

Paper for consideration by TSMAD
Data Quality Model Harmonization

Submitted by:	DQWG and SNPWG
Executive Summary:	This paper updates TSMAD on the data quality model for nautical charts and nautical publications.
Related Documents:	(1) S-100 Ed. 2.0.0 (draft) (2) S-101 DCEG (Baseline version – April 2014)
Related Projects:	(1) S-101

1 Introduction/Background

SNPWG developed a model of data quality for IHO product specifications in the nautical publications domain (NPUBS) that was based on the model for S-101 ENC's as developed by DQWG and subsequently modified by TSMAD. The NPUBS data quality model was very similar to both the DQWG and S-101 data quality models but included some extensions and adaptations. These differences were discussed at the recent DQWG9 meeting, and the recommendation from SNPWG that there should be a harmonized S-101 and NPubs data quality model was elaborated and agreed upon by DQWG, as DQWG noted that that a common harmonized data quality model would be advantageous in the future when several data streams like ENC, Surface Current, Nautical Pubs could be overlaid one another and a data quality representation of the image on the screen would be generated.. This paper reports on the S-101/NPUBS data quality model and its differences from the previous DQWG and S-101 models.

2 References

ISO 19115-1: Geographic Information – Metadata – Part 1: Fundamentals, 2014-04-01.

3 Discussion/Analysis

3.1 Nautical Publications data quality model

This version of the data quality model updates the original NPUBS data quality model with:

- The most recent available UML diagram (post-DQWG9) of the DQWG model for S-101 data quality;
- Data quality meta-features **Quality of Bathymetric Data**, **Quality of Nonbathymetric Data**, and **Quality of Survey** from the April 2014 DCEG baseline for S-101;
- Changes to enumerated attribute **category of temporal variation**, resulting from SNPWG 17, SNPWG Letter 03/2014 (2 July 2014) and the subsequent email discussion;
- Updates of the commonly used complex attribute **source indication** and its sub-attributes, discussed at and after SNPWG 16;
- Changes to attributes used for specifying scope, arising from the most recent version of ISO 19115 (published in April 2014).

The result is shown in the figure below. The structure of the model is similar to the DQWG structure but adds two attributes to **Quality of Bathymetric Data** and **Quality of Nonbathymetric Data**. Detailed specifications are provided in an accompanying draft proposal for changes to S-101 feature and attribute types.

