Preparatory Actions for
European Marine Observation and Data Network

High Resolution Seabed Mapping
WP1: Data provider contribution

Completing metadata elements for the generation of the Quality Index for the EMODnet DTM

Service Contract No. EASME/EMFFM2016/005

Document history	Date	Contributors
Initial Draft	19/05/2017	T. Schmitt with ideas from B. Loubrieu, C. Pertuisot, C. Monpert, L. Dorst
Corrections including better definition of the QI_Vertical and Horizontal Classes. Definition of the QI_Purpose	22/05/2017	T.Schmitt, B. Loubrieu, C. Monpert, C. Pertuisot
Implementation in Mikado	23/05/2017	C. Pertuisot, B. Loubrieu
Minor correction (XML layout as Annex following D. Schaap comments)	31/05/2017	T. Schmitt
Implementation in Sextant catalogue (CPRD)	01/06/2017	C. Pertuisot, B. Loubrieu
Editing document	14/06/2017	DMA Schaap
Editing the document	16/06/2017	G. Spoelstra

Table of contents

1. Introduction

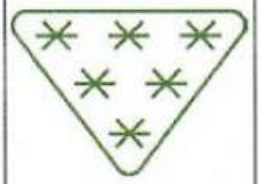
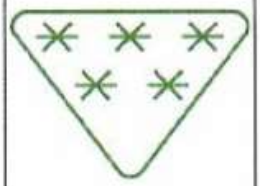
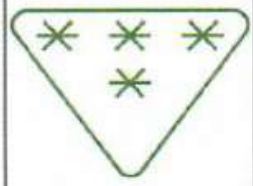
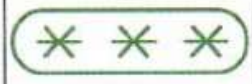
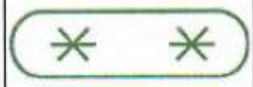
EMODnet bathymetry is composed of a multitude of datasets from a multitude of data providers. Users of the resulting grid and associated datasets need to be able to evaluate at the grid node level the quality of the bathymetric data and product they will be using.
Prior work has been done to provide an estimate of a quality index based on age of the survey and number of soundings per unit of areas. It is proposed to use further qualitative information related to the data source (such as type of sensor) to better define a quality index (QI).
The aims of such a quality index are to:

- help data users to evaluate quickly the dataset they are about to request
- indicate to the Basin coordinators what are the limitations of the dataset they are about to merge while building the EMODnet DTM
- be used as the basis of the evaluation of the quality of the EMODnet DTM

Prior to proposing an approach adapted to the EMODnet Bathymetry community (provider and users), we will provide notes on the existing CATZOC (Category of Zone of Confidence) which will be our source of inspiration.
We provide a framework for the generation of the quality indicator that we want to be homogeneous between the different types of data, relatively easy to implement (with respect to the metadata and statistical attributes readily available during the generation of the metadata and the data products), meaningful to the users and coherent with expert knowledge. In this paper we want to describe the elements requested from the data provider. In a future document we might describe how to use these elements to select and merge the datasets in the EMODnet DTM and the computation of an associated Quality indicator.

