST								M																									X													
SR-0410-INF-FRMT	Data Standards: Filenaming Convention	MUST								M																									X													
SR-0060-INF-DATA	System Initialization Data	MUST																									M																					
SR-0420-INF-DATA	L2 Data Interfaces - Manual	MUST																									M																					
SR-0421-INF-DATA	L2 Data Interfaces - Automatic	SHOULD																									M																					
SR-0425-INF-DATA	Ancillary Data Interfaces - Manual	MUST																									M							c														
SR-0430-INF-ORG	ECV Data Prod. User Feedback	MUST	M	X	X	X		X	c																																							c
SR-0429-INF-ORG	Coop. with Hosting Organization/s	MUST																				X	c	M	c	c															c		c		c	X		c
SR-0440-INF-ORG	Scientific Cooperation	MUST		c					c	c	M	c	X	c																	c	c		c		c		X	X	X	c		c		c	X		c
SR-0441-INF-ORG	Professional Cooperation	MUST		X	M	M		X				X			X	M	X	X	M	X	X	M	M	M	X	X		M	X												c		c		c	X		c
SR-0450-INF-ORG	Climate Research and Mod. Community	MUST		M						X			c																										c									
SR-0451-INF-ORG	Community Valid. of ECV Data Prod.	MUST																												X										M	c		c			c		X
SR-0460-OPL-USER	Data User Reg. - Selfregistration	MUST		M																										c																		
SR-0470-OPL-USER	Data User Reg. - Operator Notification	MUST		M																										c																		
SR-0480-OPL-USER	Data User Mgmt. - Setting Access Rights	MUST		M																										c																		
SR-0490-OPL-USER	Data User Mgmt. - Perm. Change Message	MUST		M																										c																		
SR-1050-RES-HMN	Human Resources - Operator	MUST		X			X		X					X	c		X	X	X		c	X	X	X		c	X	X	X	X			X		X						X	X	X	X	X	X	X	X

			F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
SR-1060-RES-HMN	Human Resources - Sci. Maintainer	MUST	c		X	X		X		X	X	X	X	X	c		c				c					c					X	X		X		X	X	X	X		X		c	X			X	X
SR-1070-RES-HMN	Human Resources - SW Maintainer	MUST												c	X	X	c	c		X	X				X	X															c		c	X			X	X
SR-1080-RES-FCLY	Office Space	MUST	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
SR-1090-RES-FCLY	Server Rooms	MUST																				X		X																								
SR-1100-RES-HW	Network Resources	MUST																				X		X																								
SR-1110-RES-HW	Server Resources	MUST																				X		X																								
SR-1120-RES-HW	Storage and Backup Resources	MUST																				X		X				c	c																			
SR-1130-RES	Office Workplace Resources	MUST	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
SR-1140-RES	Printing Resources	MUST	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
SR-0530-CON-DSGN	Data Source Integration	MUST									c			M	c												X	c	c			X	c		X	c												
SR-0535-CON-DSGN	L1 Inter-Calibr. Bias Quant. and Removal	COULD									M	X		M	c												X	c	c		X		c		X	c												
SR-0540-CON-DSGN	Processing Alg. - Dev. in Isolation	MUST										M				c	c																															
SR-0541-CON-DSGN	Processing Alg. - Dev. in IDL	MUST										M																																				
SR-0550-CON-DSGN	Modular Design	MUST													c				M																													
SR-0555-CON-DSGN	Layered System Architecture	MUST													c				M			M		M		M																						
SR-0561-CON-DSGN	Processor Modules Chaining	MUST													c											c							X		M													
SR-0562-CON-DSGN	Technical Platform	COULD																				X		M		M																						
SR-0570-CON-DSGN	Cross-ECV Synergies	SHOULD		X					X			X		X	M	X			X		M	M	M	M		M	X	X	X	X			X		X													
SR-0580-CON-DSGN	System Integration	MUST																				M	M	M	X	M																						
SR-0590-CON-DSGN	SEWG Standards	MUST		X					X			X		X	M	M	M		X		M	M	M	M		M	X	X	X	X			X		X													
SR-0600-CON-DSGN	Sci. Advisory Groups Recommendations	MUST		X					X	X	M	M		X	M							M	M	M		M	X	X	X	X			X		X	X	X	X		X								
SR-0610-CON-DSGN	ECSS Space Engineering SW Standards	MUST		X					X			M		X	M	M	M		X		M		M	M		M	X	X	X	X			X		X													
SR-0620-CON-DSGN	ECV Data Prod. Visualization Tool	MUST		M																								X	X																			
SR-0630-CON-DSGN	ECV Data Prod. Intercomparison Tool	MUST		M																								X	X																			
SR-0640-CON-IMPL	Open Data	MUST		M																																										X		
SR-0650-CON-IMPL	Open Source	MUST													M				M																											X		
SR-0660-CON-IMPL	ECV Independence	MUST		X					X			X		X	M	X			X		M	M	M	M		M	X	X	X	X			X		X													X
SR-0670-CON-IMPL	System Scope	MUST																									X			X			X		M													X
SR-0680-CON-IMPL	System Hosting	MUST																				M		M		M																						X
SR-0690-CON-IMPL	IDL Runtime Environment	MUST										X												M		M							X		X													
SR-0691-CON-IMPL	Processor Module/s Impl. in C++	MUST										X																					X		M													X

