Joint Earth Observation Mission Quality Assessment Framework - SAR Guidelines

National Aeronautics and Space Administration

Joint Earth Observation Mission Quality Assessment Framework - SAR Guidelines

Signature/Approval Page

Accepted by:

2 July 2024

Henri Laur
Head of the EO Management & Product Quality Division
EO Mission Management & GS Department
Directorate of EO Programmes
ESA

Date

6/28/2024.

Melissa Yang Martin
Commercial SmallSat Data Acquisition Program Manager
Earth Science Division
NASA

Date

Preface

This document is part of the Commercial SmallSat Data Acquisition (CSDA) Program under the CSDA Project configuration control. Changes to this document shall be verified by a document change notice (DCN) and implemented by change bars or by complete revision.

The material in this document is advisory in nature and is not legally binding. Conformity with this guidance is voluntary only, and nonconformity will not affect the rights or obligations of any entity. This document does not constitute an international agreement.

Abstract

This Joint Earth Observation Mission Quality Assessment Framework – SAR Guidelines document is created for the benefit of the Earthnet Data Assessment Project (EDAP) project and Commercial SmallSat Data Acquisition (CSDA) Program as part of a collaboration between ESA and NASA.

Keywords: configuration management, document control, quality assessment, SAR

Cover Art: Cover art is AI generated graphic using Microsoft Copilot Designer using term "commercial satellite multispectral radar observation across Atlantic AND Northern Hemisphere AND digital downlink"

Authored and prepared by

Samuel E. Hunt
National Physical Laboratory
United Kingdom

Davide Giudici
Advanced REmote-sensing SYStems (ARESYS)
Italy

Franz Meyer
University of Alaska, Fairbanks
United States of America

Bruce Chapman
Jet Propulsion Laboratory
National Aeronautics and Space Administration
United States of America

Clément Albinet
European Space Agency
Intergovernmental

Juval Cohen
Finnish Meteorological Institute
Finland

Batuhan Osmanoglu
National Aeronautics and Space Administration
United States of America

Alfreda Hall
National Aeronautics and Space Administration
United States of America

Jaime Nickeson
Science Systems and Applications Inc
National Aeronautics and Space Administration
United States of America

Change History Log

Revision	Effective Date	Description of Changes
1.0	04/23/24	First public release
0.2	1/31/22	Updated Figure 2. Removed Figure 4.
0.1	9/12/21	Development draft

Table of Contents

1. Introduction **9**
1.1. Scope 9
2. EO Mission Quality Assessment Framework Summary **9**
2.1. Quality Assessment Report 10
2.2. Cal/Val Maturity Matrix 10
2.2.1. Summary Cal/Val Maturity Matrix 10
2.2.2. Detailed Validation Cal/Val Maturity Matrix 11
2.3. Approach to Grading 12
2.4. Considerations for the SAR domain 12
3. Data Provider Documentation Review **13**
3.1. Product Information 13
3.1.1. Product Details 14
3.1.2. Availability & Accessibility 15
3.1.3. Product Format, Flags and Metadata 16
3.1.4. User Documentation 17
3.2. Metrology 18
3.2.1. Sensor Calibration & Characterization 18
3.2.2. Geometric Calibration & Characterization 20
3.2.3. Metrological Traceability Documentation 21
3.2.4. Uncertainty Characterization 22
3.2.5. Ancillary Data 23
3.3. Product Generation 24
3.3.1. Image Formation and Calibration Algorithm 24
3.3.2. Geometric Processing 25
3.3.3. Higher Level Product Generation 26
3.3.4. Mission Specific Processing 27
4. Detailed Validation **28**
4.1. Detailed Validation Grading Criteria 28
4.1.1. Validation Method 28
4.1.2. Validation Results Compliance 29
4.2. Performance Metrics 30
4.2.1. Measurement Validation 30
4.2.2. Geometric Validation 32

4.3. Validation Summary 34
5. References 34

List of Figures

Figure 1. Summary Cal/Val Maturity Matrix 11
Figure 2. Validation Cal/Val Maturity Matrix for the SAR domain 12
Figure 3. Data Provider Documentation Review Matrix 13

List of Tables

Table 1. Product Information > Product Details – Assessment Criteria 15
Table 2. Product Information > Availability and Accessibility – Assessment Criteria 16
Table 3. Product Information > Product Format, Flags and Metadata – Assessment Criteria 17
Table 4. Product Information > User Documentation – Assessment Criteria 18
Table 5. Metrology > Sensor Calibration & Characterization – Assessment Criteria 20
Table 6. Metrology > Geometric Calibration & Characterization – Assessment Criteria 21
Table 7. Metrology > Metrological Traceability Documentation – Assessment Criteria 22
Table 8. Metrology > Uncertainty Characterization – Assessment Criteria 23
Table 9. Metrology > Ancillary Data – Assessment Criteria 24
Table 10. Product Generation > Image formation and radiometric Calibration Algorithm – Assessment Criteria 25
Table 11. Product Generation > Geometric Processing – Assessment Criteria 26
Table 12. Higher Level Product Generation > Retrieval Algorithm – Assessment Criteria 27
Table 13. Product Generation > Mission Specific Processing – Assessment Criteria 27
Table 14. Validation > Validation Method – Assessment Criteria 29
Table 15. Validation > Validation Results –Assessment Criteria 30

Acronyms & Abbreviations

APA	Absolute Position Accuracy
ARD	Analysis Ready Data
ATBD	Algorithm Theoretical Basis Document
CF	Climate & Forecast (Metadata Convention)
CEOS	Committee on Earth Observation Satellites
DOI	Digital Object Identifier
EDAP	Earthnet Data Assessment Project
ENL	Equivalent Number of Looks
EO	Earth Observation
ESA	European Space Agency
FAIR	Findable, Accessible, Interoperable and Reusable
FRM	Fiducial Reference Measurement
GUM	Guide to the Expression of Uncertainty in Measurement
HH	Sensor polarization: Horizontal transmit, Horizontal receive
HV	Sensor polarization: Horizontal transmit, Horizontal receive
INSPIRE	Infrastructure for Spatial Information in Europe
IRF	Impulse Response Function
ISLR	Integrated Side Lobe Ratio
L1	Level 1
L2	Level 2
NES0	Noise Equivalent Sigma Nought
NESZ	Noise Equivalent Sigma Zero
NetCDF	Network Common Data Form
NPL	National Physical Laboratory, UK
PSLR	Peak Side Lobe Ratio
PUG	Product User Guide
PUM	Product User Manual
QA	Quality Assessment
QA4EO	Quality Assurance Framework for Earth Observation
QA4ECV	Quality Assurance Framework for Essential Climate Variables
RCS	Radar Cross Section
SAR	Synthetic Aperture Radar
SI	Système International (International System of Units)
URL	Universal Resource Locator
VH	Sensor Polarization: Vertical transmit, Horizontal receive
VV	Sensor Polarization: Vertical transmit, Vertical receive

1. Introduction

In recent years, the increasing range of applications of Earth Observation (EO) data products and availability of low-cost launch service has resulted in a growing number of commercial EO satellite systems, developed to deliver end-to-end information services, with an increasing number in the Synthetic Aperture Radar (SAR) domain. This evolution in the marketplace has led to increasing interest from Space Agencies in the acquisition of commercial EO data products, as they may provide complementary capabilities and services to those currently offered.

To this end, both ESA and NASA have initiated activities aimed at assessing the quality and utility of such data products. On the ESA side, the Earthnet Data Assessment Project (EDAP) project (Mannan et al. 2019) performs early data assessments on EO missions in optical, atmospheric and SAR domains. In the meantime, the NASA Earth Science Division has initiated a Commercial SmallSat Data Acquisition Program, completed a pilot study (NASA Earth Science Division 2020), and entered the sustained use phase for some of the commercial data sets.

To ensure that decisions on acquisition of commercial data are acquisitions can be made with confidence, there is a need for an objective framework with which to assess the data quality of these commercial sources. The ESA EDAP project established such an EO mission quality assessment framework (Hunt 2021). Presented here is the framework, now developed as a collaboration between ESA and NASA.

