

19th Meeting of the Data Quality Working Group

Report on the Cross check of DQ chapters of S-100 based product specification

Agenda Item 3.1A



1. INTRODUCTION/OVERVIEW

International Hydrographic Organization According to the action DQWG 18/05, S-100 Product Specification subWG, comprised of Chair, NL, SE, US and UNH, has implemented the Cross check of DQ chapters of 7 published S-100 PSs.

Product ID	♦ Name ♦	Version	Status	Domain	Date updated	4
S-100	Universal Hydrographic Data Model	5.1.0	Published	IHO Hydro	2023-10-31	
S-101	Electronic Navigational Chart	1.1.0	Published	IHO Hydro	2023-11-21	
S-102	Bathymetric Surface	2.2.0	Published	IHO Hydro	2023-07-05	
S-104	Water Level Information for Surface Navigation	1.1.0	Published	IHO Hydro	2023-08-11	
S-111	Surface Currents Product Specification	1.2.0	Published	IHO Hydro	2023-08-11	
S-121	Maritime Limits and Boundaries	1.0.0	Published	IHO Hydro	2021-10-29	
S-122	Marine Protected Areas	1.0.0	Published	IHO Hydro	2021-05-06	
S-123	Marine Radio Services	1.0.0	Published	IHO Hydro	2021-05-06	
S-124	Navigational Warnings	1.0.0	Published	IHO Hydro	2023-07-28	
S-127	Marine Traffic Management	1.0.0	Published	IHO Hydro	2021-05-06	
S-128	Catalogue of Nautical Products	1.0.0	Published	IHO Hydro	2023-09-14	
S-129	Under Keel Clearance Management Product Specification	1.0.0	Published	IHO Hydro	2020-08-21	
S-130	Polygonal Demarcations of Global Sea Areas	1.0.0	Published	IHO Hydro	2023-07-20	
S-131	Marine Harbour Infrastructure	1.0.0	Published	IHO Hydro	2023-08-17	
S-201	Aids to Navigation (AtoN) Information	1.1.0	Published	IALA AtoNs	2023-04-27	
S-240	DGNSS Station Almanac	1.0.0	Published	IALA AtoNs	2021-09-17	
S-401	Inland ENC Product Specification	1.0.0	Published	Inland ENC	2020-11-24	
S-421	Route Plan	1.0.0	Published	IEC	2021-07-02	
S-98	Data Product Interoperability in S-100 Navigation Systems	1.0.0	Published	IHO Hydro	2022-05-23	



2. PRINCIPLES OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization

The Cross check of DQ chapters follow principles as below:

- a) The Cross check shall be carried out in accordance with S-97 Ed 1.1.0 "IHO Guidelines for Creating S-100 Product Specifications".
- b) b) Only the DQ chapters of S-100 PSs shall be checked and the results presented in the form of a cross-check matrix.
- c) c) DQ elements included in the other parts of each S-100 PS shall be described in the form of Notes under the respective cross-check matrix.



International Hydrographic Organization

3.1 Identify the Required DQ Measures

10 recommendations for the development of S-100 based Product Specifications.

10 Recommendations
1.Completeness
2.Conceptual consistency
3. Domain consistency
4. Format consistency
5. Topological consistency
6. Positional Accuracy
7. Thematic Accuracy
8. Temporal Quality
9. Aggregation
10. Introduction to DQ Paragraph





International Hydrographic Organization

3.1 Identify the Required DQ Measures

The recommended data quality measures of each S-100 PS were identified by the subWG according to S-97 part C Table C-7-1 – Recommended data quality measures.