Requirement ID	Requirement Title	Priority	F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
			F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
Requirement ID	Requirement Title	Priority																																														
SR-0700-VRF-SYS	System Testing	MUST															X	c																										M				
SR-0710-VRF-SYS	System Test Report (STR)	MUST															M	c																										c				
SR-0720-VRF-DOC	URD Review	MUST	X				X																																					M				
SR-0730-VRF-DOC	Data Specification Documents Review	MUST					X			X																																		M				
SR-0740-VRF-DOC	Scientific Methods Review	MUST					X				X																																	M				
SR-0750-VRF-DOC	Processing Algorithms Publication	MUST					X				c	c	X																										c					M				
SR-0760-DOC-QC	Procedure for Open Reviews	MUST					X																																					M				
SR-0770-DOC-QC	Procedure for Reviews by Auth. Bodies	MUST					X																																					M				
SR-0780-DOC-QC	Software Quality Assurance Plan	MUST					X									c	c	c	c	c			c	c	c																			M				
SR-0785-DOC-QC	Coding Standard	MUST					X									c			c																									M				
SR-0790-DOC-USER	Data Specification Documentation	MUST					M			X																																						
SR-0800-DOC-USER	User Requirements Document (URD)	MUST	X				M																																									
SR-0810-DOC-USER	Prod. User Guide (PUG)	MUST					M			X																																						
SR-0820-DOC-USER	Processing Algorithms Documentation	MUST					M					X																																				
SR-0830-DOC-USER	Detailed Processing Model (DPM)	MUST					M					X																																				
SR-0840-DOC-SYS	System Specification Documentation	MUST					M								X																																	
SR-0845-DOC-SYS	In-line Documentation	MUST					M									X																																
SR-0860-DOC-TEST	Integration and Test Plan (ITP)	MUST					X										X	X	X																									M				
SR-0850-DOC-TEST	(Data) Prod. Validation Plan (PVP)	MUST					X																														X			X				M				
SR-0855-DOC-TEST	Accept. Test Proc. for Processor Modules	MUST					X										X	X	X																X									M				
SR-0862-DOC-OPS	Operations and Maintenance Plan	MUST		X			M				c								X	X		c	c	X	X										c		c				c	c		c	c			
SR-0864-DOC-OPS	Operations Procedures Document	MUST		X			M												X					X																								
SR-0868-DOC-EVOL	Future System Support Plan	MUST		X			X												X		X			X		X															X				X		X	M
SR-1150-SCY	Access Control for Computing Resources	MUST																	M			M		M																								
SR-1160-SCY	Demilitarized Zone (DMZ) Subnetwork	MUST																				M																										
SR-1170-SCY	Virtual Private Network (VPN)	MUST																				M		X																								
SR-1180-SCY	Secured Operations Subnetwork	MUST																				M																										
SR-1190-SCY	Secure Server Rooms	MUST																				M																										
SR-1200-SCY	Offsite Backup Storage Space	MUST																											M																			
SR-0869-SCY	Privacy for User Data	MUST		M																								X	X																			
SR-0870-PTY	Porting	MUST		X								X			X	X			M		M	M	M	M		M	X	X	X	X			X		X													