This document is developed with the evolution of the marketplace and the advance of Earth sciences and applications of EO data products in mind, and thus maybe revised as appropriate.

1.1. Scope

This document is intended to provide specific guidelines for mission quality assessment of SAR sensors, as part of the implementation of the generic EO mission quality assessment (Hunt 2021) for this domain. Section 2 provides a summary of the mission quality assessment framework. Section 3 provides a review of the SAR mission quality, as evidenced by its documentation. Finally, Section 4 provides guidelines for verifying the mission data quality is consistent with the stated performance of the sensor.

2. EO Mission Quality Assessment Framework Summary

This section outlines the overall EO mission data product quality assessment framework. The evaluation is primarily aimed at verifying that mission data has achieved the claimed mission performance and, where applicable, reviews the extent to which the missions follow community best practice in a manner that is “fit for purpose”.

The approach taken to assess data product quality is based on the QA4EO principle (QA4EO Task Team 2009) and builds on the structure and reporting style developed in other similar work (Nightingale et al. 2019). This quality assessment framework was initially developed within the

ESA Earthnet Data Assessment Project (EDAP) project and aims to build on the experience of previous work targeting the satellite Cal/Val context.

The assessment itself is conducted in two parts, as follows:

- Documentation Review – review of mission quality as evidenced by its documentation.
- Detailed Validation – quantitative assessment of product compliance with stated performance.

These parts of the assessment, along with their grading criteria, are described in Sections 3 and 4, respectively. The activities are divided into sections and subsections constituting each of the different aspects of data product quality that are assessed and graded. Assessment results are provided in a separate Quality Assessment (QA) Report and are also summarized in a color-coded Cal/Val maturity matrix.

It is expected that all relevant mission information needed to perform the assessment would be available to all users, however it is understood that confidentiality may be required for some aspects of a mission. Where this is the case, it maybe indicated as confidential in the quality assessment report. In general, pertinent key conclusions of confidential documentation should nevertheless be published openly.

2.1. Quality Assessment Report

The quality assessment for a given mission is reported using the QA Report template. The template ensures consistency of reporting and facilitates comparison between the assessments of similar missions. The QA Report covers each section of analysis, providing more detailed information, as well as including a completed mission Cal/Val maturity matrix (see following subsection) presenting the results of each subsection of analysis in a color-coded table.

2.2. Cal/Val Maturity Matrix

A Cal/Val maturity matrix provides a high-level color-coded summary of the quality assessment results. The matrix contains a column for each section of analysis, and cells for each subsection of analysis. Subsection grades are indicated by the color of the respective grid cell, which are defined in the key. A padlock symbol in the corner of a given cell indicates that the information used to assess the respective subsection is not available to the public. The reporting of assessment results is divided between two Cal/Val maturity matrices, as follows:

- Summary Cal/Val Maturity Matrix
- Detailed Validation Cal/Val Maturity Matrix

These matrices are described below.

2.2.1. Summary Cal/Val Maturity Matrix

The Summary Cal/Val Maturity Matrix provides an overall summary of the quality assessment results (see Figure 1). The matrix on the left (in dark blue) summarizes the results of the Documentation Review, while the additional column on the right (in light blue) summarizes the results of the Detailed Validation. The validation summary column is separated from the main table to make clear the results can come from multiple assessment sources.

Data Provider Documentation Review		
Product Information	**Metrology**	**Product Generation**
Product Details	Radiometric Calibration & Characterization	Radiometric Calibration Algorithm
Availability & Accessibility	Geometric Calibration & Characterization	Geometric Processing
Product Format, Flags & Metadata	Metrological Traceability Documentation	Retrieval Algorithm
User Documentation	Uncertainty Characterization	Mission Specific Processing
	Ancillary Data	

Validation Summary
Radiometric Validation Method
Radiometric Validation Results Compliance
Geometric Validation Method
Geometric Validation Results Compliance

Key
Not Assessed
Not Assessable
Basic
Good
Excellent
Ideal

Not Public

Figure 1. Summary Cal/Val Maturity Matrix (To be filled in to report the result of assessment.)

2.2.2. Detailed Validation Cal/Val Maturity Matrix

The Detailed Validation Cal/Val Maturity Matrix (see Figure 2) provides more complete reporting of the Detailed Validation Summary assessment – breaking down the validation methodologies used and the results. This section is aimed at the more technically focused reader. Since, for a given mission, multiple validation studies may be performed – for example, by the mission/vendor and/or by independent assessors – there may be multiple Detailed Validation Maturity Matrices produced and reported.

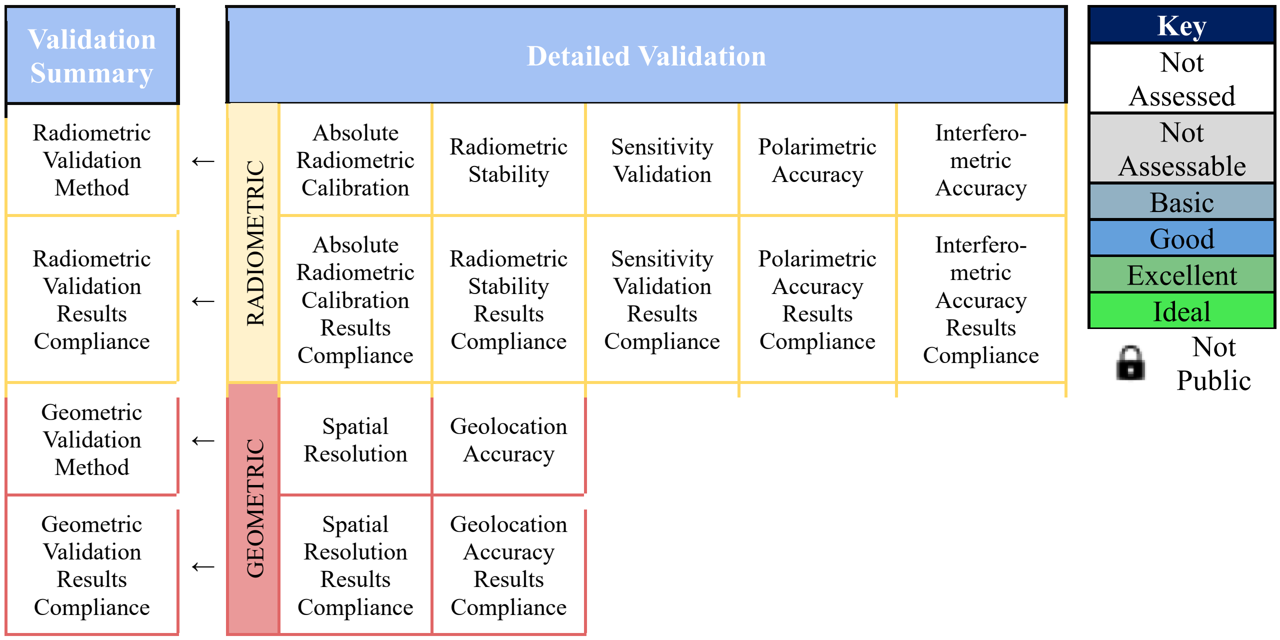

Figure 2. Validation Cal/Val Maturity Matrix for the SAR domain, showing the Validation Summary column from the Summary Cal/Val Maturity Matrix.

2.3. Approach to Grading

The assessment framework is aimed at verifying claimed mission performance and that the mission follows community best practice to an extent that is "fit for purpose". The grading criteria for each category are determined based on a logical interpretation of this principle. For example, pre-launch calibration quality grading is based on the comprehensiveness of activity with respect to the target instrument performance.

Grades of Basic, Good, Excellent, or Ideal may be given. The grade Ideal is generally reserved to provide recognition for achieving the highest standard of quality with respect to community best practice. This high bar of quality may be aspirational but is the benchmark EO data providers should aim for. Note that a grade of Basic can also be considered acceptable in a given context.

Additionally, a subsection may be graded as Not Assessable where certain aspects of product quality will not be assessed – either because the mission is not yet mature enough and has insufficient information available to allow the assessment, or because it is out of scope of the assessment. Not Assessed is only used during the evaluation phase during which a grade may not yet be available.