Data Quality Measure	DQ measure / description	Applicable to spatial representation types	124	130	131	201	240	401	42
Completeness / Commission	numberOfExcessItems / This data quality measure indicates the number of items in the dataset, that should not have been present in the dataset.	All S-100 based PS	Y	Υ	Υ	Y	Υ	Υ	Υ
Completeness / Commission	numberOfDuplicateFeatureInstances / This data quality measure indicates the total number of exact duplications of feature instances within the data.	All S-100 based PS	Y	Y	Y	Y	Y	Y	Υ
Completeness / Omission	numberOffMissingItems / This data quality measure is an indicator that shows that a specific item is missing in the data.	All S-100 based PS	Y	Y	Y	Y	Y	Y	Y
Logical Consistency / Conceptual Consistency	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.	PS with geometric surfaces.	Y	Y	Y	Y	Y	Υ	Y
Logical Consistency / Domain Consistency	numberOfNonconformantItems / This data quality measure is a count of all items in the dataset that are not in conformance with their value domain.	All S-100 based PS	Y	Υ	Υ	Y	Y	Y	Υ
Logical Consistency / Format Consistency	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.	All S-100 based PS	Y	Υ	Υ	Y	Y	Y	Y
Logical Consistency / Topological Consistency	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.	PS with curves.	Y	Y	Y	Y	Y	Y	Y
Logical Consistency / Topological Consistency	numberOffMissingConnectionsUndershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to undershoots.	PS with curves	Y	Υ	Υ	Y	Y	Y	Y
Logical Consistency / Topological Consistency	numberOffMissingConnectionsOvershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots.	PS with curves.	Y	Y	Y	Y	Y	Y	Y
Logical Consistency / Topological Consistency	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.	PS with geometric surfaces.	Y	Y	Y	Y	Y	Y	Υ
Logical Consistency / Topological Consistency	numberOfInvalidSelfIntersects / This data quality measure is a count of all items in the dataset that illegally intersect with themselves.	PS with curves / geometric surfaces.	Y	Y	Y	Y	Y	Y	Y
Logical Consistency / Topological Consistency	numberOfInvalidSelfOverlap./ This data quality measure is a count of all items in the dataset that illegally self-overlap.	PS with curves / geometric surfaces.	Y	Υ	Υ	Y	Υ	Υ	Y
Positional Accuracy / Absolute or External Accuracy	Root Mean Square Error / Standard deviation, where the true value is not estimated from the observations but known a priori.	PS with objects that have coordinative values associated.	Y	Υ	Υ	Y	Y	Y	Y
Positional Accuracy / Vertical Position Accuracy	linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.	PS with objects that have a vertical coordinative values associated.	Y	N	N	Y	Y	Y	N
Positional Accuracy / Horizontal Position Accuracy	linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.	PS with objects that have a horizontal coordinative values	Y	Υ	Υ	Y	Y	Υ	Y
		associated.							
Positional Accuracy / Gridded Data Position Accuracy	Root mean square error of planimetry / Radius of a circle around the given point, in which the true value lies with probability P.	PS with objects that have a gridded coordinative values associated.	N	N	N	N	N	N	N
Temporal Quality / Temporal Consistency	Correctness of ordered events or sequences, if reported.	PS with objects that have a time value associated.	Y	N	Y	Y	Y	Y	N
Classification Correctness	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.	All S-100 cased I S.	¥	¥	Y	¥	Y	£	¥
Aggregation Measures / Aggregation Measures	DataProductSpecificationPassed / This data quality measure is a boolean indicating that all requirements in the referred data Product Specification are fulfilled.	PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types	Y	N	Y	Y	Y	Y	Υ
Aggregation Measures / Aggregation Measures	<u>DataProductSpecificationFailRate</u> / 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.	PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types	Y	N	Y	Y	Y	Y	Υ



3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization 3.2 Carry Out the Cross Check 3.2.1 S-124

3.2.1.1 DQ chapters of S-124

The DQ chapter of S-124 is as follow:

7 Data Quality

7.1 Introduction

Datasets conforming to S-124 should always be created with the best available source information. Due to the urgency of the information, datasets may be based on incomplete or unconfirmed information and mariners will need to take this into account when deciding what reliance to place on the information contained therein. It is often not possible to determine quantifiable values to measures of data quality. Generally the quality of information can be made evident from the navigational warning amplifying text by the use of qualitative words such as 'approximate', 'reported', 'in the vicinity of' and 'about'.

Example: Dredging operations will be taking place in the vicinity of Goldwood Sawmill (49°12.47'N / 123°04.83'W), in the Mitchell Slough starting on Saturday, February 2 to Monday, February 5.

3.2.1.2 Result of cross check

Data Quality Measure	S-124
1.Completeness	N
2.Conceptual consistency	N
3. Domain consistency	N
4. Format consistency	N
5. Topological consistency	N
6. Positional Accuracy	N
7. Thematic Accuracy	N
8. Temporal Quality	N
9. Aggregation	N
10. Introduction to DQ Paragraph	Y

Y=YES, N=NO

NOTES:

- a) Regarding Measure 10 ("Introduction to DQ Paragraph"), S-124 does include an introduction, however, this is not in accordance with S-97.
- **3.2.1.3 DQ elements included in the other parts of S-124** No DQ element is included in the other parts of S-124.



