INTERNATIONAL HYDROGRAPHIC ORGANIZATION

IHO GUIDELINES FOR CREATING S-100 PRODUCT SPECIFICATIONS

IHO Publication S-97

Edition 1.0.0 – Xxxx 2019

PART C

DATA QUALITY

Error! Reference source not found.

Document History

Changes to this Specification are coordinated by the IHO S-100 Working Group. New editions will be made available via the IHO web site. Maintenance of the Specification shall conform to IHO Resolution 2/2007 (as revised).

Version Number	Date	Author	Purpose
--	09 May 2018	R. Broekman	Data Quality Checklist, Draft 0.3, IHO DQWG
0.2	20 Aug 2018	RM; EM	First draft of S-97 Part C prepared from DQWG Data Quality Checklist draft 0.3
1.0.0 RC1	12 Mar 2019	RM; JW	DQWG revisions; IHO Secretariat styling; S-100 WG4 decisions; references updated following publication of S-100 4.0.0
1.0.0	Xxx 2019	S-100WG	First Edition.

Page intentionally left blank

<u>Contents</u>

Page intentionally left blank

Part C – Data Quality

1   Overview

This Guideline is intended for developers and maintainers of Product Specifications based on the IHO framework standard S-100 (Universal Hydrographic Data Model).

Creating an S-100 based Product Specification can be a big challenge for groups with little experience with S-100, especially since S-100 is a comprehensive framework with many details that may need to be considered for any particular Product Specification. A guide to assist development teams through the process can help significantly and decrease the time it takes to create or extend an S-100 based Product Specification.

A core aim of this guideline is to assist in creating harmonized Product Specifications that can be used in the e-Navigation eco-system. The term e-Navigation eco-system is meant to encompass all Product Specifications created for use in IMO defined e-Navigation systems, both on shore and at sea, such as ECDIS.

This guideline is intended to serve as a guide for anyone planning to develop or extend an S-100 compliant Product Specification. The guideline consists of three parts:

- Part A (a separate document) is an in-depth description of the various components of an S-100-based Product Specification;
- Part B (a separate document) describes the typical steps and activities involved in creating an S-100-based Product Specification. Part B describes the overall process, specific activities, and tasks, and includes hints for solving specific problems while the Product Specification is being developed;
- Part C (this document) describes the data quality measures deemed appropriate for use in S-100 based Product Specifications.

2   Introduction

Part C is intended to ensure that data quality aspects are addressed in an appropriate and harmonized way for all S-100 based Product Specifications.

This Data Quality Guideline can be used by HSSC Working Groups developing S-100 based Product Specifications. It provides 10 recommendations for appropriate data quality measures as deemed necessary to be used within S-100 based Product Specifications.

When drafting a Product Specification, the Data Quality Guideline will serve as a guidance document to verify if the appropriate Data Quality Elements have been included in the Product Specification. A Data Quality Element is a quantitative component documenting the quality of a dataset. The applicability of a Data Quality Element to a dataset depends on both the dataset's content and its Product Specification, the result being that all available Data Quality Elements may not be applicable to all datasets.

The place of data quality measures in dataset and exchange set metadata and the encoding of data quality in metadata is described in various ISO standards (ISO 19115, 19139, 19115-1/2/3, ISO 19157) and in S-100 Parts 4a–4c.

3   References

IG-D2.8.II.1   D2.8.II.1 INSPIRE Data Specification on Elevation – Technical Guidelines.

ISO 8211	Specification for a data descriptive file for information interchange structure implementations. ISO/IEC 8211, 1994.
ISO 19115	Geographic information – Metadata (2003). As amended by Amendment 1, 2006.
ISO 19115-1	Geographic information – Metadata – Part 1 – Fundamentals. ISO 19115-1, 2014, as amended by Amendment 1, 2018.
ISO 19115-2	Geographic information - Metadata - Part 2 – Extensions for imagery and gridded data. ISO 19115-2, 2009.
ISO 19115-3	Geographic information - Metadata - XML schema implementation for fundamental concepts. ISO/TS 19115-3, 2016.
ISO 19139	Geographic information – Metadata – XML schema implementation.
ISO 19157	Geographic information – Data Quality. ISO 19157, 2013, as amended by Amendment 1, 2018.
S-44	IHO S-44 - Standards for Hydrographic Surveys.
S-100	IHO S-100 - Universal Hydrographic Data Model, Edition 4.0.0, December 2018.
S-101	IHO S-101 - Electronic Navigational Chart Product Specification, Edition 1.0.0, December 2018.
S-102	Bathymetric Surface Product Specification, Edition 1.0.0, April 2012.
S-121	Maritime Limits and Boundaries. (Under development)

NOTE: In this document, "S-100" means S-100 Edition 4.0.0 unless a different edition is explicitly identified.