2.4. Considerations for the SAR domain

Since the SAR domain covers a broad range of instruments, some assessment sub-sections may not be applicable to all SAR sensor types. Distinctions may be drawn in terms of acquisition modes, sensor resolution, polarimetric capability, interferometric capability etc. This complexity also applies for mission data products of different processing levels, where distinctions may be made for reconstructed data and geophysical products.

Finally, it is important to note that these guidelines do not intend to provide absolute criteria on whether any aspect of a given mission attains a given grade – often "expert judgement" is required, especially when considering what is "fit for purpose" in a given context.

3. Data Provider Documentation Review

In this section we provide detailed guidelines for Documentation Review. This assessment aims to review mission quality as evidenced by its documentation. It is divided into the follow sections:

- Product Information
- Metrology
- Product Generation

In the following we look at each of these sections in turn and discuss the grading criteria.

The results of the Documentation Review are reported on the left portion of the Summary Cal/Val Maturity Matrix. This portion is shown in Figure 3.

Data Provider Documentation Review		
Product Information	**Metrology**	**Product Generation**
Product Details	Sensor Calibration & Characterization	Image Formation & Calibration Algorithms
Availability & Accessibility	Geometric Calibration & Characterization	Geometric Processing
Product Format, Flags & Metadata	Metrological Traceability Documentation	Retrieval Algorithm
User Documentation	Uncertainty Characterization	Mission-Specific Processing
	Ancillary Data	

Figure 3. Data Provider Documentation Review Matrix

3.1. Product Information

The Product Information section covers the top-level product descriptive information, product format, and the supporting documentation. Its subsections are defined below.

3.1.1. Product Details

Certain basic descriptive information should be provided with any EO data product and is required for assessments of all mission domains. The list of this required information is as follows, with specific requirements for SAR sensors added:

- Product name
- Sensor Name
- Sensor Type
- Sensor Mode
  - Imaging mode

    Stripmap, ScanSAR, Bistatic, etc.
  - Sensor Frequency

    Center frequency of the observation, e.g., L-Band or 1.254GHz
  - Viewing Geometry (e.g., incidence angles)
- Sensor Polarization
- Mission Type

  Either single satellite or constellation of a given number of satellites.
- Mission Orbit

  For example, Sun Synchronous Orbit with Local Solar Time.
- Product version number
- Product ID
- Processing level of product
  - For SAR products - defined as (EOSDIS 2011):
    - Unprocessed Data: Unfocused radar echoes
    - Reconstructed Data: Single look complex (SLC), focused SAR imagery
    - Geophysical Observations: Geocoded imagery and higher-level products
- Measured quantity name

  Backscatter Coefficient
- Measured quantity units
- Stated measurement quality.

  To provide context to the reader for the rest of the assessment, provide the product "quality" as specified by the provider.

  For SAR sensors – this should cover both radiometric and geometric quality. In the radiometric case, quality could be given as absolute radiometric accuracy. Typically, data providers only give a single mission uncertainty value, which may even be the sensor's required accuracy from its specification.
- Spatial Resolution

  Sensor resolution (e.g., azimuth and range) in addition to ground distance
- Acquisition Time and Date
- Spatial Coverage

  Define swath or scan and pixel width, and footprint of a scene or single acquisition.
- Temporal Resolution

  Define repeat/revisit time, i.e., time between successive observations of a given location.
- Temporal Coverage

Define period of mission operation (expected if current mission), including any periods of inactivity during the mission.
- Point of contact (Responsible organization, including email address)
- Product access (e.g., URL, DOI, if applicable)
- Restrictions for access and use, if any

Table 1 shows how provision of data product information relates to the grade it achieves for this subsection of the quality assessment.

Table 1. Product Information > Product Details – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside of the scope of study.
Not Assessable	Relevant information not made available.
Basic	Many pieces of important information are missing.
Good	Some pieces of important information are missing.
Excellent	Almost all required information is available.
Ideal	All required information is available.

3.1.2. Availability & Accessibility

This section is about how readily the data are available to those who wish to use them. Does the data set follow the FAIR (Findable, Accessible, Interoperable, Reusable) Data Principles for scientific data management and stewardship (Wilkinson 2016), that provide valuable principles for all applications. These principles state that:

Data should be **findable**
- Metadata and data are assigned a globally unique and persistent identifier
- Data are described with rich metadata
- Metadata clearly and explicitly include the identifier of the data it describes
- Metadata and data are registered or indexed in a searchable resource

Data should be **accessible**
- Metadata and data are retrievable by their identifier using a standardized communication protocol
- The protocol is open, free and universally implementable
- The protocol allows for an authentication and authorization procedure where necessary

Data should be **interoperable**
- Metadata and data use a formal, accessible, shared and broadly applicable language for knowledge representation
- Metadata and data use vocabularies that themselves follow FAIR principles
- Metadata and data include qualified references to other (meta)data

Data should be **reusable**

- Metadata and data are richly described with a plurality of accurate and relevant attributes
- Metadata and data are released with a clear and accessible data usage license
- Metadata and data are associated with detailed provenance
- Metadata and data meet domain-relevant community standards

Table 2 shows how provision of the above information relates to the grade a data product achieves for this subsection of the quality assessment.

Table 2. Product Information > Availability and Accessibility – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Relevant information not made available.
Basic	The data set does not appear to be following the FAIR principles
Good	The data set meets many of the FAIR principles and/or there is an associated data management plan that shows progress toward the FAIR principles
Excellent	The data set meets many of the FAIR principles and has an associated data management plan. The data are available through an easy-to-access license.
Ideal	The data set fully meets the FAIR principles and has an associated data management plan. The data are available through an easy-to-access license.

3.1.3. Product Format, Flags and Metadata

An important aspect of EO data products that ensures ease of access to the widest variety of users is their format. Product metadata and flags offer users important extra layers of useful descriptive information in addition to the measurements themselves that can be crucial to their analysis.

In the ideal case, product format would meet the appropriate Committee on Earth Observation Satellites (CEOS) Analysis Ready Data (ARD) metadata guidelines, such as CEOS ARD for Land (CARD4L) (CEOS LSI 2020) requirements in the case of surface reflectance products.

In the case where such a standard does not exist, product format is graded based on the following:

- the extent to which it is documented,
- whether a standard file format is used (e.g. NetCDF),
- whether it complies with standard variable, flag and metadata naming conventions, such as the Climate and Forecast (CF) metadata Conventions (Eaton et al. 2003), or, for data from the European Union, the Infrastructure for Spatial Information in the European Community (INSPIRE) directive (INSPIRE Drafting Team 2013),
- whether flags and metadata provide an appropriate breadth of information.

If a product is derived from a constellation of satellites, the specific satellite it is observed by should be included in the product metadata. Table 3 shows how a given EO data product should be graded for its format.

Table 3. Product Information > Product Format, Flags and Metadata – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Non-standard, undocumented data format.
Basic	Non-standard or proprietary data format, or poorly-documented standard file format. Little useful metadata or data flags.
Good	Data exists in a documented standard file format. Non-standard naming conventions used. Reasonable set of documented metadata and data flags.
Excellent	Data is organized in a well-documented standard file format, meeting community naming convention standards. Comprehensive set of metadata and data flags.
Ideal	Analysis Ready Data standard if applicable, else as *Excellent*.

3.1.4. User Documentation

Data products include the following minimum set of documentation for users, which should be regularly updated as required:

- Product User Guide/Manual (PUG/PUM)
- Algorithm Theoretical Basis Document (ATBD)

It may be that for a given mission, a combination of articles, publications, webpages, and presentations provides a similar set of information in place of these documents. To achieve the highest grades this information should be presented as formal documents, and users should not be expected to search for this information.

The QA4ECV project provides generic guidance for the expected contents of these documents (Scanlon 2017a, [b] 2017). The user guide should provide general information on the product, including:

- Description of available products (as specified in Section 3.1.1).
- Description of how to read the products, i.e., product format and metadata.
- Contact information.
- References

More specifically for SAR sensors, the ATBD should include the following:

- Basic overview of the instrument design concept (not necessarily proprietary details), including viewing geometry.

