Long Term Preservation of Earth Observation Space Data

Earth Observation Preserved Data Set Content

CEOS-WGISS

Data Stewardship Interest Group

Doc. Ref.: CEOS/WGISS/DSIG/EOPDSC

Date: September, 2015

Issue: Version 1.0

Document Status Sheet

Issue	Date	Comments	Editor
1.0	15 September 2015	New document evolved from EO Preserved Data Set Content, Issue 4.0	M. Albani, R.Leone, I. Maggio, R. Cosac

Table of Contents

List of Tables

1. INTRODUCTION

1.1 Intended Audience

The document is intended to assist data managers in Earth Observation data centres in applying the CEOS Best Practices for Long Term Preservation of Earth Observation Space data (www.ceos.org) to ensure EO mission data set assets preservation, curation, and valorisation for long-term accessibility and exploitation.

1.2 Background

In 2006, the European Space Agency (ESA) initiated a coordination action to share a common approach towards the long-term preservation of Earth Observation space data among all European and Canadian data holders and archive owners. A Long Term Data Preservation (LTDP) Working Group was formed in Europe in 2007 to define and promote a coordinated approach for long-term data preservation and curation of European Earth Observation space data assets. One of the outputs of the group consisted of the 'EO Preserved Data Set Content', a best practice document guiding Earth Observation data holders in their preservation activities [RD-1]. The 'CEOS Preserved Data Set Content' generated in the frame of the CEOS WGISS Data Stewardship Interest Group (DSIG), has evolved from the European document to become a global reference for Earth Observation data preservation.

1.3 Document Scope

This document identifies the EO mission data set assets content (i.e. data records and associated knowledge) that should be preserved to ensure long-term usability and exploitation of Earth Science data.

The document is intended to provide the content description (the “what”) for all the items of the EO mission data records and knowledge that should be preserved beyond the mission lifetime. It is intended as a guideline on how to use the content description list to support CEOS Best Practices associated documents.

The document is also intended to assist data managers in making sure that the recommended and mandatory content is collected, certified for completeness and quality at data set generation, during each mission stage, thereby providing the list of expected documents, content and quality information to be generated and preserved at each stage.

In accordance with the CEOS Best Practices, the composition of the PDSC varies by sensor category and needs to be tailored for the specific data set at hand, taking into consideration the designated community, preservation objective, requirements, quality information, metadata generation and dependencies.

The data manager shall tailor the PDSC to meet the needs of the specific mission, stating which data records and knowledge should be preserved during each phase of the Preservation Workflow in accordance with [AD-1] and maintain the Preserved Data Set Content inventory table with the data records, information, and software available under configuration, in accordance with [AD-2].

1.4 Applicable and Reference Documents

ID	Resource
[AD-1]	CEOS Best Practices on Long Term Preservation of Earth Observation Space Data - EO Data Stewardship Definition
[AD-2]	CEOS Best Practices on Long Term Preservation of Earth Observation Space Data – Preservation Workflow – March 2015 Version 1.0
[AD-3]	CEOS Best Practices on Long Term Preservation of Earth Observation Space Data –EO Data Preservation Guidelines, CEOS/WGISS/DSIG/EOPG, June 2015
[AD-4]	CEOS Best Practices on Long Term Preservation of Earth Observation Space Data – Generic Earth Observation Data Set Consolidation Process version 1.0 March 2015
[AD-5]	CEOS Best Practices on Long Term Preservation of Earth Observation Space Data – Persistent Identifier
[AD-6]	CEOS EO Data Purge Alert Procedure, http://wgiss.ceos.org/purgealert/
[AD-7]	Quality Assurance Framework for Earth Observation - Guidelines Framework (QA4EO). www.qa4eo.org.

Table 1: Applicable Documents

ID	Resource
[RD-1]	EO Preserved Data Set Content v 4.0, LTDP-GSEG-EOPG-RD-11-0003, July 2012
[RD-2]	ISO 14721 - OAIS standard (ISO reference model for Open Archival Information System) Pink Book, Consultative Committee for Space Data Systems, Greenbelt, MD. August 2009.
[RD-3]	European Cooperation for Space Standardization (ECSS), http://www.ecss.nl/
[RD-4]	Producer-Archive Interface Methodology Abstract Standard (PAIMAS), 05/2004, CCSDS 651.0-M-1
[RD-5]	Producer-Archive Interface Specification (PAIS), CCSDS 651.1-R-1, 02/2012

Table 2: Reference Documents

2. HOW TO USE THIS EO PDSC

The need for accessing historical Earth Observation data information has greatly increased, driven by long term scientific and environmental monitoring.

This document is meant to provide assistance to the practical implementation at working level of [AD-1] to [AD-7], providing recommended guidelines in reply to:

- the "what" dimension
- the "when" dimension
- the "quality" dimension
- the "preservation metadata" dimension
- the "how" dimension

2.1 "What" Dimension

This document has undergone a significant public review of the "what" dimension, i.e. the content specification of what is mandatory to preserve beyond the mission lifetime (i.e. the measurements for which the instrument was designed for), either raw data (as acquired by the satellite and recorded at the stations or received via Third Parties), or otherwise, global or higher level mission products when systematically generated and/or reprocessed as part of the mission requirements.