4   Terms and abbreviations

4.1   Terms

data product specification

a detailed description of a dataset or dataset series together with additional information that will enable it to be created, supplied to, and used by another party. A data product specification provides a description of hydrographic concepts and a specification for mapping the universe of discourse to a dataset. It may be used for production, sales, end-use or other purposes

data quality element

quantitative component documenting the quality of a dataset [ISO 19101:2002; S-100 4.0.0 Annex A]

NOTE: The applicability of a data quality element to a dataset depends on both the dataset's content and its product specification, the result being that all data quality elements may not be applicable to all datasets

data quality evaluation procedure

the whole of operations used in applying and reporting quality evaluation methods and their results

data quality measure

an evaluation of a data quality sub-element

data quality overview element

the non-quantitative component documenting the quality of a dataset. Information about the purpose, usage, and lineage of a dataset is non-quantitative quality information

data quality result

> a value or set of values resulting from applying a data quality measure or the outcome of evaluating the obtained value or set of values against a specified conformance quality level

data quality scope

> the extent or characteristic(s) of the data for which quality information is reported.

> NOTE: The scope for a dataset can comprise a dataset series to which the dataset belongs, the dataset itself, or a smaller grouping of data located physically within the dataset sharing common characteristics. Common characteristics can be an identified feature type, feature attribute, or feature relationship; data collection criteria; original source; or a specified geographic or temporal extent.

data quality sub-element

> a component of a data quality element describing a certain aspect of that data quality element

Metadata ISO19115

> quality report in the format specified in the relevant ISO standards (ISO 19115-1/2/3 and ISO 19157)

metaquality

> information about the reliability of data quality results [ISO 19157]

result scope

> scope of the (data quality) result [ISO 19157, adapted]

> NOTE: Result scope is a subset of the data quality scope.

standalone quality report

> free text document providing fully detailed information about data quality evaluations, results and measures used [ISO 19157:2013]

4.2   Abbreviations

DQWG	IHO Data Quality Working Group
ECDIS	Electronic Chart Display and Information System
ECS	Electronic Chart System
ENC	Electronic Navigational Chart
GI	Geospatial Information (generally followed by "registry" or "register," meaning either the IHO Geospatial Information Registry or a specific register in it).
GML	Geography Markup Language
GSD	Ground Sampling Distance
HDF-5	Hierarchical Data Format Version 5
IEC	International Electrotechnical Commission
IHO	International Hydrographic Organization
IMO	International Maritime Organization
ISO	International Organization for Standardization
PS	Product Specification
RMSE	Root Mean Square Error