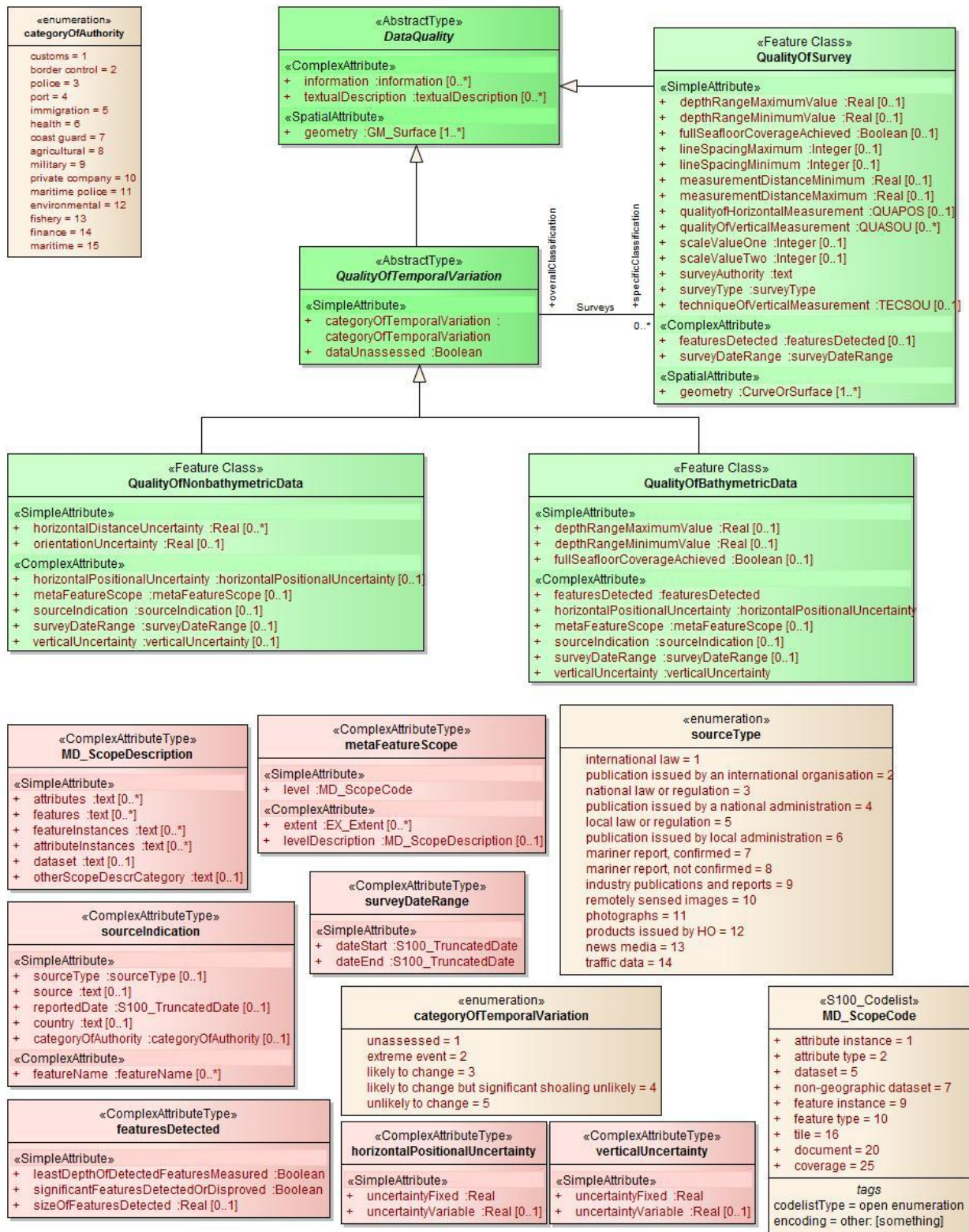


Figure 1. Data quality features and attributes for NPUBS (updated after DQWG9)

We expect that NPUBS data is likely to use **Quality of Nonbathymetric Data** rather than **Quality of Bathymetric Data** or **Quality Of Survey** but are retaining all three features in case of future need (or for use with context features – see section 3.2).

3.1.1 Source indication

Complex attribute **sourceIndication** is the successor to the S-57 attribute **source indication** (SORIND). Its function is to allow encoders to do what its name suggests, i.e., provide information about the source of information to which it applies. (E.g., the name of the rule-making agency can be indicated in sub-attribute **featureName**, etc.) Source indication is considered necessary for NPUBS data sets because assessment of publications information quality by mariners is often non-quantitative, taking the identity of the source into account, and cannot be captured using the numerical measures of uncertainty which are appropriate in ENC data (which is largely numeric, compared to nautical publications information which is often textual or categorical, e.g., rules based on ship and cargo characteristics).

3.1.2 Scope attributes

Complex attribute **metaFeatureScope** allows instances of the quality features to restrict their scopes to named feature or information classes or attributes. For example, an instance of **Quality of Nonbathymetric Data** can be restricted to all instances of **Marine Protected Area** features and only to instances of **Marine Protected Area** features by setting

QualityOfNonbathymetricData.metaFeatureScope.levelDescription = {MarineProtectedArea}

A consequence of this functionality is that data quality features of the same class may coincide or overlap. To resolve ambiguities in overlap areas, overlapping features of the same class must either apply to disjoint sets of features, or a preference order must be defined to determine which of two (or more) overlapping data quality features of the same class overrides the other(s). The figure below illustrates the concept, with a Quality feature scoped as applying to specified instances overrides Quality features scoped for feature classes which in turn overrides quality features without scope attributes (i.e., default scope).