It identifies all additional information required to correctly understand and interpret the primary data, including in particular ancillary data (e.g. spacecraft ephemeris information, attitude, etc.), auxiliary data (required to process the telemetry payload data to generate the nominal mission products), CAL/VAL databases, whenever available (including processing/reference validation data sets) and mission-related documentation, including description of mission products and of the algorithms needed to obtain them.

The detailed list is provided in Chapter 3.

2.2 "When" Dimension

The experience with historical mission recovery has underlined the need to ensure that the mission asset content is qualified as fit for purpose for the long term preservation, during the mission lifetime, in accordance with quality certifying processes. The "when" dimension described in Chapter 3 is intended to assist data managers in ensuring that the recommended and mandatory content is collected, certified for completeness and quality at data set generation, during each mission stage, thereby providing the list of expected documents, content and quality information.

2.3 "Quality" Dimension

The need to preserve EO mission data assets indefinitely has led to the establishment of several CEOS Best Practices on Long Term Preservation of Earth Observation Space Data [AD-1] to [AD-6]. To guarantee that the preserved data set is "fit for purpose", it is mandatory to preserve its quality information. This is the objective of the Quality Assurance for Earth Observation (QA4EO) framework [AD-7] developed by the CEOS Working Group on Calibration and Validation (CEOS WGCV).

The Quality Assurance for Earth Observation (QA4EO) framework aims to provide EO data users with sufficient but simple information to enable them to evaluate the fitness for purpose of data/information for their applications, while also facilitating harmonisation and interoperability of

data sources. The key principle is stated in QA4EO Study results as: Data and derived products shall have associated with them an indicator of quality to enable users to assess their suitability for particular applications, i.e. their "fitness for purpose".

This can be expanded further, requiring that all EO data and derived products have associated a documented and fully traceable quality indicator, where:

- A Quality Indicator shall provide sufficient information to allow all users to readily evaluate the "fitness for purpose" of the data or derived product.
- A Quality Indicator shall be based on a documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards.

A Quality Indicator may be a number, set of numbers, graph, uncertainty budget, or a simple "flag" (see Table 8 for a, non-exhaustive, list of Quality Indicators).

To address this, QA4EO contains a set of guiding principles, supported by a suite of "key guidelines" based on existing best practises [AD-7].

However, the concept of a quality indicator is of limited use for the purposes of deciding which one needs to be preserved, as a quality indicator appropriate for one set of users, may not apply to all users who might need different indicators.

Instead, in this document, the concept of quality information is defined. This represents the information needed to define a quality indicator, i.e. to assess the fitness for purpose of the EO data records. This information is part of the Preserved Data Set Content specification and this document indicates where the quality information should be found and at which stage of the mission lifetime (chapter 3).

2.4 "Preservation metadata" Dimension

Preservation metadata is defined as the metadata information that the data manager and steward need in support to the digital preservation process, stewardship and curation objectives as defined by the CEOS Best Practices in [AD-1] to [AD-6]. According to [RD-4] preservation metadata shall be generated during the life cycle of the asset to be preserved.

There are different types of descriptive metadata: domain specific, administrative (including rights and permissions), technical, documenting digital provenance, documenting relationships and links in the preservation repository.

2.5 "How" Dimension: Tailoring the PDSC

In accordance with the CEOS Best Practices, the composition of the PDSC varies by sensor category and needs to be tailored for the specific data set at hand, taking into consideration the designated community, preservation objective, requirements, quality information, metadata generation and dependencies.

The data manager shall tailor the PDSC to meet the needs of the specific mission, stating which data records and knowledge should be preserved during each phase of the Preservation Workflow in accordance with [AD-1] and maintain the Preserved Data Set Content inventory table with the data records, information, and software available under configuration, in accordance with [AD-2].

This tailoring should involve mission experts (e.g. instrument designers, quality working groups), but also the data end user communities, to ensure that their needs have been taken into account. The

tailored document should have a defined owner and should be kept under review throughout the mission, at a minimum at the end of each mission stage.

Procedures must be in place to ensure that all quality information identified in the PDSC as being needed is in fact saved in the correct place. A common issue for historical missions was that significant information was captured in less formal ways, making it almost impossible to retrieve afterwards.

These procedures should include acceptance of a document being conditional on the quality information being complete, review of the quality information at milestones, and the transfer of all quality information to archives at the appropriate time.

Should some quality information be required but not present (e.g. because it is recorded in a different document than that specified by the PDSC), then the tailored PDSC should be updated to reflect the actual situation.

All quality information must be stored using the processes described in the Preservation Workflow CEOS Best Practices [AD-2].