SD   Standard Deviation

XML   eXtensible Markup Language

5   Overview of data quality measures

The relations between data quality terms are presented in Figure 5-1. The terms have been defined in clause 4.1.
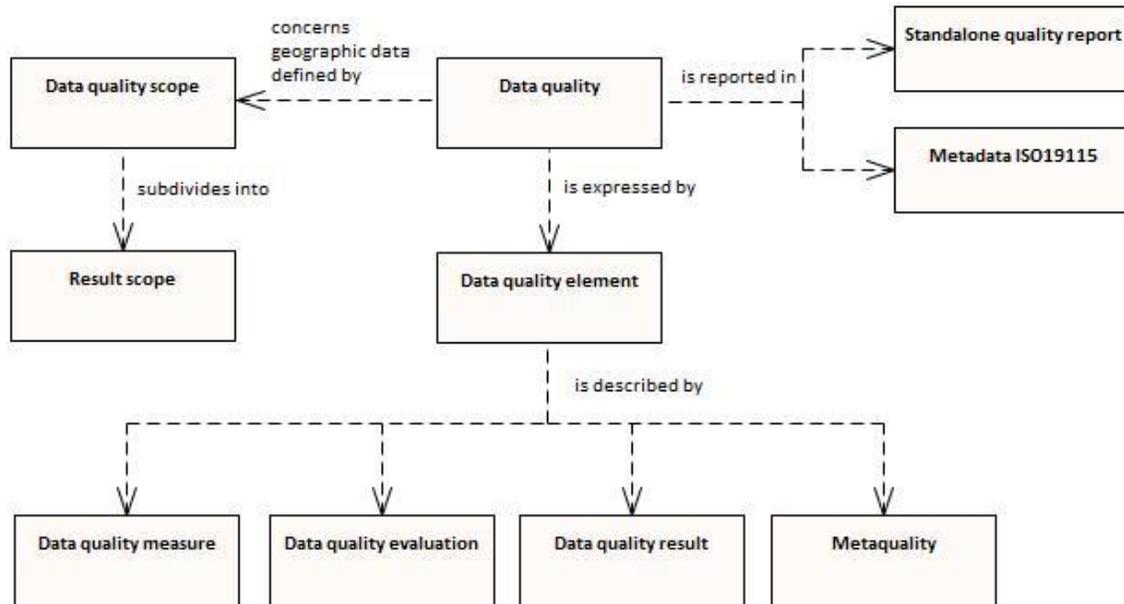
Figure 5-1 - Conceptual model of quality for geographic data[1]

The components of Data Quality Measure can be divided into the following elements[2]:
1. Completeness
2. Logical Consistency
3. Positional Accuracy
4. Thematic Accuracy
5. Temporal Quality
6. Aggregation
7. Usability

Data Quality Evaluation can be divided into the following elements:
1. Data Quality Full Inspection
2. Data Quality Sample Based Inspection
3. Data Quality Indirect Evaluation
4. Data Quality Aggregation Derivation

Data Quality Result can be divided into the following elements:
1. Data Quality Conformance Result
2. Data Quality Quantitative Result

---

[1] ISO 19157, Geographic Information – Data Quality, page 6.

[2] ISO 19157 – Geographic Information, Data Quality page 7.

3. Data Quality Descriptive Result

5.1 Data quality measures

The description of data quality measures in this section is quoted from ISO 19157.

Completeness is defined as the presence and absence of features, their attributes, and relationships. It consists of two Data Quality Elements:

- Commission – excess data present in a dataset;
- Omission – data absent from a dataset.

Logical Consistency is defined as the degree of adherence to logical rules of data structure, attribution, and relationships (data structure can be conceptual, logical or physical). If these logical rules are documented elsewhere (for example in a Product Specification) then the source should be referenced (for example in the data quality evaluation). It consists of four Data Quality Elements:

- Conceptual consistency – adherence to rules of the conceptual schema;
- Domain consistency – adherence of values to the value domains;
- Format consistency – degree to which data is stored in accordance with the physical structure of the dataset;
- Topological consistency – correctness of the explicitly encoded topological characteristics of a dataset.

Positional Accuracy is defined as the accuracy of the position of features within a spatial reference system. It consists of three Data Quality Elements:

- Absolute or external accuracy – closeness of reported coordinate values to values accepted as or being true;
- Relative or internal accuracy – closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true;
- Gridded data positional accuracy – closeness of gridded data spatial position values to values accepted as or being true.

Thematic Accuracy is defined as the accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships. It consists of three Data Quality Elements:

- Classification correctness – comparison of the classes assigned to features or their attributes to a Universe of Discourse (for example ground truth or reference data);
- Non-quantitative attribute correctness – measure of whether a non-quantitative attribute is correct or incorrect;
- Quantitative attribute accuracy – closeness of the value of a quantitative attribute to a value accepted as or known to be true.

Temporal Quality is defined as the quality of the temporal attributes and temporal relationships of features. It consists of three Data Quality Elements:

- Accuracy of a time measurement – closeness of reported time measurements to values accepted as or known to be true;
- Temporal consistency – correctness of the order of events;

- Temporal validity – validity of data with respect to time.