International Hydrographic Organization 3.2.2 S-130

3.2.2.1 DQ chapters of S-130

The DQ chapter of S-130 is as follow:

9 Data quality

9.1 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 S-130 the following Data Quality Elements have been included:

- 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;



International Hydrographic Organization

3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

- Thematic Accuracy;
- Temporal Quality;
- Aggregation measures;
- Validation checks or conformance checks including:
 - General tests for dataset integrity;
 - Specific tests for a specific data model.

9.2 Completeness

.z.ı Commission

Commission is applicable for S-130.

S-130 products must be tested with Commission checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

In terms of Commission, S-130 products shall at least populate numberOfExcessItems that indicates the number of items that should not have been present in the dataset, and numberOfDuplicateFeatureInstances that indicates the total number of exact duplications of feature instances within the data.

9.2.2 Omission

Omission is applicable for S-130.

S-130 products must be tested with Omission checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

In terms of Omission, S-130 products shall at least populate numberOfMissingItems that is the total number of missing items.

9.3 Logical Consistency

9.3.1 Conceptual Consistency

Conceptual Consistency is applicable for S-130 and follows the guidelines from S-100 Part 1.

S-130 products must be tested with Conceptual Consistency checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. Data should only be published if it passes the test.

In terms of Conceptual Consistency, S-130 products shall at least populate numberOfInvalidSurfaceOverlaps that is the total number of erroneous overlaps within the

9.3.2 Domain Consistency

Domain Consistency is applicable for S-130 and follows the guidelines from S-100 Part 5.

S-130 products must be tested with Domain Consistency checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

In terms of Domain Consistency, S-130 products shall at least populate numberOfNonconformantItems that is a count of all items in the dataset that are not in conformance with their value domain.

9.3.3 Format Consistency

Format Consistency is applicable for S-130 and follows the guidelines from S-100 Part 10b. S-130 products must be tested with Format Consistency checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. Data should only be published if it passes the test.

In terms of Format Consistency, S-130 products shall at least populate physicalStructureConflicts that is a count of all items in the dataset that are stored in conflict

9.3.4 Topological Consistency

Topological Consistency is applicable for S-130 and follows the guidelines from S-100 Part 7. S-130 products must be tested with Topological Consistency checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. Data should only be published if it passes the test.

In terms of Topological Consistency, S-130 products shall at least populate rateOfFaultyPointCurveConnections that is the number of faulty link-node connections in relation to the number of supposed link-node connections, numberOfMissingConnectionsUndershoots that is a count of items in the dataset within the parameter tolerance that are mismatched due to undershoots, numberOfMissingConnectionsOvershoots that is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots, numberOfInvalidSlivers that is a count of all items in the dataset that are invalid sliver surfaces, numberOfInvalidSelfIntersects that is a count of all items in the dataset that illegally intersect with themselves, and numberOfInvalidSelfOverlap that is all items in the dataset that illegally self-overlap.

9.4 Positional Uncertainty and Accuracy

9.4.1 Vertical Position Accuracy

Vertical Position Accuracy is not applicable for S-130.

9.4.2 Horizontal Position Accuracy

Horizontal Position Accuracy is applicable for S-130 and follows the guidelines from S-100 Part 4c

S-130 products must be tested with Horizontal Position Accuracy checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

In terms of Horizontal Position Accuracy, S-130 products shall at least populate circular Error95 that indicates the radius describing a circle in which the true point location lies with the probability of 95%.

9.4.3 Gridded Data Positional Accuracy

Gridded Data Position Accuracy is not applicable for S-130.



3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization 9.5 Thematic Accuracy

9.5.1 Thematic Classification Correctness

Thematic Classification Correctness is applicable for S-130 and follows the guidelines from S-100 Part 4c.

S-130 products must be tested with Thematic Classification Correctness checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. Data should only be published if it passes the test.

In terms of Thematic Classification Correctness, S-130 products shall at least populate miscalculationRate that is the number of incorrectly classified features in relation to the number of features that are supposed to be there.

9.5.2 Non-Quantitative Attribute Accuracy

Non-Quantitative Attribute Accuracy is applicable for S-130 and follows the guidelines from S-100 Part 4c.

S-130 products must be tested with Non-Quantitative Attribute Accuracy checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

The accuracy of non-quantitative attributes can be correct or incorrect. S-130 products shall at least populate numberOfIncorrectAttributeValues that is a count of all attribute values where the value is incorrect.