The following requirements should apply for the tailoring:

- **R01: The PDSC document should be tailored for each mission and instrument.**
- **R02: The PDSC tailoring should be reviewed at least at the start of each mission stage.**
- **R03: The tailored PDSC should be made available to the designated community for review and feedback.**
- **R04: All quality information identified in the tailored PDSC should subsequently be documented and saved, with clear procedures in place to ensure this.**
- **R05: To facilitate checking that the PDSC has been complied to, all items of quality information should be given an identifier specifying the row that they correspond to.**
- **R06: To ensure that quality information required is available, the PDSC should be used to define deliverables for (sub) contracts.**
- **R07: The project office should maintain a directory of the knowledge information, and specifically of the linkages between items.**
- **R08: A suitable tool shall be developed to record and allow traceability of linkages between items of quality information.**
- **R09: The tool shall be used to record the quality information and the linkages between items, for a given mission or instrument.**
- **R10: A copy of the knowledge information identified in the PDSC should be stored in the same archiving centre as the data records.**
- **R11: Whenever possible, an automatic tool shall be provided to allow tracing the knowledge information relevant for a given data record.**
- **R12: Effort shall be made to ensure that all documentation, tools, calibration data and other associated knowledge are free from any legal or commercial restriction.**

3. PRESERVED DATA SET CONTENT SPECIFICATION

The Preserved Data Set Content specification is intended to provide the content description (the "what") for all the items of the EO mission data records and knowledge that should be preserved beyond the mission lifetime.

In the document the term Mission is used generically and includes the concept of "Experiment", "Campaign", "Project", etc.

EO Missions/Sensors Dataset is defined as:

- **Data Records:** these include raw data and/or Level-0 data, higher-level products, browse images, auxiliary and ancillary data, calibration and validation data sets, and descriptive metadata;
- **Associated Knowledge:** this includes all the Tools used in the Data Records generation, quality control, visualization and value adding, and all the Information needed to make the Data Records understandable and usable by the Designated Community (e.g. mission architecture, products specifications, instruments characteristics, algorithms description, calibration and validation procedures, mission/instruments performances reports, quality related information). It includes all Data Records Representation Information, Packaging Information and Preservation Descriptive Information according to the OAIS information model (part of this information might be included in the descriptive metadata depending on the specific implementation).

3.1 Data Records

Data records are identified as:

1. Raw data[1]
2. Level 0 data (L0)
3. Level 1 (L1) to higher levels mission data products when generated as part of the mission requirements and/or reprocessed
4. Browses whenever generated
5. Ancillary data (spacecraft ephemeris information, attitude, etc.)
6. Auxiliary data (required to process the telemetry payload data to generate the nominal mission products)
7. Calibration and validation datasets[2] (needed to calibrate the satellite instruments and monitor data quality)
8. Metadata

[1] Raw data shall be preserved whenever conversion to Level 0 cannot be adequately certified.

[2] Including processing/reference validation data sets.

3.2 Tools

This includes:

1. L0 consolidation software[3]
2. Data processing software (for products generation from Level 0 to higher levels according to mission requirements)[4]
3. Quality control software
4. Data/products visualization tools
5. Value adding tools

3.3 Information

It is assumed that each document part of the "Information" generated and identified in one of the stages below is maintained and updated in the following stages according to mission evolution. Documents that might evolve are included below only in the first stage during which they are generated even if they are maintained and updated during the subsequent stages of the mission.

Mission or project related documentation is generally identified by:

1. Mission architecture documents describing purpose, scope and performances of the mission and of the on-board instruments, information regarding relevant orbits, platform position, attitude, ground coverage (acquisition footprint), head-roll-pitch.
2. Documents describing data and product format specifications.
3. Documents describing measurement requirements and/or measurement performances (theoretical models). Documents regarding instruments characteristics, performances and instrument description (physical implementations). Documents describing models and/or algorithms needed/used to obtain mission data and products, including specific/special cases, known errors and configuration necessities. In other words, all documents covering the conceptual environment, its implementation and its operations.
4. Reports concerned with measurement trends, failures, changes of performances, un-availabilities
5. Reports and outcomes from events such as: congresses, studies, communities and investigators concerned with models' review, algorithm changes, and Cal/Val changes affecting data processing chains.
6. Documents related to the process of data qualification: precision, numerical representations, formats, uncertainties, errors, adjustment/correction methods (e.g. Cal/Val procedures and documents).
7. Document related to workflows, work procedure, documentation three and bi-directional link
8. Scientific publications based on the data exploitation or relevant to them (properly linked to the data) and outreach material.
9. Administrative (Memorandum, Intellectual Property Rights, etc.)
10. Mission Data Records and Documentation Tree

[3] Whenever raw data are preserved.

[4] Data Processing Software could be maintained in operation to generate mission products or all products could be generated through a Bulk Processing Campaign and Software code and algorithms archived.

Mission documentation shall include Representation Information, Packaging Information and Preservation Descriptive Information according to OAIS Information Model [RD-4].