- Description of the radiometric calibration processing, including the sensor measurement function.
- Description of the geometric processing.
- Description of the of the geophysical retrieval processing, if required
- Description of any other mission specific processing, as necessary.
- Description of the uncertainty analysis performed on this processing.
- Details of assumptions and limitations of the algorithm.

There are a variety of relevant technical details of varying significance which are important to include in such processing descriptions, for example if the product is in units of backscatter, defining the type of coefficient. The mission assessor should apply expert judgement to decide the extent to which necessary details are included.

Note that the PUG and ATBD will likely be the source of much of the information required for the other subsections of the assessment. In particular, the technical review of the fitness for purpose of the processing algorithms is undertaken in the Product Generation section of the assessment.

Table 4 describes how the product user documentation is assessed.

Table 4. Product Information > User Documentation – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	No user documentation provided or documentation out-of-date.
Basic	Limited PUG available, no ATBD. Information is up to date.
Good	Some PUG and ATBD-type information available. These may be as formal documents or from multiple sources, e.g. articles. Documentation is up to date.
Excellent	PUG meets QA4ECV standard, reasonable ATBD. Documents are up to date.
Ideal	PUG ATBD available meeting QA4ECV standard**.** Documents are up to date.

3.2. Metrology

Metrology is the science of measurement. This section covers the aspects of the mission related to measurement quality, including calibration, traceability, and uncertainty. The Metrology subsections are now defined.

3.2.1. Sensor Calibration & Characterization

The pre-flight and post-launch sensor calibration and characterization should encompass a given sensor's behavior to an extent and sufficient quality that is "fit for purpose" within the context of the mission's stated performance.

In the SAR case, the objective of calibration and characterization is to provide knowledge about the SAR acquisition and image formation system, such that this information is properly exploited to improve the product quality (e.g., remove errors or biases).

In the SAR context, errors may originate from various sources, e.g., from the SAR antenna or from the SAR electronics, including geometric errors such as pointing errors, gain and phase errors due to thermal drift or ageing, for example. Further errors may be introduced also on-ground during image formation, being generated by limited knowledge or approximations implemented by the processing algorithms.

In this context, the pre-flight characterization of the instrument is paramount, to have the best possible knowledge of the instrument behavior already before operations, and to limit the in-orbit characterization effort, which often requires specific acquisition modes or targets, conflicting with the nominal operations. Pre-flight characterization includes a set of measurements such as antenna pattern characterization in anechoic chambers, thermal response characterization in gain and phase and others. Furthermore, during the pre-flight phase, the definition, deployment, and validation of the calibration facilities that will be used in-flight need to be carried out. In this phase, specific electronic ground support equipment systems may be exploited to generate test signals or simulate acquisition scenarios.

Of course, the on-ground characterization methods have a limited capability to create the in-flight conditions of full system operations. Therefore, a residual set of activities are necessary, the majority of these being performed during the initial phases of the mission (launch and early orbit phase and commissioning phase). The main in-flight calibration and characterization activities include sensor pointing characterization, antenna model verification, radiometric bias and stability verification, among others.

Although the sensor should be well calibrated after the pre-flight and in-flight characterization activities, monitoring tasks need to be planned and performed throughout the whole mission lifetime, to correctly follow possible drift or due to ageing and seasonal effects. This brings the need to define a set of calibration facilities to be exploited during the mission lifetime. These facilities include specific circuits, for on-board calibration, and on-ground systems such as dedicated calibration targets (e.g., corner reflectors) or calibrators (e.g., transponders).

In addition to the mission-specific calibration facilities, natural targets having suitable characteristics are useful for monitoring SAR sensor stability in the long term. Some examples are:

- The tropical rainforest, due to its high canopy density, gives an homogeneous response to SAR signals, thus being useful to check the radiometric accuracy and to validate the antenna model. These areas have been mapped by the CEOS working group on calibration and validation (WGCV) SAR subgroup (http://calvalportal.ceos.org/sar_subgroup/).
- The low-backscatter areas, such as the Atlantic and Pacific Doldrums or smooth desert surfaces, are useful to check the SAR sensitivity or noise equivalent sigma zero (NESZ).
- The very coherent areas of rocky or salty deserts, such as the Atacama in Chile, are useful for SAR interferometric quality assessment.

Table 5 shows how to grade pre- and post-launch sensor calibration and characterization.

Table 5. Metrology > Sensor Calibration & Characterization – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside of the scope of study.
Not Assessable	Pre-flight and post-launch sensor calibration & characterization not documented or information not available.
Basic	Pre-flight and post-launch sensor calibration & characterization misses some important aspects of instrument behavior and/or is not entirely of a level of quality to be judged fit for purpose.
Good	Pre-flight and post-launch sensor calibration & characterization covers the most important aspects of instrument behavior at a level of quality to be judged fit for purpose.
Excellent	Pre-flight and post-launch sensor calibration & characterization covers all reasonable aspects of instrument behavior to a quality that is "fit for purpose" in terms of the mission's stated performance. Pre-flight calibration traceable to SI or community reference, characterization methods meet good practice. Post-launch Cal/Val uses appropriate community infrastructure/methods (e.g., from CEOS).
Ideal	Meets *Excellent* criteria, additionally, the calibration and characterization include the measurements needed to assess uncertainties at the component level and their impact on the final product. Post-launch Cal/Val uses appropriate community infrastructure/methods traceable to SI.

3.2.2. Geometric Calibration & Characterization

As for sensor calibration and characterization, geometric calibration and characterization, pre-flight and on-orbit, should encompass a given sensor's behavior to an extent and sufficient quality that is "fit for purpose" within the context of the mission's stated performance.

Pre-launch includes the calibration and characterization of the geometric aspects of the sensor, such as antenna pattern, as well as other components of the satellite that influence the geometric processing should be characterized, such as star trackers or attitude control systems. Post-launch relevant performance parameters should be temporally monitored.

This is specific to given instrument types and calibration methods and will require a degree of expert judgement. However, for post-launch calibration and characterization where SI-traceable test-sites are available, these should be used.

For the SAR case, the geometric calibration aims at improving the capability of the correct identification of the position on the Earth surface of a specific pixel in the SAR image (geolocalization). The geolocalization can be affected by systematic errors such as delays introduced by the instrument electronics, errors in reconstructing the sensor's trajectory, errors introduced by approximations made in the image-formation algorithms implemented on the ground. In addition, external factors such as unknown phase delays in signal path, plate tectonics and solid earth tides introduced by the propagation media, impact the overall geolocation accuracy.

To assess the geolocation performance and identify possible systematic components, targets with known coordinates are exploited as reference points. These targets may be man-made targets such as the already mentioned corner reflectors or transponders, or natural targets. The geolocation assessment methodology is described in (Freeman 1992).

Table 6 shows how the geometric calibration and characterization is graded within the assessment framework.

Table 6. Metrology > Geometric Calibration & Characterization – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Geometric calibration & characterization not documented or not available.
Basic	Geometric calibration & characterization misses some important aspects of instrument behavior and/or is not entirely of a level of quality to be judged “fit for purpose”.
Good	Geometric calibration & characterization covers most important aspects of instrument behavior at a level of quality to be judged “fit for purpose”.
Excellent	Geometric calibration & characterization covers all reasonable aspects of instrument behavior to a quality that is “fit for purpose” in terms of the mission’s stated performance. Post-launch characterization uses appropriate community infrastructure/methods (e.g., from CEOS).
Ideal	In addition to meeting *Excellent* criteria, geometric calibration and characterization includes the measurements needed to assess uncertainties at the component level and their impact on the final product. The quality is “fit for purpose” in terms of the mission’s stated performance and meets the science-user expectations.

3.2.3. Metrological Traceability Documentation

Traceability is defined in the vocabulary of metrology (VIM) (JCGM 2012) as the

“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”

Traceability is therefore a key aspect of achieving reliable, defensible measurements. In this definition, an important part of measurement traceability is that it is well documented. Various diagrammatic approaches have been developed to present the traceability chains for EO data products (e.g., the QA4ECV guidance includes a traceability chain drawing tool (Scanlon 2017c)). A traceability diagram should be included in the documentation for every EO mission. Guidance for a detailed measurement function centered on “uncertainty tree diagram”, more suitable for reconstructed data (and some geophysical products) processing and should be the aspiration for missions in the future (Mittaz, Merchant, and Woolliams 2019).