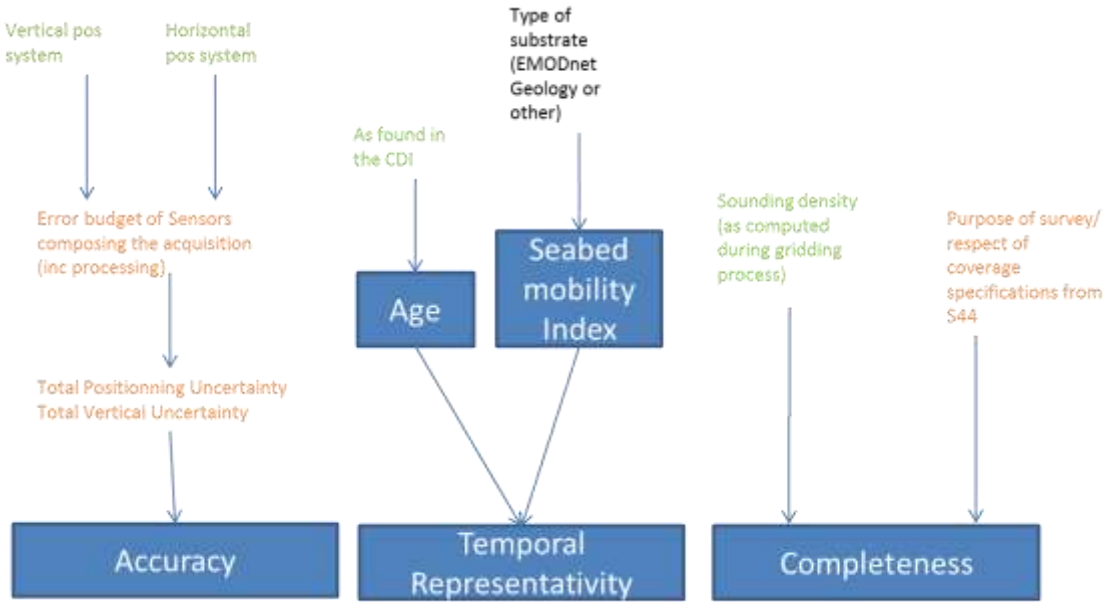
2. Existing approach – general understanding of the CATZOC

The CATZOC (CATegory Zone Of Confidence) is a an IHO categorization of the level of accuracy of bathymetric data. It aims at providing qualitative indications on the uncertainties attached to bathymetric data underlying the paper charts or ENCs. The primary intention of the CATZOC is for the chart/ENC users to assess how confident one should be with respect to representation of obstacles to navigation on the navigation documents.
In order to do so, Hydrographic Offices, mainly rely on elements of uncertainty on the vertical and horizontal positions of the sounding, the sampling strategy (density) and potential temporal variation of the seafloor supposing to have happened since the acquisition. Those attributes are gathered through metadata associated per surveys (POSACC, SOUACC, TECSOU, SUREND, etc associated under the M_QUAL S-57 list of attributes).
The table below describes the recent implementation of this ZOC categorization.

1	2	3		4	5	6
ZOC	Position Accuracy	Depth Accuracy		Seafloor Coverage	Typical Survey Characteristics	Symbol
A1	± 5 m	= 0.50 + 1%d		Full area search undertaken. All significant seafloor features detected and depths measured.	Controlled, systematic survey, high position and depth accuracy achieved using DGPS or a minimum three high quality lines of position (LOP) and a multibeam, channel or mechanical sweep system.	
		Depth (m)	Accuracy (m)			
		10 30 100 1000	± 0.6 ± 0.8 ± 1.5 ± 10.5			
A2	± 20 m	= 1.00 + 2%d		Full area search undertaken. All significant seafloor features detected and depths measured.	Controlled, systematic survey achieving position and depth accuracy less than ZOC A1 and using a modern survey echosounder and a sonar or mechanical sweep system.	
		Depth (m)	Accuracy (m)			
		10 30 100 1000	± 1.2 ± 1.6 ± 3.0 ± 21.0			
B	± 50 m	= 1.00 + 2%d		Full area search not achieved; uncharted features, hazardous to surface navigation are not expected but may exist.	Controlled, systematic survey achieving similar depth but lesser position accuracies than ZOC A2, using a modern survey echosounder, but no sonar or mechanical sweep system.	
		Depth (m)	Accuracy (m)			
		10 30 100 1000	± 1.2 ± 1.6 ± 3.0 ± 21.0			
C	± 500 m	= 2.00 + 5%d		Full area search not achieved, depth anomalies may be expected.	Low accuracy survey or data collected on an opportunity basis such as soundings on passage.	
		Depth (m)	Accuracy (m)			
		10 30 100 1000	± 2.5 ± 3.5 ± 7.0 ± 52.0			
D	worse than ZOC C	Worse Than ZOC C		Full area search not achieved, large depth anomalies may be expected.	Poor quality data or data that cannot be quality assessed due to lack of information.	
U	Unassessed – The quality of the bathymetric data has yet to be assessed					