4. PRESERVED DATA SET CONTENT FOR EARTH OBSERVATION MISSIONS

The PDSC should be tailored appropriately for each mission/instrument. The tailoring of the PDSC should involve mission experts (e.g. instrument designers, quality working groups) and the designated user communities. The tailored document should have a defined owner, and should be kept under review throughout the mission lifecycle, at the end of each mission phase/stage. Procedures must be in place to ensure that all quality information identified in the PDSC as traceable and preserved. These procedures should include acceptance of a document being conditional on the quality information being complete, review of the quality information at milestones, and the transfer of all quality information to archives at the appropriate time

An Earth Observation space mission is generally divided into the following stages:

1. **Mission Concept (MC).** Defines the mission to a sufficient level to show the scientific value and technical feasibility. During this stage, identification of the science requirements by the Science study team and study scientist are carried out. Additional activities include the identification of a reference platform to be used in the preliminary system level studies. Feasibility verification documents, mission technology and programmatic estimates for the future mission stages are also generated. According to ECSS standards [RD-3] the Mission Concept stage can be identified as Phase A of mission design.

2. **Mission Definition (MD).** This stage is concerned with the mission scientific requirements detailed definition and the selection of technical solutions for system concept. During this stage, types of scientific instrument measurements (e.g. spectral analysis, temperature measurement, etc.) are identified and defined eventually combining existing sensors/instruments in different modes or with different scientific models. According to ECSS standards the Mission Definition stage can be identified as Phase B of mission design.

3. **Mission Implementation (MI).** According to Mission Definition results, this stage produces the detailed definition and implementation of the mission system and components: sensors/instruments; algorithms and their relationship in the frame of scientific domains; methods of measurement and any other context necessary to perform measures. Production, development testing and pre-qualification of selected critical elements and components lead to the conclusion of the technology development activities. According to ECSS standards the Mission Implementation stage can be identified as Phases C/D of mission design and implementation.

4. **Mission Operations (MO).** This stage identifies the operational timeframe of the mission being the period during which data are captured, algorithms are revised and improved, activities concerned with input analysis, calibration and validation of sensor/instrument as well as activities concerned with qualification of processed data are performed. According to ECSS standards the Mission Operations stage can be identified as the Phase E Operations till the end of mission lifecycle.

5. **Post Mission (PM).** This represents the ***Post-Operations and Preservation stages***. The Post Mission stage is usually identified according to current ECSS standards as the Phase F of a mission. In this document the Post Mission stage has been extended and augmented and mainly focuses on the archived data to accommodate the need to preserve them in the long term for further reuse and exploitation. The post mission stage starts with the satellite end of life (e.g. for an Earth Observation mission with the event of satellite disposal or failure). The Post Mission stage focuses on datasets (data and information) consolidation and appraisal, datasets reprocessing to align to the latest version, ground segment and media disposal (depending on specific mission), and data and associated information migration to a

long-term preservation environment. During the Post Mission stage, a limited set of functions (e.g. data discovery and access) is provided by the mission ground segment (still in operation) according to the adopted strategy and depending on mission requirements until its disposal and data migration to long-term preservation. This stage also focuses on historical data reuse and exploitation, on data and concerned information preservation against aging and technological changes, and on data curation and enrichment).

4.1 Mission Concept Stage (MC)

Rationale – Information produced during this stage provides a snapshot of the scientific and technical framework in which the mission was born. Mission and sensors requirements, assessment studies, technology readiness review and cost analysis are performed during this stage. Preserving this information – both for approved and not approved missions – would allow future users to have reference material for new missions' evaluation and definition. Traceability of this information is also useful to compare initial expectations to what was actually achieved by the mission and to understand which changes occurred between the pre-mission and the next stages.

ID	Type	Identification	Description	Quality Information	Notes
MC 1.1	Doc	Scientific Scenario and User Communities	Defines scientific scenario and expected goals. Also list Principal Investigator, designated user communities and third party actors.	Required uncertainty for services and applications, lifetime, data availability, data accuracy, data latency, revisit time, geographical coverage, spatial resolution.	
MC 1.2	Doc	Mission Requirement Document	Defines scientific mission and sensor requirements, processing methods, qualification, methods	Calibration plan and quality assessment plan for the mission. Uncertainty requirements for instrument product (e.g. radiometric/geometric uncertainty, coverage, revisit time, etc.) Justification for the design decisions (e.g. band selection)	Most information should be contained in the mission documentation, e.g. the Mission Requirement Document (MRD), Mission Operations Concept Document (MOCD) and Mission Description Document (MDD) according to ECSS or equivalent standards.
MC 1.3	Doc	Mission Operation Plan	Defines the plan on how the mission will be conducted	Plan for handling quality information	

Table 3: Assets to be preserved during the Mission Concept Stage

4.2 Mission Definition Stage (MD)

Rationale: The Mission Definition stage produces the entire mission and data detailed definition documents. It includes Sensor/Instrument requirements, characteristics, calibration methods, etc. Preserving this information is fundamental to understand changes that may have occurred over time while in operation.