It is important that traceability documentation remains up to date. It is common that aspects of a sensor's calibration may be modified or completely changed over the course of a mission, which changes the sensor's traceability chain, and such updates should be documented.

Table 7 shows how the metrological traceability documentation is graded, based on its completeness.

Table 7. Metrology > Metrological Traceability Documentation – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	No traceability chain documented.
Basic	Traceability chain diagram and/or uncertainty tree diagram included, missing some important steps.
Good	Traceability chain and/or uncertainty tree diagram documented identifying most important steps and sources of uncertainty.
Excellent	Rigorous uncertainty tree diagram, a traceability chain documented, identifying all reasonable steps of and accompanying sources of uncertainty.
Ideal	Rigorous uncertainty tree diagram and traceability chain documented, identifying all reasonable steps and accompanying sources of uncertainty. Establishes traceability to SI.

3.2.4. Uncertainty Characterization

To ensure measurements are both meaningful and defensible, it is crucial that they come with rigorously evaluated uncertainty estimates. A comprehensive description of how to evaluate sources of uncertainty in a measurement, and propagate them to a total uncertainty of the final measurement and, is provided by the metrological community in the Guide to the Expression of Uncertainty in Measurement (GUM) (JCGM 2008). The GUM approach should be applied to all EO missions.

The application of Earth Observation metrology has progressed greatly in recent years. Increasingly, providers of operational and reprocessed data products are applying different approaches to evaluate and distribute metrologically rigorous error-covariance information for L1 and L2 products at the per pixel level, as required by climate studies. For example, ESA's Sentinel-2 mission has developed an on-the-fly, pixel-level uncertainty evaluation tool (Gorroño et al. 2017). There have also been some initiatives, like the previously mentioned FIDUCEO project, that have applied metrology to historical sensor data records (Taylor et al. 2019).

With that said, it is typical for uncertainties (or performance estimates) to be evaluated in a manner that does not comply with the GUM, for example, the performance specification value or single offset from a comparison sensor may be quoted as the uncertainty.

For the SAR case, the uncertainty in the radiometric measurement is provided by the key performance figures of the Absolute radiometric accuracy and stability. The geometric uncertainty is quantified instead by the geolocation accuracy.

Table 8 shows the uncertainty characterization grading under the assessment framework.

Table 8. Metrology > Uncertainty Characterization – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	No uncertainty information provided.
Basic	Uncertainty established by limited comparison to measurements by other sensor/s.
Good	Limited use of GUM approach, and/or, an expanded comparison to measurements by other sensors. Most important sources of uncertainty are included.
Excellent	Full GUM approach is used to estimate measurement uncertainty, all important sources of uncertainty included. Uncertainty per pixel provided.
Ideal	Full GUM approach is used to estimate measurement uncertainty, including a treatment of error-covariance. Per pixel uncertainties are provided for appropriate error correlation components, e.g., random systematic.

3.2.5. Ancillary Data

Throughout the processing chain there may be a requirement for external input data, for example, a digital elevation model or reference data for algorithm tuning. The ancillary datasets used during the processing should be identified to the user, where possible due to commercial sensitivity. Ideally this should be traceable on a per product level.

Ancillary datasets must be of a sufficient quality, including the application of suitably rigorous metrology, for example, in the form of SI traceability.

The suitability of the ancillary data for its application must also be considered, with respect to the mission's stated performance requirements. For example, the quality, size, and representativeness of algorithm input data (in terms of factors like surface type). The requirements will be specific to the retrieval method used and may require some expert judgement.

Table 9 shows how the ancillary data are graded under the assessment framework.

Table 9. Metrology > Ancillary Data – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Use of ancillary data undocumented.
Basic	Ancillary data used in product generation, specified to some extent, though incomplete. Not entirely of a sufficient quality to be judged "fit for purpose" in terms of the mission's stated performance.
Good	Ancillary data used in product generation, specified, though not necessarily on a per product basis. Mostly of a sufficient quality to be judged "fit for purpose" in terms of the mission's stated performance.
Excellent	Ancillary data used in product generation, fully specified per product and traceable. Ancillary data used are of sufficient quality to be judged "fit for purpose" in terms of the mission's stated performance.
Ideal	Ancillary data used in product generation, meets the Excellent criteria, and are traceable to SI where appropriate.

3.3. Product Generation

The Product Generation section covers the processing steps undertaken to produce the data product. This starts with an assessment of the application of image formation and calibration of the instrument measurements to L1. If the mission under assessment produces a L2 data product, then additional steps of assessment must be undertaken.

3.3.1. Image Formation and Calibration Algorithm

The applied L1 image formation and calibration algorithm should be of a sufficient quality that is "fit for purpose" within the context of the mission's stated performance across all stated use cases and scene types (e.g., land, ocean, etc.). What this requires is specific to the sensor-domain and will require a degree of expert judgement. This should be based on the same reasoning applied to the pre-launch and in-flight calibration assessment and reviewed based on the ATBD.

For the SAR case, a specific note needs to be mentioned for the image formation process. A large variety of image formation algorithms (also called focusing algorithms) have been presented in literature. Typically, the focusing algorithm is tailored to the specific acquisition mode (e.g., stripmap or spotlight), the specific geometry (e.g., monostatic or bistatic) and the specific instrument features, such as number of antennas, bandwidth, operating frequency, and many others. Additional image formation steps may include filtering to reach the specified radiometric and geometric resolution (e.g., multi-looking).

Generally, the calibration algorithm is also integrated, and complementary, to the image formation algorithm to generate the final SAR image. Calibration steps include the compensation of a sensor's specific biases, such as power variations during the acquisition, or the spatially and temporally variant gain applied by the SAR antenna.

Table 10 shows how the image formation and calibration algorithm are graded.

Table 10. Product Generation > Image formation and radiometric Calibration Algorithm – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Image formation and calibration algorithm not documented.
Basic	Image formation and calibration algorithm somewhat documented. Calibration algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance.
Good	Image formation and calibration algorithm documented. Reasonable retrieval algorithm used, judged "fit for purpose" in terms of the mission's stated performance for most expected use cases.
Excellent	Image formation and calibration algorithm documented. Calibration used "fit for purpose" in terms of the mission's stated performance in all expected use cases.
Ideal	Image formation and calibration algorithm well-documented. State-of-the-art calibration algorithm used, easily "fit for purpose" in terms of the mission's stated performance.

3.3.2. Geometric Processing

Different geometric processing methodologies may be applied to remote sensing data depending on the domain and application of the data product. The applied geometric processing should be of a sufficient quality that is "fit for purpose" within the context of the mission's stated performance for all mission products. Again, this constitutes a technical review of the ATBD from the data provider.

In the SAR case, the geometric processing typically considers the projection of the SAR image pixels, from the native radar coordinates to geographic coordinates. This process includes the modelling of the radar acquisition geometry and the proper interpolation of the image on an Earth referenced grid, by means of an Earth model, including the usage of a digital elevation model. In addition, the geometric processing includes the computation of the scaling factors to convert the image intensity according to the physical backscatter quantity (e.g. $sigma_0$, $beta_0$, $gamma_0$, see e.g. (Raney et al. 1994; Small 2011).

Table 11 shows how the geometric processing is graded.

Table 11. Product Generation > Geometric Processing – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Geometric processing not fully documented.
Basic	Geometric processing documented. Missing all or part of the calibration parameters. Calibration algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance. Confidence in the calibration quality is minimal.
Good	Geometric processing documented. Missing part of the input calibration parameters. Reasonable retrieval algorithm used. Confidence in the calibration quality is considered sufficient.
Excellent	Geometric processing documented. All input calibration parameters exist. Methodology used is considered "fit for purpose" in terms of the mission's stated performance for all expected use cases. Quality flags indicate good geometric accuracy with less than 5% exceptional.
Ideal	Geometric processing well-documented. State-of-the-art methodology used, easily "fit for purpose" in terms of the mission's stated performance. Quality flags indicate excellent geometric accuracy.