Usability is based on user requirements. All quality elements may be used to evaluate usability. Usability evaluation may be based on specific user requirements that cannot be described using the quality elements described above. In this case, the usability element shall be used to describe specific quality information about a dataset's suitability for a particular application or conformance to a set of requirements.

6   Recommendations for Product Specification developers

This document applies the Data Quality concept from ISO 19157 to the development of S-100 based Product Specifications. This document provides ten recommendations which assist in finding and applying applicable Data Quality Measures as described in S-100.

Completeness > Commission / Omission

DQ_CompletenessCommission: Defined in S-100 Appendix 4c-C, Hydrographic Quality Metadata Attribute Definitions.

DQ_CompletenessOmission: Defined in S-100 Appendix 4c-C, Hydrographic Quality Metadata Attribute Definitions.

> **Recommendation 1**: Data Quality Measure Completeness (Commission/Omission) to be included in the Product Specification.

Logical Consistency > Conceptual Consistency: The Conceptual Schema Language is described in S-100 Part 1 – Conceptual Schema Language. It provides the description of:

- classes;
- attributes;
- basic data types;
- primitive types;
- complex types;
- predefined derived types;
- enumerated types;
- codelist types;
- relationships and associations;
- composition and aggregation;
- stereotypes and optional, conditional, and mandatory attributes and associations;
- naming and name spaces;
- notes;
- packages.

> **Recommendation 2**: Data Quality Measure Conceptual Consistency to follow the guidelines from S-100 Part 1 and to be included in the Product Specification.

Logical Consistency > Domain Consistency: this is described in S-100 Part 5 – Feature Catalogue. This Part provides a standard framework for organizing and reporting the classification of real world phenomena in a set of geographic data. It defines the methodology for classification of the feature types and specifies how they are organized in a Feature Catalogue and presented to the users of a set of geographic data. This methodology is applicable to creating catalogues of feature types in previously uncatalogued domains and

to revising existing Feature Catalogues to comply with standard practice. It applies to the cataloguing of feature types that are represented in digital form. Its principles can be extended to the cataloguing of other forms of geographic data.

> **Recommendation 3**: Data Quality Measure Domain Consistency to follow the guidelines from S-100 Part 5 and to be included in the Product Specification.

Logical Consistency > Format Consistency this is described in S-100 Part 10 – Encoding formats. S-100 does not mandate particular encoding formats so it is left to developers of Product Specifications to decide on suitable encoding standards and to document their chosen format. The issue of encoding information is complicated by the range of encoding standards that are available, which include but are not limited to: ISO/IEC8211, GML, XML, GeoTiff, HDF-5, JPEG2000.

> **Recommendation 4**: Data Quality Measure Format Consistency to follow the guidelines from S-100 Part 10 and to be included in the Product Specification.

Logical Consistency > Topological Consistency this is described in S-100 Part 7 – Spatial Schema. It supports 0, 1, 2, and 2.5 dimensional spatial schemas and two levels of complexity – geometric primitives and geometric complexes.

The conditions for topological consistency are provided in S-100 Part 7 – Spatial Schema, clause 7-4.3 and Appendix 7-A. The Figures in clause 7-4.3 and Appendix 7-A should be referred to for more details.

> **Recommendation 5**: Data Quality Measure Topological Consistency to follow the guidelines from S-100 Part 7 and to be included in the Product Specification.

Positional Accuracy is described by S-100 Part 4c – Metadata - Data Quality.

This is further subdivided into Absolute or External Accuracy, Vertical Position Accuracy, Horizontal Positional Accuracy, Gridded Data Position Accuracy.

One should take notice of the different ways in which spatial data referencing is applied. Point set data includes a coordinate direct position for each point in the point set (points/curves). Gridded data references the grid as a whole. The two spatial properties of gridded data describe how the spatial extent was tessellated into small units and spatial referencing to the earth. The ISO 19123 Standard indicates that a grid may be defined in terms of a coordinate reference system. This requires additional information about the location of the grid's origin within the coordinate reference system, the orientation of the grid axes, and a measure of the spacing between the grid lines. A grid defined in this way is called a rectified grid. If the coordinate reference system is related to the Earth by a datum, the grid is a georectified grid. The essential point is that the transformation of grid coordinates to coordinates of the external coordinate reference system is an affine transformation.