9.5.3 Quantitative Attribute Accuracy

Quantitative Attribute Accuracy is applicable for S-130 and follows the guidelines from S-100 Part 4c.

S-130 products must be tested with Quantitative Attribute Accuracy checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. It is allowable to publish the data with a quality statement which indicates non-conformance.

The accuracy of quantitative attributes can be measured in terms of uncertainty intervals. S-130 products shall at least populate attributeValueUncertainty3Sigma that is half the length of the interval defined by an upper and lower limit in which the true value for the quantitative attribute lies with a probability of 95%.

9.6 Temporal Quality

9.6.1 Temporal Consistency

Temporal Consistency is not applicable for S-130.

9.6.2 Temporal Validity

Temporal Validity is not applicable for S-130.

9.6.3 Temporal Accuracy

Temporal Accuracy is not applicable for S-130.

9.7 Aggregation

Aggregation is not applicable for S-130.

9.8 Quality measure elements

The data quality measures recommended in S-97 (Part C) and their applicability in S-130 are indicated in Table 9.1 below. NA indicates the measure is not applicable. The application schema above has indicated how the data quality elements will be related to the data items, and the encoding description below will indicate how the quality elements will be encoded.

No.	Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Scope in S-130
1	Complete ness / Commissi on	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 features and info types
2	Complete ness / Commissi on	Excess data present in a dataset, as described by the scope.	numberOfDuplicateFeatur eInstances / This data quality measure indicates the total number of exact duplications of feature instances within the data.	dataset/ dataset series	All features and info types
3	Complete ness / 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 features and info types
4	Logical Consisten cy / Conceptu al Consisten cy	Adherence to the rules of a conceptual schema.	numberOfInvalidSurfaceO verlaps / 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	Features with surface geometry; spatial objects of type surface
5	Logical Consisten cy / Domain Consisten cy	Adherence of the values to the value domains.	numberOfNonconformantit ems / 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 features and info types



International Hydrographic Organization

3.2.2.2 Result of cross check

Data Quality Measure	S-130
1.Completeness	Y
2.Conceptual consistency	Y
3. Domain consistency	Y
4. Format consistency	Y
5. Topological consistency	Y
6. Positional Accuracy	Y
7. Thematic Accuracy	Y
8. Temporal Quality	N/A
9. Aggregation	N/A
10. Introduction to DQ Paragraph	Y

Y=YES, N=NO, N/A=not applicable

NOTES:

-,

3.2.2.3 DQ elements included in the other parts of S-130

No DQ element is included in the other parts of S-130.



3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization 3.2.3 S-131

3.2.3.1 DQ chapters of S-131

The DQ chapter of S-131 is in the pictures.

3.2.3.2 Result of cross check

Data Quality Measure	S-131
1.Completeness	Y
2.Conceptual consistency	Y
3. Domain consistency	Y
4. Format consistency	Y
5. Topological consistency	Y
6. Positional Accuracy	Y
7. Thematic Accuracy	Y
8. Temporal Quality	Y
9. Aggregation	Y
10. Introduction to DQ Paragraph	Υ

Y=YES, N=NO

NOTES: a) Regarding Measure 10 ("Introduction to DQ Paragraph"), S-131 follows the template provided by S-97 but makes the, unnecessary, distinction of positional accuracy to depth and positional (horizontal). b) Lacks a detailed text description of each quality element.

3.1 Introduction

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 scope for 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 S-131 the following data quality elements have been included:

- Conformance to this Product Specification;
- Intended purpose of the data product;
- Completeness of the data product in terms of coverage;
- Logical consistency;
- Depth uncertainty and accuracy;
- Positional uncertainty and accuracy;
- Thematic accuracy;
- Temporal quality;
- Aggregation measures;
- Validation checks or conformance checks including:
- · General tests for dataset integrity; and
- Specific tests for compliance against the S-131 data model.

8.2 Quality measure elements

The data quality measures recommended in S-97 (Part C) and their applicability in S-131 are indicated in Table 8.1 below. NA indicates the measure is not applicable. This table reproduces the first 4 columns of the data quality checklist recommended elements and replaces the final column with descriptions of the scope of the element in the context of S-131 datasets.