3.3.3. Higher Level Product Generation

For many types of higher level (e.g., geophysical) products there are typically a variety of potential retrieval methods that may be used to derive them. These may vary in ways such as model complexity and computational efficiency – resulting in higher or lower quality final products.

As with the sensor calibration, the higher-level method should be of a sufficient quality that is "fit for purpose" within the context of the mission's stated performance across all stated use cases (e.g. scene types). What this requires is specific to a given variable's retrieval methods and will require a degree of expert judgement.

Table 12 shows how the retrieval algorithm used to generate higher level products is graded.

Table 12. Higher Level Product Generation > Retrieval Algorithm – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Retrieval algorithm not documented.
Basic	Retrieval algorithm somewhat documented. Retrieval algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance.
Good	Retrieval algorithm documented. Reasonable retrieval algorithm used, judged "fit for purpose" in terms of the mission's stated performance for most expected use cases, with at least a sensitivity analysis carried out.
Excellent	Retrieval algorithm documented. Retrieval algorithm "fit for purpose" in terms of the mission's stated performance all expected use cases and validated performance against similar algorithms or with empirical evidence.
Ideal	Retrieval algorithm documented. State-of-the-art retrieval, easily "fit for purpose" in terms of the mission's stated performance, full uncertainty budget derived and validated.

3.3.4. Mission Specific Processing

Additional processing steps are separate from the main sensor calibration or retrieval processing. These may include processes like the generation of classification masks. Additional processing steps must themselves be assessed for quality based on their "fitness for purpose" in the context of the mission.

For the SAR case, the mission specific processing may be related to combination of multiple acquisitions to generate specific products, for instance the co-registration to obtain a stack of co-registered images for interferometric analyses.

Each additional processing step should be separately assessed and based on the criteria described in Table 13, and then a combined score determined.

Table 13. Product Generation > Mission Specific Processing – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	Additional processing steps not documented.
Basic	Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.
Good	Additional processing steps documented. All significant additional processing steps are fit for stated purpose.
Excellent	Additional processing steps documented. All additional processes steps are fit for stated purpose.
Ideal	All additional processing steps are fully documented and considered state-of-the-art.

4. Detailed Validation

In this section we provide guidelines for a *Detailed Validation* assessment. The overall goal here is to verify that the mission performance is consistent with the sensor stated performance.

The detailed validation assessment is broadly divided into radiometric and geometric validation activities. Within the two sections are paired sub-sections describing the performance metrics, each of which are evaluated both in terms of the quality of the validation method used and the validation results compliance.

The results are reported as part of the *Detailed Validation Cal/Val Maturity Matrix* (Figure 2) and are then summarized across all performance metrics in the Validation Summary column on the left. This Validation Summary is the same summary shown in the last column in the *Summary Cal/Val Maturity Matrix* shown in Figure 1.

The remainder of this section includes:

- The criteria for grading the quality of the validation method used and validation results compliance is given in Section 4.1.
- The Radiometric and Geometric performance metrics to be assessed are described in Section 4.2.
- Finally, in Section 4.3 the approach for synthesizing the results of the Detailed Validation into the Validation Summary column is described.

4.1. Detailed Validation Grading Criteria

This section describes how, in generic terms, the criteria for grading the quality of the Validation Method and Validation Results Compliance subsections of the Radiometric and Geometric performance metrics.

4.1.1. Validation Method

Generally, satellite validation attempts to demonstrate compliance of mission data products with respect to some claimed performance level (e.g., documented specifications) by comparison with independent reference data. A metrologically-rigorous validation of measurements goes a step further, attempting to verify both the satellite measurements and their associated uncertainties. Validated uncertainties provide evidence of the credibility of the uncertainty estimate given. Commonly used metrics such as the statistical spread of differences may be used to estimate the uncertainty, however this often may not provide a realistic estimate of the actual uncertainty.

A rigorous validation must compare mission data products with independent reference data that are fully representative of the satellite measurements being validated (e.g., scaling considerations), over the full extent of measurements the satellite may make (e.g., biomes, dynamic range, seasonal variation). This may require the use of a variety of different reference datasets to cover different observation conditions.

In the same way, these guidelines describe how to assess the quality of satellite mission data. Similar considerations must be made for the quality of reference data used to validate the satellite mission data. The highest quality validation reference data provide uncertainty-assessed validation reference data traceable to SI, and come from activities, such as the ESA Fiducial Reference Measurement (FRM) projects (Fox 2019).

Table 14 shows how the validation methods are graded. The specific interpretation of these criteria in the quality assessment of a particular validation activity depends on several factors, for example the method used or the sensor target performance, therefore some level of expert judgement may be required when determining the grading.

Table 14. Validation > Validation Method – Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	No validation activity performed.
Basic	Methodology is simple comparison, covering a limited range of satellite measurements. Uncertainty information not available for reference data.
Good	Methodology covers a range of satellite measurements that represents typical use cases, using representative reference measurements. Uncertainty information not available for reference data.
Excellent	Methodology assesses satellite measurements and reference data with respect to their characterized uncertainties. Reference measurements are assessed to be well representative of the satellite measurements.
Ideal	Methodology assesses satellite measurements and reference data with respect to their error-covariance and attempts to validate those uncertainties. Reference measurements independently assessed to be fully representative of the satellite measurements.

4.1.2. Validation Results Compliance

This section assesses the actual results of the validation activities themselves. In the best case these will show both validated satellite measurements and their associated uncertainties and will have been obtained by a group independent of the satellite data provider.

The results should be documented in a Validation report from a user community, see the QA4ECV guidance for expected content (Scanlon 2017d).

Grading for this subsection is based on the compliance of the validation results with the data performance claimed by the data provider and with the possibly more stringent standards from the user community.

Table 15 how the validation results are graded within the assessment framework.

Table 15. Validation > Validation Results –Assessment Criteria

Grade	Criteria
Not Assessed	Assessment outside the scope of study.
Not Assessable	No validation activity performed.
Basic	Claimed mission performance shows some agreement with validation results.
Good	Claimed mission performance shows good agreement with validation results.
Excellent	Claimed mission performance shows excellent agreement with validation results. Analysis performed independently of the satellite mission owner.
Ideal	Claimed mission performance shows excellent agreement with validation results, measurement uncertainties also validated. Analysis performed independently of the satellite mission owner.

4.2. Performance Metrics

This section describes the performance metrics that define the *Detailed Validation Cal/Val Maturity Matrix* structure. This is divided into the Measurement and Geometric sections.

4.2.1. Measurement Validation

Different SAR satellite missions are aimed at a broad range of applications and thus are subject to various design and performance trade-offs in order to meet their goals. The domain spans a large variety of solutions, ranging from small satellites operating in formation (Distributed SAR (Kraus et al. 2019)) to larger satellites often operated in a constellation. The performance characteristics of these different types of missions may in general be very different. Here we assess their compliance with their claimed performance, which in absolute terms is mission/application specific.

The performance metrics are defined to characterize different aspects of measurement integrity, which may be of different relative importance depending on the intended application. For data products intended for quantitative analysis, the validation of calibration is clearly necessary to provide credibility to the measurements. For temporal analysis, calibration stability of the data record must be demonstrated. Finally, low measurement noise performance may be important for data where instantaneous images are analyzed, but less important in long term data where it will tend to average out.

For the *Measurement Validation* section, the following metrics are used to validate SAR satellite sensors:

- Absolute Radiometric accuracy validation
- Radiometric Stability validation
- Sensitivity validation
- Polarimetric accuracy
- Interferometric quality validation

4.2.1.1. Absolute Radiometric Accuracy Validation

For a distributed target, the absolute radiometric accuracy is defined as the uncertainty resulting from the measurement of the reflectivity ($sigma_0$) of a uniform and invariant target situated anywhere within the operating dynamic range of the system, anywhere in the swath, assuming that the uncertainty in measurement of the $sigma_0$ is zero (i.e., speckle is ignored). Typically, the validation of the absolute radiometric accuracy for a distributed target involves the acquisition of SAR images over a homogeneous target suitable for the sensor considered, e.g., the Amazonian rainforest.