For Positional Accuracy, currently in the hydrographic community the 95% confidence level (Gaussian distribution) is commonly used. The Root Mean Square Error (RMSE) is commonly used in the scientific community. RMSE is the square root of the average of the set of squared differences between dataset coordinate values and coordinate values from an independent source of higher accuracy for identical points.

Other calculation methods are also possible, depending on the specific Product Specification. Comparisons of S-44, S-101, S-102, and S-121 were done. They use various methods of calculating and/or expressing the same concept (uncertainty). Some of the methods in use are:

- S-44: 95% (2*SD)
- S-121: Standard Circular Error (=0.7071*SQRT(SD(X)+SD(Y))), converted to a category attribute.
- S-101: All horizontal positional (2D), vertical (1D), horizontal distance (1D) and orientation (1D) uncertainty attributes concern the 95% confidence level of the variation associated with all sources of measurement, processing and visualization error. Uncertainty is encoded as: (1) <u>uncertainty fixed</u> (The maximum absolute value of the one-dimensional error (for vertical) or two-dimensional error (for horizontal). The error is assumed to be positive and negative.) and (2) <u>uncertainty variable factor</u> (The factor to be applied to a quantity to calculate its uncertainty. The fraction that equates to the factor (or percentage) contributing to the variable uncertainty component is indicated, that is a factor of 5% is encoded as 0.05.)

The situation above is like using nautical miles, imperial miles, and kilometers separately between different Product Specifications – which may work well in isolation, but once you start combining the different Product Specifications and use computer algorithms to create a "smart" system based on these quality parameters, great care must be taken in developing systems to ensure confusion is not introduced.

The IHO DQWG is considering different approaches to addressing this situation, and may leave the separate Product Specifications with their own parameters but inform developers and other users of the Specifications how to convert from one accuracy standard to another.

> **Recommendation 6**: Data Quality Measure Positional Accuracy to follow the guidelines from S-100 Part 4c and to be included in the Product Specification. The calculation of the Positional Accuracy is to be further harmonized where possible with other S-100 based Product Specifications.

Thematic Accuracy is described in ISO 19157.

This is further subdivided into: Quantitative Attribute Accuracy, Non-quantitative Attribute Accuracy, and Thematic Classification Correctness.

The data content of a geographic application is defined in accordance with a view of real world features and in the context of the requirements of a particular application. The content is structured in terms of objects. This document considers two types of object:

1. Features – features are defined together with their properties.
2. Information Types – information types are used to share information among features and other information types. Information types have only thematic attribute properties.

For Thematic Classification Correctness, the assignment of an item to a certain class can either be correct or incorrect.

For Quantitative Attribute Accuracy, the accuracy of quantitative attributes can be measured in terms of uncertainty intervals.

For Non-quantitative Attribute Accuracy, the attribute values of non-quantitative attributes can be correct or incorrect.

> **Recommendation 7**: Data Quality Measure Thematic Accuracy to follow the guidelines from S-100 Part 4c and to be included in the Product Specification.

Temporal Quality is described by Part 4c – Metadata - Data Quality.

Temporal Consistency and Temporal Validity are recommended to be included as this provides the user with the guarantee that any information in the temporal-spatial domain is registered correctly. For data elements with a very precise temporal attribute (for example remote sensing), the temporal accuracy may also be provided.

> **Recommendation 8**: Data Quality Measure Temporal Quality to follow the guidelines from S-100 Part 4c and to be included in the Product Specification. Temporal Consistency and Temporal Validity should be included.

Aggregation: Data Quality specified at upper level (for example series) is applicable at lower level (for example dataset), see Table 6-1 (quoted from ISO 19157) below. If the Data Quality differs between upper and lower level, then supplemental information should be provided at lower level.

Upper level	Series	
	Dataset	
	Subset	
	Feature type	Attribute type
Lower level	Feature instance	Attribute instance

Table 6-1 - Hierarchical levels

In the case of aggregation of different quality results, the standalone quality report will provide full information on the original results (with evaluation procedures and measures applied); the aggregated result; and the aggregation method, whereas the metadata may describe only the aggregated result with a reference to the original results described in the standalone quality report.