International Hydrographic Organization

Table 8.1 - Quality measure elements

No.	Data quality element	Definition	DQ measure / description	Evaluation scope	Scope in S-131
1	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/datas et series	All features and info types.
2	Completeness I 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/datas et series	All features and info types.
3	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/datas et series/spatial object type	All features and info types.
4	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	Features with surface geometry, spatial objects of type surface.
5	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 features and info types.
6	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/datas et series	All features and info types.
7	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	Features with curve geometry; spatial objects of curve types.
8	Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfMissingConnectionsUnders hoots / 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	Features with curve geometry, spatial objects of curve types.
9	Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOfMissingConnectionsOversh oots / This data quality measure is a court of items in the dataset within the parameter tolerance that are mismatched due to overshoots.	spatial object / spatial object type	Features with curve geometry, spatial objects of curve types.

No.	Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Scope in S-131
10	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	Features with surface geometry; spatial objects of type surface.
11	Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOflnvalidSelfintersects / This data quality measure is a count of all items in the dataset that illegally intersect with themselves.	spatial object / spatial object type	Features with surface geometry; spatial objects of type surface or curve.
12	Logical Consistency / Topological Consistency	Correctness of the explicitly encoded topological characteristics of the dataset, as described by the scope.	numberOffnvalldSelfOverlap / This data quality measure is a count of all items in the dataset that illegally self- overlap.	spatial object / spatial object type	Features with surface geometry; spatial objects of type surface or curve.
13	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	Objects that have coordinative values associated.
14	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	N/A S-131 does not use 3-D coordinates.
15	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	Objects that have a horizontal coordinate values associated.
16	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	NA. S-131 does not have features with gridded geometry.
17	Temporal Quality / Temporal Consistency	Consistency with time.	Correctness of ordered events or sequences, if reported.	dataset/datas et series/spatial object type	Features with time intervals, fixed/periodic date ranges, schedules.



International Hydrographic Organization

No.	Data quality element and sub element	Pefinition	DQ measure / description	Evaluation scope	Scope in S-131
18	Thematic Accuracy / ThematicClassification Correctness	comparison of the lasses assigned to satures or their tirributes to a niverse of iscourse.	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 1915 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 is 1 liem 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/datas et series/spatial object type	All features and info types.
19	Aggregation Measures / AggregationMeasures	n a data product pecification, everal equirements are et up for a product o conform to the pecification.	DataProductSpecificationPassed / This data quality measure is a boolean indicating that all requirements in the referred data product specification are fulfilled.	dataset/datas et series/spatial object type	Dataset as a whole.
20	Aggregation Measures / AggregationMeasures	n a data product pecification, everal equirements are et up for a product o conform to the pecification.	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/datas et series/spatial object type	Dataset as a whole.

8.3 Test methods for data compliance and usability

Test methods for evaluating data compliance consist of executing the relevant tests from the Validation Checks (Annex D) for each quality element in Table 8.1 and counting the number of instances in the dataset which fail the checks for that quality element.

Note that in some cases "executing the relevant test" may involve comparing the encoded S-131 dataset to the source material by visual means (e.g., for measures 17 and 18). For tests requiring visual comparison of encoded data to source material, sampling methods may be used if the volume of data precludes checking all the relevant data objects.

Accuracy computations for Positional Accuracy / Absolute or External Accuracy should use the following recommendations:

Maximum RMSE (horizontal) = E / 10000

E = Denominator of intended scale of mapping

In addition to the above, dataset usability must be assessed against:

- Intended user requirements in regard to coverage, scale and specific content requirements as defined by the Producing Agency and key stakeholders;
- . Conformance to established maintenance processes (see Clause 10); and

Overall compliance with the S-131 Product Specification, including context-specific evaluation
of individual encoding instances for requirement of conformance to checks classified as "Error"
and "Warning" in Annex D – Validation Checks.

For dataset integrity requirements, see clause 13.8.

8.4 Data quality testing and reporting

S-131 products must be tested with the S-131 specific checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. The checks are a mix of data format validation checks, conformance to standard checks and logical consistency checks. The checks are listed in Annex D.

Production and certification processes for S-131 data may include a standalone quality report which provides full information on the original results (with evaluation procedures and measures applied). This report may be included in the exchange set as a support file (discovery metadata for data quality reports should indicate the datasets to which they apply and resource purpose "other" (ISOMetadataFile if the report is in ISO format).

Alternatively, a quality report distributed with the exchange set may describe only the aggregated results along with a reference to a location where the full results may be obtained.

Data Quality Measure Aggregation results should be computed to indicate if the dataset/dataset series have passed the Product Specifications. The elements which must be included are described in Table 8.2 below.