For a point target, the absolute radiometric accuracy measures the uncertainty in measurement of the radar cross section (RCS), considering an invariant and well-known ground target. Typically, the validation of the absolute radiometric accuracy involves the acquisition of SAR images over proper calibration targets such as corner reflectors or active transponders.

Absolute radiometric calibration involves antenna pattern gain correction, incidence angle correction, range spreading loss correction and an absolute calibration constant. The absolute calibration constant, and the antenna pattern, should be provided by the data provider in the metadata. Similarly, datasets containing multiple channels (e.g., polarizations) will likely require separate calibration parameters. These assessments should be conducted per methods presented in (Shimada et al. 2009). The absolute radiometric accuracy validation also considers the accuracy of $beta_0$, $gamma_0$, and radiometrically terrain corrected (Small 2011) imagery.

4.2.1.2. Radiometric Stability Validation

Radiometric stability is defined as the deviation resulting from repeated independent measurements of the reflectivity ($sigma_0$) or the RCS of a stable target. This target can be situated anywhere within the system dynamic range, and swath, assuming that the uncertainty in measurement of the $sigma_0$ is zero (i.e., speckle is ignored). As for the absolute radiometric accuracy, typically the Amazonian rainforest and dedicated calibration targets are exploited to generate the measurement time series on which the deviation is computed.

Assessment of radiometric stability will be handled through tracking of known target amplitude over time, for various targets across the image swath (Schmidt, Ramon, and Schwerdt 2018).

4.2.1.3. Sensitivity Validation

The sensitivity is defined as the backscatter value that corresponds to a signal-to-noise ratio equal to 1 (0 dB) in the SAR image. The sensitivity is also called Noise Equivalent Sigma Zero (a.k.a. noise equivalent sigma nought, NES0), and is an important performance parameter for any SAR instrument. The validation of the NESZ is generally performed by acquiring data over very low or null backscatter targets, such as calm water or deserts. The sensitivity validation is important for validation of the specifications provided by the system manufacturer, but also to check that the noise is correctly modelled by the SAR processor, resulting in a radiometrically compensated image (*de-noising*).

NESZ can be estimated by sensor parameters and can be estimated using signal-free regions (Calabrese and Episcopo 2014) or through modeling of interferometric pairs (Leanza et al. 2018).

4.2.1.4. Polarimetric Accuracy Validation

The SAR sensor may transmit and receive electromagnetic waves in one or more polarizations. Linear polarizations are often used, in which case the transmit or receive polarizations are either horizontally polarized (H) or vertically polarized (V). The corresponding SAR image therefore contains different information depending on the combination of the transmit and receive polarizations. Typical classes of SAR sensors are “single-pol”, “dual-pol”, or “full-pol” (a.k.a. quad-pol). Single-pol is when one polarization is used in transmit and receive. Dual-pol is when one polarization is used in transmit and two simultaneously are recorded in receive (e.g., dual-V: VV and VH or dual-H: HV and HH). Full-pol means that the full set of combinations are available (e.g., four combinations VV, VH, HV, HH). Circular polarization transmission and H and V polarization reception is also employed and referred to as “compact-pol”. The polarimetric accuracy quantifies the capability of the SAR image to correctly measure the backscatter in a certain polarization and characterizes the cross-talk between polarimetric channels. The definition of the polarimetric accuracy can be borrowed from the definition of the absolute radiometric accuracy, considering the particular polarization combination (VV, VH, HV or HH as mentioned). Polarization accuracy can be measured using corner reflectors if quad polarization imagery is available (van Zyl 1990).

4.2.1.5. Interferometric Accuracy Validation

In addition to backscatter intensity, SAR images also carry information about the sensor-target distance in the phase of each pixel. While the absolute distance from a single SAR image is hardly exploitable, the differential distance between two subsequent passes is widely exploited in SAR interferometry to assess surface deformations occurring between two passes. The aim of interferometric quality validation is to assess the capability of the SAR system to provide an interferogram with a limited amount of phase aberrations introduced by the instrument or by the processing chain. Highly stable targets such as rocky or salt flats are suitable as they should ideally yield a perfectly “flat” interferogram (i.e., no phase differences should be measured). Any deviations from this ideal interferogram are considered as artefacts generated by the SAR imaging system.

Assessment on interferometric quality will be handled through tracking of point target phase over time for various targets across the image swath (Marinkovic, Ketelaar, and van Leijen 2007).

4.2.2. Geometric Validation

Geometric performance assessment of SAR remote sensing data includes three (3) major aspects: 1) instrument Impulse Response Function (IRF); 2) geolocation accuracy on the Earth’s surface, or absolute position accuracy (APA); (3) polarimetric channel co-registration. In geometric assessment, it is also important to consider temporal stability and global consistency in all aspects (Shimada et al. 2009).

For geometric assessment, first it is important whether the data are provided in a swath or gridded format. Swath data products have not been resampled and have the original time-tagged observations as sampled by the instrument. Gridded products typically contain observations that have been resampled to a fixed Earth grid with a fixed pixel interval and may be orthorectified to correct for terrain distortions.

Swath products must be accompanied by additional information regarding geometry of the observations in the product, either within the product or as a separate geolocation product. This additional information usually includes time-tagged geodetic latitude and longitude of each observation (sample or pixel), and for many data sets, the terrain height. It may also include information such as, the flight direction and incidence angles, quality flags, satellite position and its velocity and attitude. This data may be available for each observation or at a coarser time resolution, e.g., at the scene start, center and end. For multi-channel instruments there may be additional information about relative alignment of the individual channels.

Gridded products are typically provided as scenes (or tiles) and may be accompanied by additional information such as acquisition time, incidence angle, and digital elevation model used to process the data. This information may be provided at a coarser resolution than the product resolution.

For *Geometric Validation* of satellite imagery, we define the following metrics used for evaluation:

- Impulse Response Function validation
- Geolocation accuracy validation

The impulse response function provides the accuracy with which the geolocation can be determined. In other words, the precision with which the main lobe of a point target can be estimated provides the maximum achievable geolocation accuracy.

The *geolocation accuracy validation* assessment combines geometric specification and the uncertainty criteria in one evaluation matrix for each metric. This is achieved by comparing the precise position of known targets with the measured target locations from the imagery.

4.2.2.1. Spatial Resolution Validation

The Impulse Response Function of a SAR system is the intensity and phase signal that is obtained in the SAR image when the SAR instrument acquires a point-like target, with sufficient RCS such that the noise can be neglected. Of interest is both the shape of the IRF, in terms of main lobe and side lobes, and the localization. The spatial resolution of the SAR image can be assessed from the width of the main lobe. The ratio between the side lobes and the main lobe power gives an indication of the influence of very bright point targets on the response of neighboring lower RCS targets. An image acquired over calibration point targets, such as corner reflectors or transponders, are often used for IRF assessment. In order to assess the localization accuracy, the IRF obtained over a point target with known coordinates is exploited. The spatial resolution in radar coordinates will be measured using the 3 dB width of the IRF in azimuth and range directions. Further assessment will be conducted through the analysis of peak side lobe ratio (PSLR) and integrated sidelobe ratio (ISLR) (Shimada et al. 2009; Freeman 1992).

4.2.2.2. Geolocation Accuracy Validation

The geolocation accuracy is defined as the deviation between the position of a target in the geocoded SAR image and its actual position. An ideal point-like target with perfectly known coordinates is considered. Typically, the same target that is used for IRF assessment can be exploited for geolocation accuracy validation, if the coordinates are known with an accuracy higher than the one achievable by the SAR (i.e., the coordinates have been measured for a stable point target during calibration campaigns). The effects of the propagation path (atmosphere) can have an impact on the geolocation accuracy and therefore must be accounted for in the assessment (Schubert et al. 2015).

4.3. Validation Summary

The *Validation Summary* provides a synthesis of the per performance metric assessments provided in the *Detailed Validation Cal/Val Maturity Matrix* (Figure 2). It is also presented as part of the *Summary Cal/Val Maturity Matrix*.