3. Proposed approach

Recognizing the fact that all data contributors of the EMODnet HRSM project do not have necessarily the ability to provide a CATZOC value associated with all their datasets, the intent of the proposed approach is to get inspired by this classification although simplifying it.
As a matter of fact, 3 main parameters will be used to compute the HRSM Quality Index (QI): the accuracy (vertical, horizontal), the temporal representativity, and the completeness of the survey/sampling of the seabed. For each of the parameters an integer value will be given (see below for details). The score reached by a dataset will be the concatenation of each value. For example a recent shallow single beam with a natural GPS survey with a poor density will be coded like 2231.

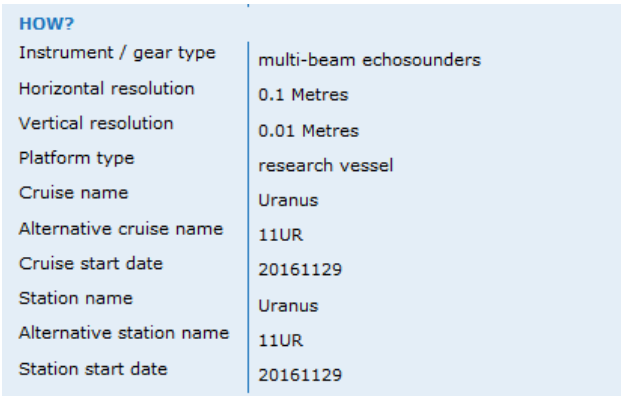


3.1. Approach in the Common Data Index (CDI) metadata

The CDI metadata format is used in EMODnet HRSM to describe survey data sets. The CDI has several elements which are important for the later computation of the Quality Index for individual surveys.

3.1.1. Filling the accuracy component

The accuracy is defined by both the horizontal and the vertical part. In the best case, data providers can provide an estimate of their vertical and horizontal accuracy **in the "Horizontal resolution" and "Vertical resolution" fields in the How section** (see Figure on the right).
The term resolution is not fully adapted for bathymetry. It originates from the global vocabulary of Seadatanet.



In all cases, it will be requested to fill the **QI_Horizontal and QI_Vertical** (new fields that will be added to the Other Info section→Quality info section (see Figure below) within the existing CDI Schema).

OTHER INFO

Quality info

Name	Date	Comment
COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata	2008-12-04	See the referenced specification
IHO S-44	2016-12-20	Validated survey

For the **QI_Horizontal** component, the proposed classification is:

0: Unknown or > 500m (That is grossly equivalent to TACAN, OMEGA systems or similar)

1: between 500m and 50m (That is grossly equivalent to LORAN, DECCA systems or similar)

2: between 50m and 20m (That is grossly equivalent to natural GPS systems)

3: < 20m (GPS with correction) (That is grossly equivalent to aided GPS system DGPS, RTK …)

For the **QI_Vertical** component, the proposed classification is based on the sounding measurement devices:

0: Unknown, plummet, leadline

1: SBES Low Frequency, SDB (similar than 2+5%d)

2: MBES low frequency (lower than 100kHz) (similar than 1+2%d)

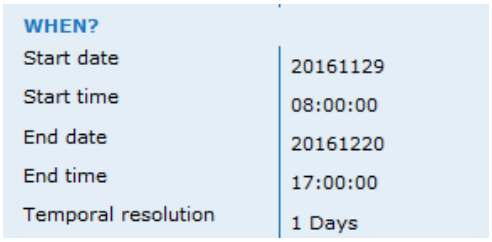
3: Lidar, SBES High Frequency

4: MBES High frequency (higher that 100kHz) (1+0.5%d)

The data provider should complete the existing field Horizontal and Vertical resolution (when possible and in meters) and name and complete the QI_Horizontal and QI_Vertical fields.