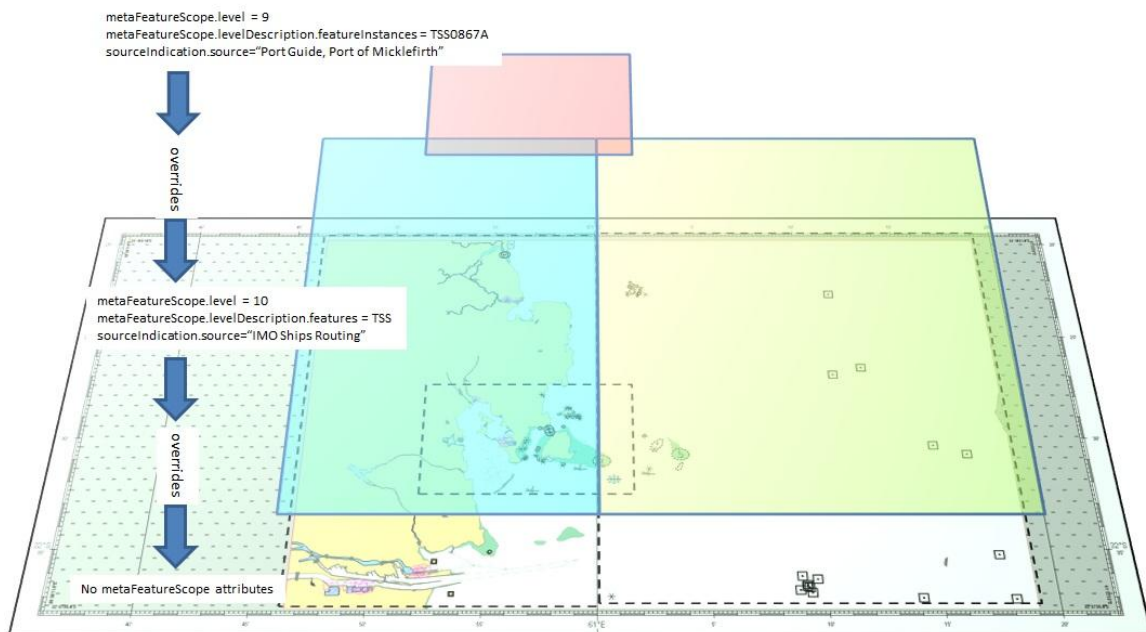


Figure 2. Scoped quality features

3.1.3 Listed values and definitions

Noting that while the April 2014 S-101 baseline binds **category of temporal variation** to **Quality of Bathymetric Data** only and that the S-101 definitions probably reflect this, SNPWG proposes to conform to the original DQWG model in which both **Quality of Bathymetric Data** and **Quality of Nonbathymetric Data** inherit the attribute **category of temporal variation** from a common super-type. Accordingly SNPWG recommended some changes to

the listed values of **category of temporal variation**, which changes need to be harmonised with DQWG and TSMAD:

Label	S-101 definition	proposed definition	Comments
un-assessed	temporal variation not assessed	(No change.)	No change to the definition.
extreme event (SNPWG change: insert “extreme” before “event”)	no new survey conducted after an event (e.g. hurricane, earthquake, volcanic eruption, landslide, etc), which is considered likely to have changed the seafloor significantly	No new survey conducted after an event (e.g. hurricane, earthquake, volcanic eruption, landslide, etc.), which is considered likely to have resulted in significant change .	Any occurrence, significant or not is an “event”. Adding “extreme” to the label conveys that changes may be very important for navigation purposes. Generalizes definition so it can be used for changes to other types of information as well as bathymetry.
likely to change	Continuous or frequent change (e.g. river siltation, sand waves, seasonal storms, ice bergs , etc).	Continuous or frequent change (e.g. river siltation, sand waves, seasonal storms, construction , etc.).	Construction also has the potential to cause changes at the location, including to bathymetry.
(DQWG: likely to change, but significant shoaling not expected) (S-101: likely to change, but significant shoaling unlikely)	(not yet defined)	Significant horizontal change, but vertical change is minimal.	Definition proposed by SNPWG18.
unlikely to change	significant change to the seafloor is not expected	Significant change not expected.	Generalizes definition so it can be used for changes to other types of information as well as bathymetry.

3.2 Context features and data quality

Some NPUBS product specifications allow for NPUBS exchange sets to be used either as an overlay in conjunction with S-101 ENC's (or similar products) or as standalone datasets, i.e., without an ENC or similar layer. Application schemas for some NPUBS product specifications will therefore include ‘context features’. These are generally S-101 features which provide background and context to the main NPUBS data sets, for example, land and water areas that provide the appropriate graphical background or marks indicating the boundaries of protected areas.

Distinct instances of data quality feature instances are likely for the ‘domain’ and ‘context’ features since they are derived from different sources, and since context features are not core information in NPUBS datasets. Since the bounding extents of ‘domain’ and ‘context’ features in NPUBS exchange sets will coincide or at least overlap, the meta-feature polygons for ‘domain’ and ‘context’ features will also coincide or overlap.

3.3 Text attributes for data quality features

While this model provisionally includes **information** and **textualDescription** as attributes of the “Quality of...” features, we believe text attributes are rarely if ever needed for data quality features, and their use should be discouraged or prohibited in these features to minimize distractions for both cartographer and end-user.

Furthermore, should SNPWG decide to retain them as attributes of the “Quality of...” features, they may be replaced by either the current S-101 approach to text (information type **SupplementaryInformation**) or SNPWG’s proposed simplification of it (either a complex attribute merging information and **textualDescription**, or a fuller model of text content).