Requirement ID	Requirement Title	Priority	F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L.-t. Planning
SR-0878-QTY	System Availability	MUST		X					X										X	X		X	X	X	X																X	X	X	M				X
SR-0879-QTY	External Data Access Availability	MUST		M															X	X		X		X	X			X		X														X				X
SR-0880-QTY	File Store Availability	MUST																				X		X	X			M	X															X				X
SR-0881-QTY	On-line Data Archive Availability	MUST																				X		X	X			X	M															X				X
SR-0882-QTY	Processor Module/s Availability	MUST										X							X	X			X										X		M									X				X
SR-0885-RLY	System MTBF	MUST		X					X										X	X		X	X	X	X																X	X	X	M				X
SR-0886-RLY	External Data Access MTBF	MUST		M															X	X		X		X	X			X		X														X				X
SR-0887-RLY	File Store MTBF	MUST																				X		X	X			M	X															X				X
SR-0888-RLY	On-line Data Archive MTBF	MUST																				X		X	X			X	M															X				X
SR-0889-RLY	Processor Module/s MTBF	MUST										X							X	X			X										X		M									X				X
SR-0890-RLY	Appl. SW Development Environment	MUST														M																																
SR-0900-RLY	Appl. SW Testing Environment	MUST															M																															
SR-0905-MTY	Appl. SW Operations Instance	MUST																	M																													
SR-0906-MTY	Appl. SW Testing Instance	MUST															M																															
SR-0907-MTY	Appl. SW Development Instance	MUST														M																																
SR-0909-ORG-CM	Configuration Control Board (CCB)	MUST		X					X							X	X	X	M	X			X	X	X																X							
SR-0910-ORG-CM	Open Collaboration Management	MUST		X			X		M							X	X		X				X	X																								
SR-0920-ORG-CM	Open Collaboration - SW Sources	MUST							M														X																									
SR-0930-ORG-CM	Open Collaboration - Config. Items	MUST							M														X																									
SR-0931-ORG-CM	Data Versioning	MUST																					c				X	M	c		X	X	X	X	X													
SR-0500-ORG-CM	Data Vers. - Prod. Versions	MUST								X																		M																				
SR-0510-ORG-CM	Data Vers. - File Versions	MUST								X																		M																				
SR-0520-ORG-CM	Data Vers. - Prod. Vers. in File Names	MUST								X																		M																				
SR-0932-ORG-CM	Software Versioning	MUST										X				X			X	X			M																									
SR-0935-ORG-OPS	Operator Organization	MUST		X					X										X	c		X	X	X	X																	X	X					
SR-0936-ORG-OPS	Scientific Maintenance Organization	MUST									X	X	X	X																						X	X			X								
SR-0937-ORG-OPS	Software Maintenance Organization	MUST																		X	X																											
SR-0940-ORG-OPS	Data User Reg. - Verif. by Operator	MUST		M																																												
SR-0942-ORG-OPS	Man. Data Acquis. - Transfer Frequency	MUST																									M																					
SR-0945-ORG-OPS	Offsite Backup - Full Backup	MUST																											M																			
SR-0946-ORG-OPS	Offsite Backup - Regular Transfer	MUST																											M																			

			F1 User Support							F2 Product & Alg. Dev.					F3 Appl. SW							F4 Platform & Infra.					F5 Data Services				F6 Data Processing					F7 Data Analysis					F8 Support Funcs.							
Requirement ID	Requirement Title	Priority	F1.1 User Survey	F1.2 Data User Supp.	F1.3 Marketing	F1.4 (Sales)	F1.5 Documentation	F1.6 Training	F1.7 Collaboration	F2.1 Prod. Spec.	F2.2 Sci. Meth.	F2.3 Alg. Dev.	F2.4 Alg. Publ.	F2.5 Mission Integr.	F3.1 SW Spec.	F3.2 SW Dev.	F3.3 SW Testing	F3.4 SW Integr.	F3.5 Appl. SW Ops.	F3.6 SW Maint.	F3.7 SW Evol.	F4.1 Sys. Admin.	F4.2 Conf. Ctrl.	F4.3 Platform Ops.	F4.4 Platform Maint.	F4.5 Platform Evol.	F5.1 Data Aquis.	F5.2 Data Mgmt.	F5.3 Data Archiving	F5.4 Product Diss.	F6.1 (Calibr.)	F6.2 Parametrization	F6.3 Data QC	F6.4 Ancil. Data Gen.	F6.5 Product Gen.	F7.1 Data Eval.	F7.2 Prod. Verif.	F7.3 Trend Anal.	F7.4 Sci. Publ.	F7.5 Prod. Valid.	F8.1 Management	F8.2 Office Admin.	F8.3 Regul. Planning	F8.4 QA	F8.5 Financial Ctrl.	F8.6 Legal	F8.7 Eval.	F8.8 L-t. Planning
SR-0950-ORG-QC	Data QC Procedure - File Listing Check	MUST																															M															
SR-0960-ORG-QC	Data QC Procedure - Report Inspection	MUST																															M															
SR-0965-ORG-USER	Data User Services	MUST		M																																												
SR-0969-ORG-PROD	Prod. Control Board (PCB)	MUST	X	X						X	M	X	X																	X					X		X		X	X	X							
SR-0970-ORG-PROD	Algorithm Improvement	MUST										M																																				
SR-0971-ORG-PROD	Defining Data Stability Methods	MUST									M																																					
SR-0980-ORG-PROD	Scientific Methods - Documentation	MUST					X				M																																					
SR-0990-ORG-PROD	Alg. Improvement - Sci. Documentation	MUST					X				M	X																																				
SR-0999-ORG-EVOL	Long-term System Evolution Planning	MUST																																														M
SR-1000-ORG-EVOL	Alg. Improvement - SW Documentation	MUST					X					M																																				
SR-1010-ORG-EVOL	Processor Module/s Evolution	MUST										X			X	M	X																															
SR-1020-ORG-EVOL	L2 Data Providing System Evolution	MUST												X													M																					
SR-1030-ORG-EVOL	Team Collaboration Framework	MUST		X			X		M		X	X				X	X	X	X	X		X	X	X	X																X	X	X	X				
SR-1035-ORG-RETR	Preservation of Data and Documents	MUST		X			X		X										X			X	X	X				X	X												M	X		X				
SR-0000-MIS	GCOS Requirements	MUST								X	M	X																																				

M - Major responsibility for requirement | X - Allocation of requirement to system function | c - Contribution to implementation of system requirement