3.1.2. Evaluating the temporal representativity

QI_Age will be calculated from the age of the survey. This is defined as the age in years (integer value) between the date of the EMODnet DTM release and the start date (to consider the worst case).



The data providers will only have to make sure this section is properly filled, with particular care on the Start date value.

The objective of this indicator will be ultimately to highlight the probability that the measured seafloor corresponds to the present day seafloor. The classification below tries to grasp morpho-dynamic time frame of processes that can affect the seabed with measurable consequences bigger than 100m resolution:

0: 30y – oldest date (geological structural, tidal basin changes,)

1: 10y – 30y

2: 5y – 10y (erosion/deposition at the scale of structure like continental shelf / canyons…)

3: 0y – 5y (time frame of dune migration or coastal shoreface modifications)

Note: as suggested above and if EMODnet Geology allows it, the age of the survey, as computed here, will be compared to the level of mobility that is expected from the nature of the seabed (eg. an old survey on a rocky area has not the same meaning as an old survey in a highly mobile area such as a sandy seabed).

3.1.3. Evaluating the purpose of the survey

This field describes what the objectives of the survey were. It describes both elements of seabed sampling and accuracy reached through data processing. For the **QI_Purpose** component, the proposed classification is:

0: Purpose of the survey unknown (historical survey with no associated information)

1: Transit and/or opportunity

2: Bathymetric/morphologic survey

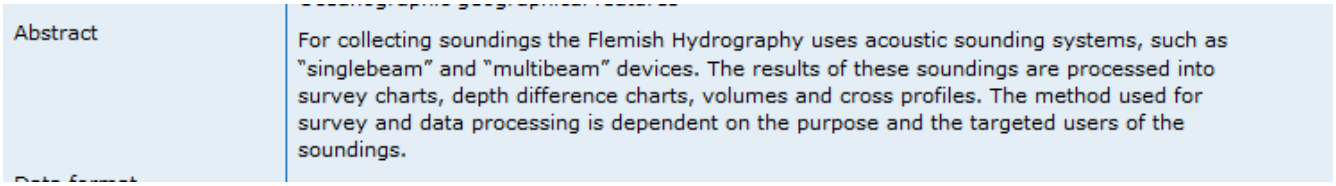
3: Hydrographic survey or compatible with hydrographic standards

The data provider will fill the new QI_Purpose field.

3.1.4. Filling the abstract section

This field is a free text area. It is strongly suggested that the data provider describes here elements that cannot be described elsewhere such as the purpose of the survey, the survey conditions, some processing considerations such as tide or SVP related.

Note that this section will not be used in the QI computation, but it can strongly help the users of the



dataset to better understand the limitations of the dataset.

3.1.5. Computing the Quality Index

As mentioned above the Index will be composed as concatenation of individual score per components as follows.

QI_Horizontal:QI_Vertical:QI_Age:QI_Purpose

A Quality Index value will be computed in a second time using the elements provided in the string above, along with intrinsic local properties of the DTM (number of soundings per grid node, interpolation yes/no, GEBCO, …)

3.2. Approach in the Sextant catalogue – Composite Product (CPRD)

In the case of composite DTM product (composed of a series of surveys), the logic and the Quality Index remain the same. **However, the data producer will have to consider giving each of the quality indicator based on the contribution with the lowest quality.**

E.g. Suppose that your composite grid includes multiple surveys including some positioned using aided GPS (QI_Horizontal=3) and some positioned using LORAN, or similar (QI_Horizontal=1), the resulting QI_Horizontal for the composite DTM will be 1.