ID	Type	Identification	Description	Quality Information	Note
MD 1.1	Doc	Mission Requirements Specifications	Defines mission requirements, mission space to ground functional and resource allocation and operational scenario. Contains the specifications for the verification and validation method for space to ground resources	Description of the quality information at a global (e.g. revisit times and mission products uncertainty) and at a subsystem level: instrument e.g. straylight, channel crosstalk, spatial sampling. FoV, observation mode, spectral channels.	Most information should be contained in the System Requirement Document and Justification File, System Functional Specification, and Design Definition File (DDF), Design Justification File (DJF) documents according to ECSS standards and equivalent
MD 1.2	Doc	Space to Ground segment ICDs	Defines the main systems / segments ICDs, system budget estimation and data flow.	Error Control (e.g. CRC) data latency, data rate, quality flags, packet lost/damaged, timeliness etc. for different scenarios (e.g. Near-Real-Time NRT, calibration mode, ground stations availability and relative position)	Most information should be contained in the space-to-ground interface control document (SGICD) according to ECSS standards.
MD 1.3A	Doc	Sensor / Instrument requirements	Defines the Sensor / Instruments requirements for design (e.g. bands, modes, performances, etc.).	Sensor uncertainty budget based on previous knowledge. Specification of uncertainty associated with optical properties e.g. noise, linearity, calibration accuracy, signal synchronisation, electrostatic protection, temperature and pressure range.	
MD 1.3B	Doc / Data Record	Sensor / Instrument characteristic	Characteristic for processing of acquired data, data processing model	Assessment of performance/ acceptability including uncertainty, linearity, sun-glint, straylight: Documented model descriptions, validation of model and software, version control. Validation by comparison with other models or reference data sets including simulated products and ground measurements.	This includes validation campaigns with in-situ products

MD 1.4A	Doc / Data Record	Sensor / Instrument qualification process	Qualification process for sensor, captured data, processed data.	Documented procedure for validation.	
MD 1.4B	Doc / Data Record	On-ground calibration and characterisation plan	Calibration requirements	Identification of reference standards, pre-flight calibration methods, re-calibration intervals Uncertainty aims	Pre-launch calibration campaign includes: • Optical Tests • Thermal Test • External Calibration Test • Field Of View determination
MD 1.4C	Doc / Data Record	Ground/Ocean calibration reference and scientific base	Calibration requirements	Traceability to International System of Units (SI) via international reference standards: Procedures, calibration certificates, traceability statement, and uncertainty analysis.	Should include description of these sites, accuracy, stability of the site conditions.
MD 1.5	Doc	Processing algorithms and data format specification	Defines: Mathematical models and algorithms for mission data processing; Auxiliary and ancillary data orchestration; Data and Products format requirements and standards.	Documented descriptions of mathematical models and algorithms for mission data processing; including: Assessment of performance / acceptability. Peer reviewed papers Simulation for validation results. Validation by comparison with test datasets Validation of performance simulator. Auxiliary and ancillary data orchestration. Data and Products format requirements and standards including: Metadata specifications (including quality information/parameters) Naming conventions Version controls specification	Should define what validation evidence is required to accept any product.

Table 4: Assets to be preserved during the Mission Definition Stage

4.3 Mission Implementation Stage (MI)

Rationale: Preserving all the information produced during the Mission Implementation stage is needed to understand procedural impacts relative to instrument, algorithm and product implementation. Data acquired during the calibration and validation campaigns of the instrument under construction (e.g. in a laboratory or dedicated campaigns) is of critical importance as a reference for the future use of the data.

ID	Type	Identification	Description	Quality Information	Note
MI 1.1	Doc	Mission Design (Space and Ground Segment)	Defines mission requirements specification and implementation design.	Clear identification of technical procedure. Record of decision made during implementation	Most information should be contained in the System and Subsystems Requirement Documents and Justification Files, System Technical and Functional Specifications including Interface Requirements, Design Definition Files and Design Justification Files according to ECSS standards.
MI 1.2A	Doc	Detailed Space to Ground Segment Operations Concept and implementation	Defines the detailed Space to Ground operational implementation and any contingency procedure/plan needed	Recording procedure for assuring the data quality. Storing of diagnostic information received.	Most information should be contained in the consolidated Mission Operations Concept and Space to Ground Technical Budget documents according to ECSS standards.
MI 1.2B	Doc	Data Handling	Data Capture and handling	Clear identification of technical procedure	
MI 1.2C	Doc	On Board Processing	On board processing	Algorithm description and software validation	
MI 1.3	Doc	Sensor/Instrument Design and Implementation	Defines the Sensor/Instrument platform design and implementation and its performances Platform and instrument design implementation/test, budges performances.	Testing results including uncertainty. Uncertainty budget with supporting evidence (from on ground characterisation). Uncertainty combination, covariance.	In this item it is possible to include the information of the other relevant subsystem with a direct impact on the mission data performances (e.g. attitude and orbit determination subsystem).
MI 1.4	Doc / Data Records	Validation and Calibration	Independent validation and calibration campaign method, data validation activities with simulated data.	Calibration results, uncertainty budget with supporting evidence, traceability to SI validation results	This includes the pre-flight calibration/validation campaign and should focus on calibration rather than validation.