Table 8.2 - Elements of data quality aggregated report (extract from S-97 Part C checklist)

Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Applicable to spatial representation types
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	All features and information types of the dataset
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	All features and information types of the dataset



3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization 3.2.4 S-201

3.2.4.1 DQ chapters of S-201

The DQ chapter of S-201 is in the right picture.

3.2.4.2 Result of cross check

Data Quality Measure	S-201
1.Completeness	N
2.Conceptual consistency	N
3. Domain consistency	N
4. Format consistency	N
5. Topological consistency	N
6. Positional Accuracy	Y
7. Thematic Accuracy	N
8. Temporal Quality	N
9. Aggregation	N
10. Introduction to DQ Paragraph	Y

Y=YES, N=NO

NOTES: Regarding Measure 10 ("Introduction to DQ Paragraph"), S-201 does include an introduction, however, this is not in accordance with S-97.

3.2.4.3 DQ elements included in the other parts of S-201No DQ element is included in the other parts of S-201.

6 Data Quality

6.1 Introduction

Data Quality is considered to be meta information and for S-201 it is divided into three parts:

- Data Compliance and integrity
- Positional accuracy
- lineage

6.2 Data Compliance and Integrity

S-201 datasets must be validated using the conformance checks that are listed in S-58. The data quality elements listed in S-100 Part 4C that are applicable to S-201 are indicated in the table of conformance checks listed in S-58.

Datasets must not be published unless they pass all the compliance checks designated as "Critical".

The detailed results of applying compliance checks listed in S-58 are not required to be reported as part of the exchange set. They may be conveyed as support files or by separate arrangement.

6.3 Positional accuracy

S-201 positional accuracy data quality indicators do not use the model of data quality elements defined in S-100 Part 4c.

However, positional accuracy must be evaluated and must be indicated in dataset metadata or spatial attribute metadata as provided in the application schema (Figure 12). This specification does not prescribe a specific required level of positional accuracy

6.4 Lineage

Lineage and process step information elements are not required to be present in S-201 datasets distributed to end user. They may be included as extra metadata in exchange sets distributed to vendors and distributors.

Required source information about S-201 datasets is limited to the source and distribution information contained in discovery metadata described in S-100 Part 4a and Section 12 of this product specification.



3. IMPLEMENTATION PROCESS OF THE CROSS CHECK OF DQ CHAPTERS

International Hydrographic Organization 3.2.5 S-240

3.2.5.1 DQ chapters of S-240

The DQ chapter of S-240 is in the pictures.

3.2.5.2 Result of cross check

Data Quality Measure	S-240
1.Completeness	Y
2.Conceptual consistency	Y
3. Domain consistency	Y
4. Format consistency	Y
5. Topological consistency	Y
6. Positional Accuracy	Y
7. Thematic Accuracy	Y
8. Temporal Quality	Y
9. Aggregation	Y
10. Introduction to DQ Paragraph	Y

Y=YES, N=NO

NOTES:

a) Lacks a detailed text description of each quality element.

3.2.5.3 DQ elements included in the other parts of S-240

No DQ element is included in the other parts of S-240.

6 Data Quality

6.1 Introduction

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 S-240 the following data quality elements have been included:

- · 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;
- · Elements specifically required for the data product;
- · Validation checks or conformance checks including:
 - General tests for dataset integrity;
 - Specific tests for a specific data model;

6.2 Quality measure elements

The recommended data quality measures and their applicability in S-240 are indicated in the table below. NA indicated the measure is not applicable. This table reproduces the first 4 columns of the data quality checklist recommended elements and replaces the final column with descriptions of the scope of the element in the context of S-240 datasets.

Data quality element and sub element	Definition	 Evaluation scope	Scope in S- 240	



International Hydrographic Organization

Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Scope in S- 240	
Completeness / Commission	excess data resent in a lataset, as lescribed 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 features and info types	
Completeness / Commission	excess data resent in a lataset, as lescribed 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 features and info types	
Completeness / Omission	lata absent rom the lataset, as lescribed by he scope.	numberOffMissingItems / 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 features and info types	
Logical Consistency / Conceptual Consistency	Adherence to the rules of a onceptual chema.	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	Features with surface geometry; spatial objects of type surface	
Logical Consistency / Domain Consistency	dherence of ne values to ne value lomains.	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 features and info types	
Logical Consistency / Format Consistency	legree to which ata is stored in accordance with the physical tructure of the lata set, as escribed 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 features and info types	
Logical Consistency / Topological Consistency	Correctness of he explicitly ncoded opological haracteristics if the dataset, is described by he 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	Features with curve geometry; spatial objects of curve types	
Logical Consistency / Topological Consistency	Correctness of the explicitly incoded opological haracteristics of the dataset, is 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	Features with curve geometry; spatial objects of curve types	