Discussions on the modelling of text attributes are currently in progress in SNPWG and should be completed by the time DQWG9 commences, which means a proposed model of text attributes should be available for discussions at DQWG9 and with TSMAD at TSMAD29 in February 2015.

3.4 Quality model for individual spatial objects

The data quality model for individual spatial objects is similar to that in the April 2014 baseline S-101 DCEG and is shown in the figure below.

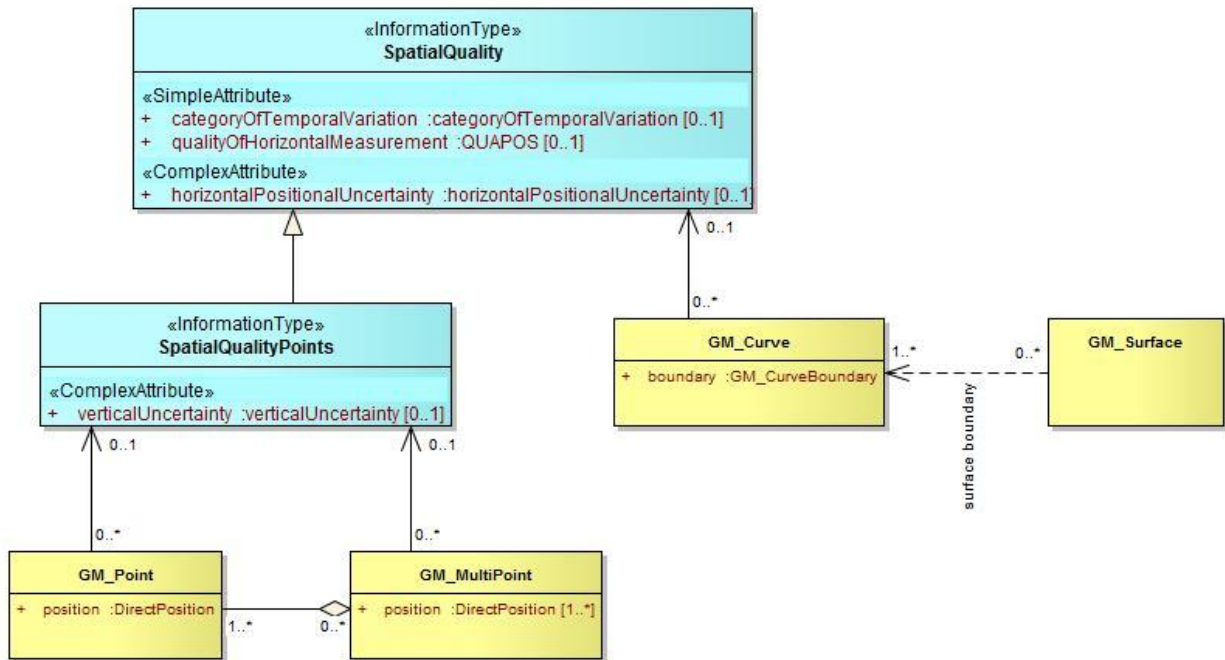


Figure 3. Quality for individual spatial objects (updated after DQWG9)

3.5 Portrayal

The quality portrayal algorithm will need to be updated to take into account the preference order of intersecting quality features as well as the ability to indicate quality down to the feature class, individual feature, and attribute levels. The outline of the approach should be:

- Quality indicators in geographic regions where quality features of the same type intersect (i.e., two **Quality of non-bathymetric data** features) should use the preference order specified in the accompanying proposal to determine which quality indicator overrides the others.
- Quality features applying to only feature or attribute types which have been suppressed by the user should not be taken into account when determining the quality indicators.
- The DQWG rule about showing the lowest-grade quality indicator applicable continues to apply.

The question of when, how and where to display indicators which have no intrinsic order (in particular, the contents of **source indication** attributes) remains open at this time, but off-graphic display, pick-report display, or other means of on-request, non-intrusive display of text should suffice for such indicators.

4 Conclusion

The proposed changes to the S-101 data quality model result in a data quality model that is more generalized and can be used in common across multiple IHO product specifications including nautical publications datasets as well as giving it the ability to capture quality characteristics at both coarse-grained and fine-grained levels, i.e., ranging from the dataset level down to feature class and attribute levels.

5 Actions Requested

TSMAD is invited to:

- endorse the resulting data model as the common data quality model for S-101
- interact with DQWG and SNPWG to maintain a harmonized data quality model to the extent possible