MI 1.5A	Doc	Ground Processors Design, Algorithms Implementation and Supporting Information for data processing.	Defines the design and implementation of the ground data processors and the algorithm. It includes also supporting information for data processing (e.g. ancillary, auxiliary data description & orchestration, etc.)	Algorithm description and software validation for all software used on ground and on board. Metadata and naming conventions, version control	
MI 1.5B	Notes/ Papers	Technical Notes Scientific Paper Peer Review	Version Control	Algorithm description and software validation for all software used on ground and on board. Metadata and naming conventions, version control	
MI 1.6A	Doc	Products Specifications	Provides a detailed description of product and their characteristic	Description of uncertainty/quality indicators and method to provide uncertainty to different users.	
MI 1.6B	Doc	Data Format Specifications	Contains information that will allow the user to read and use the data.	Data format naming conventions, performances of compression algorithm, quality indicator specification	
MI 1.6C	Doc / Data Records	Supporting Information for processing	Ancillary and auxiliary definition and identifications, orchestrations	Appropriate quality indicator for ancillary/auxiliary data to be used in the mission operations stage with the relevant metadata	
MI 1.7	Doc / Data Records	Qualification Process	Detailed qualification methods and data	Assessment of performances/acceptability	

Table 5: Assets to be preserved during the Mission Implementation Stage

4.4 Mission Operations Stage (MO)

Rationale: Data acquired during the Mission Operation stage is the concrete heritage that the mission will leave to future generations. The Mission Operation stage provides the effective data that will be analysed by the scientific community and that will the core of the mission preservation objective. The data also serve public administration and commercial applications, which depend on reliable, sustainable data availability to fulfil their public task and to set up viable business cases. Software related to this mission stage needs to be preserved in order to use, process and exploit data in the future. Documents also need to be preserved to have a comprehension of the data itself and to perform mission results qualification.

ID	Type	Identification	Description	Quality Information	Notes
MO 1.1	Doc	Mission data access and Service Requirements document and User Handbook	Defines the data archival and processing/reprocessing strategy, the data accessible to users and the services requirements & performances during the operations stage.	Clear identification of technical procedure	Phase E1 and Phase E2 ECSS standards equivalent
MO 1.2	Doc	Sensor Ground Segment Operations Plan	Defines the actual implementation of the end-to-end mission operations.	Uncertainty budget with supporting evidence	
MO 1.3	Doc / Data Records	Mission Operations Acquisition Plans and Reports	Describes the mission sensor acquisition plans and reports.	Availability of data, data quality, model evolution, calibration parameters evolutions, geo- location performance, data anomalies	Phase E1 and Phase E2 ECSS standards equivalent
MO 1.4	Data Records	Raw/Level 0	Raw or Level 0 data from the sensor or instrument data packets	Completeness of data, timeline, Certification of L0 processing (unless stored as raw). Noise – SNR & SD of the data	Raw data shall be preserved whenever conversion to Level 0 cannot be adequately certified
MO 1.5A	Data Records	Level 1	Processed image data L1 products	Associated uncertainties and evidence. Processing algorithm recorded and validated. Reference to calibrations, traceability. For geometrically located area, geometric alignment and resampling	
MO 1.5B	Data Records	Level 2	Processed image data L2 products and higher	Associated uncertainties and evidence. Processing algorithm recorded and validated. Reference to calibrations, atmospheric corrections, traceability. Reference	

				to validation where relevant	
MO 1.6	Data Records	Browses/Images	Browse Digital Catalogue	No specific quality information is needed for the browse images.	Whenever generated
MO 1.7	Data Records	Ancillary Data	Attitude, Ephemeris, Navigation parameters, Observation counters, Orbital State Vectors, Times, Sun Position, Temperatures Sensor/CCD/Amplifiers noises, Earth Relative position, Azimuth instrument parameters (e.g. optical response	Quality flags and performance parameters e.g. orbit accuracy, temperature stability	
MO 1.8	Data Records	Auxiliary Data	Band/Multispectral/ Band-by-band parameters for algorithms, Non linearity correction factors, Error/Failure/Gap correction factors, Calibration curve/Factors, Scaling correction factors, Atmospheric correction factors, geometry correction factors, drift factor, albedo parameters, instrument modes, incident angle, absolute calibration constants, solar radiance, moon temperature brightness, local seasonal variances, weather forecast/actual, wind, altimetry/geode model DEM, etc.	Associated uncertainties and evidence where appropriate otherwise performances flags and parameters e.g. drift. Sensitivity coefficients for L1 and L2 data to their parameters. Date range for auxiliary file version.	Required to process the telemetry payload data to generate the nominal mission products
MO 1.9	Doc/ Data Records	Calibration and validation data	Cal/Val data acquired during mission operations (optical/radiometric stability, Instrument availability, Internal calibration, Optic pointing pattern, etc.)	In-flight calibration reports, uncertainly with evidence, version report, instrument anomalies Parameters evolution (degradation model, DS, pixel response linearity…) Instrument validation: SNR validation, absolute and relative radiometric vicarious calibration, MTF, geolocation, L2 products… Validation reports, satellite uncertainties.	Cal/Val data acquired during mission operations through validation campaigns run to calibrate the satellite instruments and monitor data quality. It includes processing/reference validation data sets. Includes also related documentation (e.g. reports).