Data quality element and sub element			Evaluation scope	Features with curve geometry; spatial objects of curve types	
Logical Consistency / Topological Consistency	Correctness of the explicitly elecoded to pological characteristics of the dataset, as described by the scope.	e explicitly coded This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots. Ite dataset, described by			
Logical Consistency / Topological Consistency	Correctness of the explicitly elected the pological 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	Features with surface geometry; spatial objects of type surface	
Logical Consistency / Topological Consistency	Correctness of the explicitly elected the pological characteristics of the dataset, and 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	Features with surface geometry; spatial objects of type surface or curve	
Logical Consistency / Topological Consistency	Correctness of the explicitly encoded the pological characteristics the dataset, and described by the scope.	numberOfinvalidSelfOverlap / This data qualify measure is a count of all items in the dataset that illegally self-overlap.	spatial object / spatial object type	Features with surface geometry; spatial objects of type surface or curve	
Positional Accuracy / Absolute or External Accuracy	Coseness 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	objects that have coordinative values associated.	
Positional Accuracy / Vertical Position Accuracy	Coseness of reported coordinative values to values accepted as or bring 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	objects that have a vertical coordinate values associated.	
Positional Accuracy / Horizontal Position Accuracy	Coseness of reported coordinative values to values accepted as or bring 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	objects that have a horizontal coordinate values associated	
Positional Accuracy / Gridded Data Position Accuracy	Coseness of r ported coordinative v lues to values a cepted 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	NA. S-240 does not have features with gridded geometry	



International Hydrographic Organization

Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Scope in S- 240	
Temporal Quality / Temporal Consistency	Consistency Correctness of ordered events or sequences, if reported.		dataset/dataset series/spatial object type	Features with time intervals, fixed/periodi date ranges, schedules.	
Thematic Accuracy I ThematicClassificationCorrectness	comparison of he classes ussigned to eatures or their stributes to a miverse 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 features and info types	
Aggregation Measures / AggregationNeasures			dataset/dataset series/spatial object type	Dataset as a whole	
Aggregation Measures / AggregationMeasures	n a data roduct pecification, everal equirements re set up for a roduct to onform to the pecification.	DataProductSpecificationFallRate / 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	Dataset as a whole	

6.3 Test methods

Define (or provide a pointer to) testing methods for each element in the table. Some test methods may just reference Annex B (Validation Checks) with the direction to report the number of instances in the dataset which fail the checks for that quality element.

6.4 Data quality testing and reporting

S-240 products must be tested with the S-240 specific checks prior to release by the data producer. The data producer must review the check results and address any issues to ensure sufficient quality of the data products. The checks are a mix of data format validation checks, conformance to standard checks and logical consistency checks. The checks are listed in Annex B.

Production and certification processes for S-240 data should include a standalone quality report which provides full information on the original results (with evaluation procedures and measures applied).

The dataset or exchange set metadata that is distributed with the exchange set 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 an indication if the dataset has passed conformance to the Data Product Specification.

Data Quality Measure Aggregation results should be included to indicate if the dataset/dataset series have passed the Product Specifications. The elements which must be included are described in the table below.

Table 2. Elements of data quality aggregated report (extract from DQWG checklist)

Data quality element and sub element	Definition	DQ measure / description	Evaluation scope	Applicable to spatial representation types
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	All features and information types of the dataset
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	All features and information types of the dataset





International Hydrographic Organization 3.2.6 S-401

3.2.6.1 DQ chapters of S-401

The DQ chapter of S-401 is in the right picture.

3.2.6.2 Result of cross check

Data Quality Measure	S-401
1.Completeness	N
2.Conceptual consistency	N
3. Domain consistency	N
4. Format consistency	N
5. Topological consistency	N
6. Positional Accuracy	N
7. Thematic Accuracy	N
8. Temporal Quality	N
9. Aggregation	N
10. Introduction to DQ Paragraph	Y

Y=YES, N=NO

NOTES:-.