The aggregated Data Quality result provides a result if the dataset has passed conformance to the Data Product Specification.

> **Recommendation 9**: Data Quality Measure Aggregation results should be included to indicate if the dataset/dataset series have passed the Product Specifications.

Usability: Usability is based on user requirements. All quality elements may be used to evaluate usability. Usability evaluation may be based on specific user requirements that cannot be described using the quality elements described above. In this case, the usability element shall be used to describe specific quality information about a dataset's suitability for a particular application or conformance to a set of requirements.

All Product Specifications should have a paragraph describing Data Quality. To ensure harmonization across different Product Specifications, DQWG recommends that all Product Specifications share a common text

explaining the concept of Data Quality -> Introduction to Data Quality. The text below is a proposal for this common introduction:

Introduction to data quality

Data quality allows users and user systems to assess fitness for use of the provided data. Data quality measures and the associated evaluation are reported as metadata of a data product. This metadata improves interoperability with other data products and provides usage by user groups that the data product was not originally intended for. The secondary users can make assessments of the data product usefulness in their application based on the reported data quality measures.

For <*this Product Specification*> the following Data Quality Elements have been included[3]:
- Conformance to this Product Specification;
- Intended purpose of the data product;
- Completeness of the data product in terms of coverage;
- Logical Consistency;
- Positional Uncertainty and Accuracy;
- Thematic Accuracy;
- Temporal Quality;
- Aggregation measures;
- Validation checks or conformance checks including:
  - General tests for dataset integrity;
  - Specific tests for a specific data model.

> **Recommendation 10**: Paragraph "Introduction to data quality" to be used as a template by all S-100 based Product Specifications.

7   Data quality measures

This list is derived from ISO 19157.

There are twenty different data quality measures defined that can be used for validation of S-100 based Product Specifications. These are shown in Table 7-1. There is a recommendation following the table for the target result of Positional Accuracy for depth contour lines and gridded bathymetry.

The full list of ISO 19157 measures applicable to S-100 based Product Specifications is available in S-100 Appendix 4c-C – Hydrographic Quality Metadata Attribute Definitions.

Data Quality Measure	Definition	DQ measure / description	Evaluation scope	Applicable to spatial representation types
Completeness / Commission	Excess data present in a dataset, as described by the scope.	numberOfExcessItems / This data quality measure indicates the number of items in the dataset, that should not have been present in the dataset.	dataset/dataset series	All S-100 based PS
Completeness / Commission	Excess data present in a dataset, as described by the scope.	numberOfDuplicateFeatureInstances / This data quality measure indicates the total number of exact duplications of feature instances within the data.	dataset/dataset series	All S-100 based PS
Completeness / Omission	Data absent from the dataset, as described by the scope.	numberOfMissingItems / This data quality measure is an indicator that shows that a specific item is missing in the data.	dataset/dataset series/spatial object type	All S-100 based PS

---

[3] As deemed necessary by the IHO – Hydrographic Standards and Services Committee

Data Quality Measure	Definition	DQ measure / description	Evaluation scope	Applicable to spatial representation types
Logical Consistency / Conceptual Consistency	Adherence to the rules of a conceptual schema.	numberOfInvalidSurfaceOverlaps / This data quality measure is a count of the total number of erroneous overlaps within the data. Which surfaces may overlap and which must not is application dependent. Not all overlapping surfaces are necessarily erroneous.	spatial object / spatial object type	PS with geometric surfaces.
Logical Consistency / Domain Consistency	Adherence of the values to the value domains.	numberOfNonconformantItems / This data quality measure is a count of all items in the dataset that are not in conformance with their value domain.	spatial object / spatial object type	All S-100 based PS
Logical Consistency / Format Consistency	Degree to which data is stored in accordance with the physical structure of the data set, as described by the scope	physicalStructureConflictsNumber / This data quality measure is a count of all items in the dataset that are stored in conflict with the physical structure of the dataset.	dataset/dataset series	All S-100 based PS
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	rateOfFaultyPointCurveConnections / This data quality measure indicates the number of faulty link-node connections in relation to the number of supposed link-node connections. This data quality measure gives the erroneous point-curve connections in relation to the total number of point-curve connections.	spatial object / spatial object type	PS with curves.
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfMissingConnectionsUndershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to undershoots.	spatial object / spatial object type	PS with curves.
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfMissingConnectionsOvershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots.	spatial object / spatial object type	PS with curves.
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfInvalidSlivers / This data quality measure is a count of all items in the dataset that are invalid sliver surfaces. A sliver is an unintended area that occurs when adjacent surfaces are not digitized properly. The borders of the adjacent surfaces may unintentionally gap or overlap to cause a topological error.	dataset / dataset series	PS with geometric surfaces.
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfInvalidSelfIntersects / This data quality measure is a count of all items in the dataset that illegally intersect with themselves.	spatial object / spatial object type	PS with curves / geometric surfaces.
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfInvalidSelfOverlap / This data quality measure is a count of all items in the dataset that illegally self-overlap.	spatial object / spatial object type	PS with curves / geometric surfaces.