Each row in the *Detailed Validation Cal/Val Maturity Matrix* is represented by one cell in the *Validation Summary* column. Thus, there are four summary cells in total – *Measurement Validation Method*, *Measurement Validation Results Compliance*, *Geometric Validation Method* and *Geometric Validation Results Compliance*.

The grade for each of these summary cells represents a combination of the grades of the contributing cells. The approach is to effectively average the grades of the contributing cells, where each grade is valued as follows: Basic is 1, Good is 2, Excellent is 3, and Ideal is 4.

5. References

Calabrese, D., and R. Episcopo. 2014. “Derivation of the SAR Noise Equivalent Sigma Nought.” In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar*, 1–4. ieeexplore.ieee.org.

CEOS LSI. 2020. “Product Family Specification Surface Reflectance.” CEOS. https://ceos.org/ard/files/PFS/SR/v5.0/CARD4L_Product_Family_Specification_Surface_Reflectance-v5.0.pdf.

Eaton, Brian, Jonathan Gregory, Bob Drach, Karl Taylor, Steve Hankin, Jon Blower, John Caron, et al. 2003. “NetCDF Climate and Forecast (CF) Metadata Conventions.” cfconventions.org. 2003. https://cfconventions.org/latest.html.

EOSDIS. 2011. “Data Processing Levels for EOSDIS Data Products.” NASA Earth Science. June 23, 2011. https://science.nasa.gov/earth-science/earth-science-data/data-processing-levels-for-eosdis-data-products.

Fox, Nigel. 2019. “D-180 FRM4STS Final Report.” frm4sts.org. 2019. http://www.frm4sts.org/wp-content/uploads/sites/3/2020/01/OFE-D-180-V1-Iss-1-Ver-1-signed.pdf.

Freeman, A. 1992. “SAR Calibration: An Overview.” *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society* 30 (6): 1107–21.

Gorroño, Javier, Norman Fomferra, Marco Peters, Ferran Gascon, Craig I. Underwood, Nigel P. Fox, Grit Kirches, and Carsten Brockmann. 2017. “A Radiometric Uncertainty Tool for the Sentinel 2 Mission.” *Remote Sensing* 9 (2): 178.

Hunt, Samuel. 2021. "Earth Observation Mission Quality Assessment Framework." EDAP.REP.001. ESA. https://earth.esa.int/eogateway/documents/20142/37627/EDAP.REP.001_2.0%20-%20Mission%20Quality%20Assessment%20Guidelines.pdf.

INSPIRE Drafting Team. 2013. "INSPIRE Metadata Implementing Rules: Technical Guidelines Based on EN ISO 19115 and EN ISO 19119." MD_IR_and_ISO_20131029. European Commission Joint Research Centre. https://inspire.ec.europa.eu/documents/inspire-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1.

JCGM. 2008. "Evaluation of Measurement Data—guide to the Expression of Uncertainty in Measurement." Vol. 100. Joint Committee for Guides in Metrology. https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6.

———. 2012. "International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)." 200. Joint Committee for Guides in Metrology.

Kraus, Thomas, Gerhard Krieger, Markus Bachmann, and Alberto Moreira. 2019. "Spaceborne Demonstration of Distributed SAR Imaging with TerraSAR-X and TanDEM-X." *IEEE Geoscience and Remote Sensing Letters* 16 (11): 1731–35.

Leanza, A., A. Monti Guarnieri, D. Giudici, A. Recchia, and S. Mancon. 2018. "A Novel Method for Noise Equivalent Sigma Nought Estimation." In *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*, 1–5.

Mannan, Rubinder, Kevin Halsall, Clément Albinet, Giuseppe Ottavianelli, Philippe Goryl, Valentina Boccia, Andrea Melchiorre, et al. 2019. "ESA's Earthnet Data Assessment Pilot: Paving the Way for New Space Players." In *Sensors, Systems, and Next-Generation Satellites XXIII*, 11151:111511X. International Society for Optics and Photonics.

Marinkovic, P., G. Ketelaar, and F. van Leijen. 2007. "InSAR Quality Control: Analysis of Five Years of Corner Reflector Time Series." *Proceedings of Fringe*. https://www.researchgate.net/profile/Ramon-Hanssen/publication/229001941_InSAR_quality_control_Analysis_of_five_years_of_corner_reflector_time_series/links/00b7d517e380e023c6000000/InSAR-quality-control-Analysis-of-five-years-of-corner-reflector-time-series.pdf.

Mittaz, Jonathan, Christopher J. Merchant, and Emma R. Woolliams. 2019. "Applying Principles of Metrology to Historical Earth Observations from Satellites." *Metrologia* 56 (3): 032002.

NASA Earth Science Division. 2020. "Commercial SmallSat Data Acquisition Program Pilot Evaluation Report." NASA. https://cdn.earthdata.nasa.gov/conduit/upload/14180/CSDAPEvaluationReport_Apr20.pdf.

Nightingale, Joanne, Jonathan P. D. Mittaz, Sarah Douglas, Dick Dee, James Ryder, Michael Taylor, Christopher Old, et al. 2019. "Ten Priority Science Gaps in Assessing Climate Data Record Quality." *Remote Sensing* 11 (8): 986.

QA4EO Task Team. 2009. "A Quality Assurance Framework for Earth Observation: Operational Guidelines." 3.0. QA4EO. http://qa4eo.org/docs/Guidelines_Framework_v3.0.pdf.

Raney, R. K., T. Freeman, R. W. Hawkins, and R. Bamler. 1994. "A Plea for Radar Brightness." In *Proceedings of IGARSS '94 - 1994 IEEE International Geoscience and Remote Sensing Symposium*, 2:1090–92 vol.2. ieeexplore.ieee.org.

Scanlon, T. 2017a. "QA4ECV Product Documentation Guidance: Algorithm Theoretical Basis Document." QA4ECV. http://www.qa4ecv.eu/sites/default/files/QA4ECV ATBD Guidance.pdf.

———. 2017b. "QA4ECV Product Documentation Guidance: Product User Manual." QA4ECV. http://www.qa4ecv.eu/sites/default/files/QA4ECV PUM Guidance.pdf.

———. 2017c. "QA4ECV Product Documentation Guidance: Provenance Traceability Chains." QA4ECV. http://www.qa4ecv.eu/sites/default/files/QA4ECV Traceability Chains Guidance.pdf.

———. 2017d. "QA4ECV Product Documentation Guidance: Validation and Intercomparison Report." QA4ECV. http://www.qa4ecv.eu/sites/default/files/QA4ECV%20Validation%20Guidance.pdf.

Schmidt, Kersten, Núria Tous Ramon, and Marco Schwerdt. 2018. "Radiometric Accuracy and Stability of Sentinel-1A Determined Using Point Targets." *International Journal of Microwave and Wireless Technologies* 10 (5-6): 538–46.

Schubert, Adrian, David Small, Nuno Miranda, Dirk Geudtner, and Erich Meier. 2015. "Sentinel-1A Product Geolocation Accuracy: Commissioning Phase Results." *Remote Sensing* 7 (7): 9431–49.

Shimada, M., O. Isoguchi, T. Tadono, and K. Isono. 2009. "PALSAR Radiometric and Geometric Calibration." *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society* 47 (12): 3915–32.

Small, D. 2011. "Flattening Gamma: Radiometric Terrain Correction for SAR Imagery." *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society* 49 (8): 3081–93.

Taylor, Michael, Jon Mittaz, Marine Desmons, and Emma Wooliams. 2019. "FIDUCEO Project: (AVHRR): Report on the AVHRR FCDR Uncertainty," August. http://cedadocs.ceda.ac.uk/id/eprint/1405.

Wilkinson, Mark D. 2016. "The FAIR Guiding Principle for Scientific Data Management and Stewardship:Comment." 2. Guides. Arrow@TU-Dublin. https://arrow.tudublin.ie/dataguide/2/.

Zyl, J. J. van. 1990. "Calibration of Polarimetric Radar Images Using Only Image Parameters and Trihedral Corner Reflector Responses." *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society* 28 (3): 337–48.