MO 1.10	Doc/Data Records	Quality Parameters	PA/QA of instrument, raw data and products	Assessment of performance/accept ability	
MO 1.11	Doc/Data Records	Metadata	Metadata Digital Inventory	No specific quality information is needed for the metadata	The metadata can be generated from auxiliary, ancillary and similar data and can always be recovered if appropriate procedures are set in place.
MO 1.12	SW Code	Level 0 consolidation		Algorithms and software verification / Validation, version control	
MO 1.13	SW Code	Data Processing Software	Instrument processing algorithms, context and source codes, testing context	Algorithm description. Algorithms and software verification / Validation, version control	
MO 1.14	SW Code	Quality Control Software		Algorithms and software validation, Algorithms and software verification / Validation, Version control	
MO 1.15	SW Code	Visualization Tools	Processing and visualizing tools	Software validation and version control Algorithms and software verification / Validation, Version control	
MO 1.16	SW Code	Value-Added Software		Software validation and version control Algorithms and software verification / Validation, Version control	
MO 1.17	SW Code	Data/ image processing	Packed telemetry, PUS, CCSDS, Instrument source packet, product formats, and storage formats.	Software validation and version control, software developments	
MO 1.18	Doc	Product qualification and quality assurance	Defines the product qualification process	Assessment of performance/accept ability based on relevant	

		monitoring reports	outputs.	quality parameters such as uncertainty levels, flags etc.	
MO 1.19	Doc	Sensor/Instrument evolution and history records	Describes any instrument event that might affect data quality (e.g. upgrading, downgrading, LUTs). It includes also known-errors and limits of sensors/instruments.	Instrument timeline Documented supporting evidence for decisions	
MO 1.20	Doc	Referred publications and papers	Referred publications, articles and technical notes clearly referencing the used datasets.	No specific quality information has been requested	Any future publication should be enforced to provide clear reference to the utilized dataset.
MO 1.21	Doc	Tandem and/or combined campaigns, comparisons	Data and reports	Uncertainty budgets with supporting evidence Comparisons report following QAE4EO Guideline 7	
MO 1.22A	Doc / Data Records	Cross- campaign, cross-comparisons and cross- calibration activities documentation and Data	Describes the cross campaign scenario and operational context. Also describes any cross-calibration activities	Evidence of participation in appropriate comparisons. Comparison report following QA4EO Guidelines 4 and 7	
MO 1.22A	Doc / Data Records	Cross- campaign, cross-comparisons and cross- calibration activities documentation n and Data	Describes the cross campaign scenario and operational context. Also describes any cross-calibration activities	Evidence of participation in appropriate comparisons. Comparison report following QA4EO Guidelines 4 and 7	
MO 1.23	Doc	Data Access Policy	Describes the data access policy for mission in the operational stage.		

Table 6: Assets to be preserved during the Mission Operations Stage

4.5 Post Mission Stage (PM) - Post-Operations and Preservation

Rationale: After the end of a mission, datasets acquired during the operational stage need to be consolidated and aligned to the latest available version of the processors and/or improved version. All the evolution activities carried out in the previous stages and the changes to the data and associated information are properly assessed and consolidated during this stage for end-to-end consistency/coherency/provenance based on the documentation produced and preserved in the previous stages. During this stage the user communities will still need to analyse and process data. Enhanced algorithms and processors improvements could be implemented to improve data exploitation and processing performances.

ID	Type	Identification	Description	Quality Information	Note
PM 1.1A	Doc	Data consolidation & reprocessing strategy, implementation plans, and consolidated/ reprocessed data. Processing	Processing and/or Calibration change including provenance e and context	Algorithms and software validation, version control Clear description of motivation for reprocessing and improvements gained	Level 0 data consolidation should be certified in this stage and in such a case raw data could be disposed.
PM 1.1B	Doc/Data Records	Data consolidation & reprocessing strategy, implementation plans, and consolidated/ reprocessed data. Ancillary, Auxiliary	Updated Ancillary, Auxiliary	Associated uncertainties and evidence, version control	
PM 1.1 C	Doc/Data Records	Data consolidation & reprocessing strategy, implementation plans, and consolidated/ reprocessed data. PA/QA	Quality information updated as part of reprocessing	Assessment of performance/ Acceptability	
PM 1.2	Data Records (Reproce ssed data set)	Data consolidation & reprocessing strategy, implementation plans, and consolidated/ reprocessed data. L0, L1, L2	Reprocessed data & products	Associated uncertainties and evidence, version control	
PM 1.3	Doc	Data consolidation & reprocessing strategy, implementation plans and consolidated/ reprocessed data. Data/Image processing	Instrument processing algorithms,	Algorithm and software validation, version control	
PM 1.4	Data	Data consolidation & reprocessing strategy, implementation plans and consolidated/ reprocessed data. Browse Metadata	Metadata Inventory		

PM 1.5	Docume ntation	Referred publications and papers	Referred publications, articles and technical notes clearly referencing the used datasets.	PID	
PM 1.6	Doc	Historical Data Access Policy	Describes the data access policy for the historical mission in the Preservation stage.		
PM 1.7	Doc	Historical Mission User Handbook	Describes the consolidated end-to-end mission description, data formats, operational scenarios, and all information necessary for future data use. It includes also the appraisal of the mission datasets (i.e. their value).	Summary of quality information approach within mission / instrument	Generated starting from information collected in the previous stages.