Data Quality Measure	Definition	DQ measure / description	Evaluation scope	Applicable to spatial representation types
Positional Accuracy / Absolute or External Accuracy	Closeness of reported coordinative values to values accepted as or being true.	Root Mean Square Error / Standard deviation, where the true value is not estimated from the observations but known a priori.	spatial object / spatial object type	PS with objects that have coordinative values associated.
Positional Accuracy / Vertical Position Accuracy	Closeness of reported coordinative values to values accepted as or being true.	linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.	spatial object / spatial object type	PS with objects that have a vertical coordinative values associated.
Positional Accuracy / Horizontal Position Accuracy	Closeness of reported coordinative values to values accepted as or being true.	linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.	spatial object / spatial object type	PS with objects that have a horizontal coordinative values associated.
Positional Accuracy / Gridded Data Position Accuracy	Closeness of reported coordinative values to values accepted as or being true.	Root mean square error of planimetry / Radius of a circle around the given point, in which the true value lies with probability P.	spatial object / spatial object type	PS with objects that have a gridded coordinative values associated.
Temporal Quality / Temporal Consistency	Consistency with time.	Correctness of ordered events or sequences, if reported.	dataset/dataset series/spatial object type	PS with objects that have a time value associated.
Thematic Accuracy / ThematicClassification Correctness	Comparison of the classes assigned to features or their attributes to a universe of discourse.	miscalculationRate / This data quality measure indicates the number of incorrectly classified features in relation to the number of features that are supposed to be there. [Adapted from ISO 19157] This is a RATE which is a ratio, and is expressed as a REAL number representing the rational fraction corresponding to the numerator and denominator of the ratio. For example, if there are 1 items that are classified incorrectly and there are 100 of the items in the dataset then the ratio is 1/100 and the reported rate = 0.01.	dataset/dataset series/spatial object type	All S-100 based PS.
Aggregation Measures / AggregationMeasures	In a data Product Specification, several requirements are set up for a product to conform to the Specification.	DataProductSpecificationPassed / This data quality measure is a boolean indicating that all requirements in the referred data Product Specification are fulfilled.	dataset/dataset series/spatial object type	PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types
Aggregation Measures / AggregationMeasures	In a data Product Specification, several requirements are set up for a product to conform to the specification.	DataProductSpecificationFailRate / This data quality measure is a number indicating the number of data Product Specification requirements that are not fulfilled by the current product/dataset in relation to the total number of data Product Specification requirements.	dataset/dataset series/spatial object type	PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types

Table 7-1 - Recommended data quality measures

NOTE: Recommendations for Positional Accuracy / Absolute or External Accuracy[4]:

Maximum RMSE (horizontal) = E / 10000

Maximum RMSE (vertical) = Vint / 6

Recommendation for Positional Accuracy / Gridded Data Position Accuracy:

Maximum RMSE (horizontal) = GSD / 6

Maximum RMSE (vertical) = GSD / 3

Where:

E = Denominator of intended scale of mapping

Vint = normal contour line interval

GSD = Ground Sampling Distance

8 Minimum standard for data validation

A minimum standard set of checks for data validation is under development and will be added to this document when completed.

---

[4] INSPIRE D2.8.II.1 Data Specification on Elevation – Technical Guidelines, page 95

Page intentionally left blank