3.2.6.3 DQ elements included in the other parts of S-401No DQ element is included in the other parts of S-401.

6 Data Quality

6.1 Introduction

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 S-401 the following data quality elements have been included;

- Conformance to this Product Specification;
- Intended purpose of the data product;
- Completeness of the data product in terms of coverage;
- Logical Consistency;
- Depth Uncertainty and Accuracy;
- Positional Uncertainty and Accuracy;
- Thematic Accuracy;
- Temporal Quality;
- Aggregation measures:
- Validation checks or conformance checks including:
 - · General tests for dataset integrity;
 - Specific tests for compliance against the S-401 data model.

For S-401 data quality is divided into two parts – data compliance, usability and integrity against all requirements of S-401; and bathymetric data quality.

6.1.1 Data Compliance and Usability

All S-401 datasets must be validated against the above data quality elements using conformance checks that are located in Annex C – IENC Validation Checks. As a minimum requirement, all datasets must conform to all checks that are categorized as "Critical" in Annex C.

S-401 datasets must conform to all mandatory elements of Annex A – Encoding Guide, where the word 'must' is used

In addition to the above, dataset usability must be assessed against:

- Intended user requirements in regard to coverage, scale and specific content requirements as defined by the Producing Agency and key stakeholders;
- Conformance to established maintenance processes (see Section 8); and
- Overall compliance with the S-401 Product Specification, including context-specific evaluation
 of individual encoding instances for requirement of conformance to checks classified as "Error"
 and "Warning" in Annex C IENC Validation Checks.

For dataset integrity requirements, see clause 11.6.

6.1.2 Bathymetric Data Quality

Bathymetric data quality comprises the following:

completeness of data (For example, waterway coverage). currency of data (For example, temporal degradation); uncertainty of data; source of data:

Data quality can be encoded at three different metadata levels (dataset, feature, feature instance). All positional (2D), vertical (1D), horizontal distance (1D) and orientation (1D) uncertainty attributes concern



International Hydrographic Organization 3.2.7 S-421

3.2.7.1 DQ chapters of S-421

There is no DQ chapter in S-421.

3.2.17.2 Result of cross check

Data Quality Measure	S-421
1.Completeness	N
2.Conceptual consistency	N
3. Domain consistency	N
4. Format consistency	N
5. Topological consistency	N
6. Positional Accuracy	N
7. Thematic Accuracy	N
8. Temporal Quality	N/A
9. Aggregation	N
10. Introduction to DQ Paragraph	N

Y=YES, N=NO, N/A=not applicable

NOTES: -.

3.2.7.3 DQ elements included in the other parts of S-421

No DQ element is included in the other parts of S-421.



4. CROSS-CHECK MATRIX

International Hydrographic Organization

Data Quality Measure	S-124	S-130	S-131	S-201	S-240	S-401	S-421
1.Completeness	N	Y	Y	N	Y	N	N
2.Conceptual consistency	N	Y	Y	N	Y	N	N
3. Domain consistency	N	Y	Y	N	Y	N	N
4. Format consistency	N	Y	Y	N	Y	N	N
5.Topological consistency	N	Y	Y	N	Y	N	N
6. Positional Accuracy	N	Y	Y	Y	Y	N	N
7. Thematic Accuracy	N	Y	Y	N	Y	N	N
8. Temporal Quality	N	N/A	Y	N	Y	N	N/A
9. Aggregation	N	N/A	Y	N	Y	N	N
10. Introduction to DQ	Y	Y	Y	Y	Y	Y	N
Paragraph							

Y=YES, N=NO, N/A=not applicable



4. CROSS CHECK MATRIX

International Hydrographic Organization

NOTES:

- a)There is no DQ chapter in S-421.
- b) Similar to the previously reviewed product specifications, DQ chapters of S-124, S-130, S-131, S-201, S-240 and S-401 are not in a harmonized way.
- c) DQ chapters of S-124, S-131, S-201, S-240 and S-401 do not conform to the S-97.
- d) S-124, S-201 and S-401 do not implement the table C-7-1 Recommended data quality measures of S-97.



5. RECOMMENDATIONS

International Hydrographic Organization It is recommended to:

- Feedback comments and suggestions to developers of S-124, S-130, S-131, S-201, S-240, S-401 and S-421.
- Continue review of new and amended Product Specifications.
- Invite other S-100 product specifications to join the review.



6. ACTIONS REQUESTED FROM THE SUBWG

International Hydrographic Organization

The DQWG is requested to:

- a. **Note** the information provided;
- b. **Approve** the recommendations above.