Table 7: Assets to be preserved during the Post Mission Stage

ANNEX A – QUALITY INDICATORS

Term	Definition
Quality Indicator	A means of providing a user of data or derived products (resulting from processing of data) with sufficient information to assess its suitability for a particular application. This information should be based on a quantitative assessment of its traceability to an agreed reference or measurement standard (ideally SI), but can be presented as numeric or a text descriptor, providing the quantitative linkage is defined. For many missions this will mean a documented and complete uncertainty budget (see QA4EO-QAEO-GEN-DQK-006), with quantitative evidence of traceability (see QA4EO-QAEO-GEN-DQK-007), though for some applications it will be sufficient to describe biases to agreed references or other sensors. The QI is likely to be presented as a report.
Uncertainty	Non-negative parameter characterising the dispersion of the quantity values that are being attributed to a measure and (quantity), based on the information used. A measure of the standard deviation of the probability distribution for the measure. Where possible this should be derived from an experimental evaluation but can also be an estimate based on other information, e.g. experience. Uncertainty evaluation should start by identification of a measurement equation. The sensitivity of the determined measure and to each effect in the measurement equation can be calculated either through partial derivation of the measurement equation, or through experimental investigation of the effect. The different uncertainty contributions are listed in an "uncertainty budget" and combined in quadrature. The standard uncertainty can then be, as appropriate, expanded with a coverage factor, for example to obtain a 95 % confidence level. The analysis of uncertainty is described in QA4EO-QAEO-GEN-DQK-006.
Traceability	Property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty. In practice traceability is obtained by a series of comparison each of them involves calibration standard at one level in the chain using a standard at a higher level. Ideally traceability will lead back to the SI, through a National Measurement Institute. For example an irradiance-mode radiometer may be calibrated against a standard irradiance source (lamp), which was calibrated against a primary irradiance source at a National Measurement Institute (a blackbody), whose irradiance properties were known due to a filter radiometer (effectively an absolute pyrometer), which was calibrated against the primary radiometric reference (the cryogenic radiometer) and thus to SI. At each stage in the traceability chain there needs to be documented evidence of traceability, in the form of a calibration certificate, along with documented procedures and validation. This is described in QA4EO-QAEO-GEN-DQK-007.
Sensitivity (Coefficients)	This determines how sensitive the measure and (e.g. a L1 or L2 data product) is to any particular source of uncertainty. Some sensitivity coefficients can be calculated by differentiating the measurement equation (e.g. an inverse square law behaviour makes the sensitivity of irradiance to distance a factor of two: a 1 % change in distance, makes a 2 % change in irradiance). Other sensitivity coefficients are determined experimentally, e.g. by changing the temperature of the sensor, it is possible to determine how sensitive the signal on that sensor is to temperature changes. It may also be necessary to determine the sensitivity of model results to changes in the assumptions of that model.

Term	Definition
Calibration	Assessment of the correct values to the instrument's measurement scale by comparison with a reference standard of higher accuracy (higher level at the traceability chain). For example an instrument's spectral radiance responsivity is calibrated by putting it in front of a reference radiance source, whose radiance is determined traceable. Every step of a calibration chain needs documentation, including reference standard properties and suitability (see QA4EO-QAEO-GEN-DQK-003), documented procedures (see QA4EO-QAEO-GEN-DQK-002) and evidence of traceability (QA4EO-QAEO-GEN-DQK-007).
Reference Standard	Realisation of the definition of a given quantity, ideally with a stated uncertainty, which can be used as a reference; it can be individual or community defined. A reference standard can be an artefact such as a lamp or a reflectance tile of known and certified irradiance or reflectance and associated uncertainty. The measurements against a reference standard are calibrations. A reference standard can be a calibrated instrument that is compared with the test instrument. A reference standard might refer to the calibration sites that had been previously characterised and are monitored from the ground. In all cases a reference standard needs to have known properties, with formal calibration, and must be used within its range of validity and in an appropriate manner. This process must be documented (see QA4EO-QAEO-GEN-DQK-003).
Validation	Confirmation that the performance (of an instrument, algorithm, or software) that fits the intended purpose. Performance of instruments and software can be validated by testing performance against known standards, and formal auditing processes. See QA4EO-QAEO-GEN-DQK-005 (for software and algorithms).
Supporting evidence	Documentation describing how a process was carried out and its traceability. Includes calibration certificates, documentary procedures, records of software validation, records of traceability. See QA4EO-QAEO-GEN-DQK-007 for a list of suitable evidence of traceability.
Comparisons	Organised peer-to-peer comparisons, where different sensors/calibration laboratories/etc. measure the same reference standard and results are compared with each other in a formal way. A formal comparison will follow the procedure described in: QA4EO-QAEO-GEN-DQK-004.

Table 8: Quality Indicators