Note that while you can decide to provide a composite DTM grid, we strongly recommend that all datasets composing the composite grid are detailed using individual CDI sources. In that case, the list of CDI identifiers associated to the Composite Product must be filled in the field "Data source description". This field is a free text field. Therefore, in this case, the elements provided through the CDI will be used to define the quality index of the CPRD.

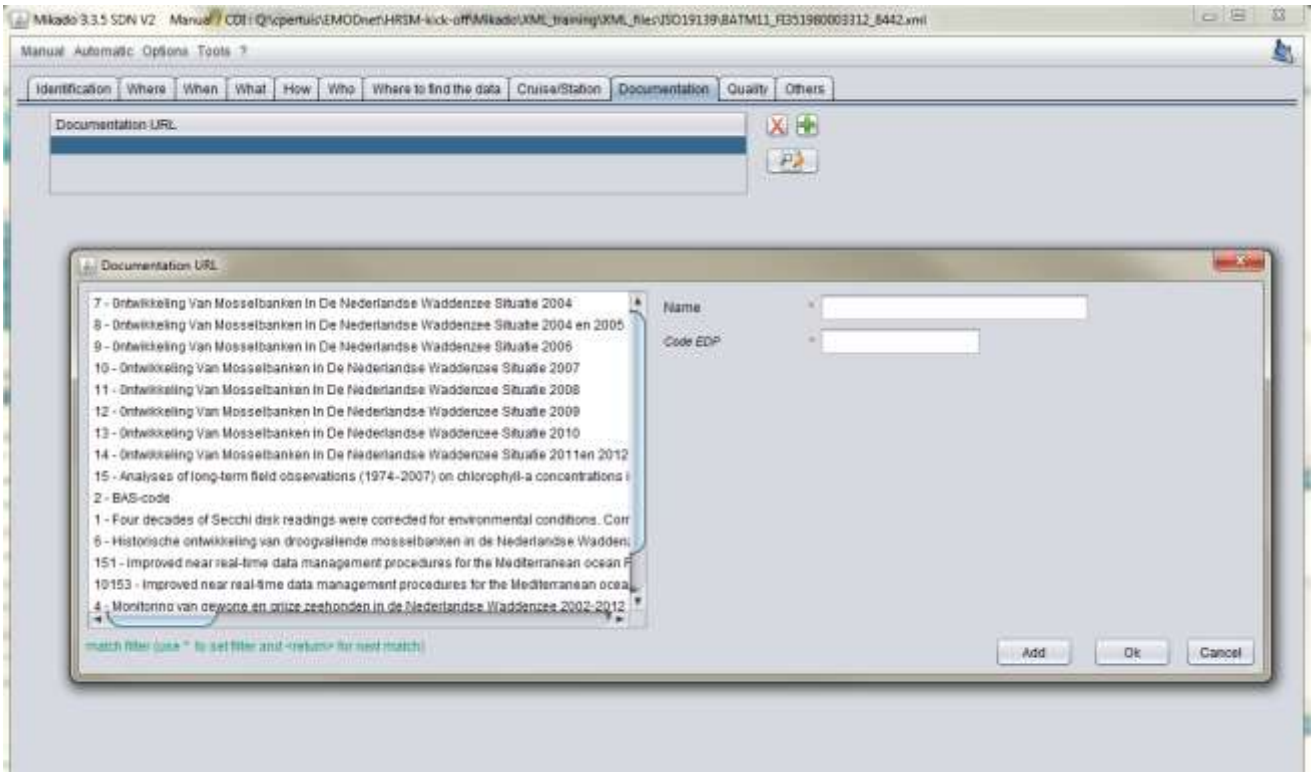
4. Quality Index Metadata Implementation

The CDI Schema will not change; the CDI XML format will change in this way that the three new indicators QI_Horizontal, QI_Vertical and QI_Purpose should be defined as fields and completed by the Data Providers. Moreover it is advised for the data providers to pay attention to completing the Abstract, Horizontal and vertical fields.

4.1. Implementation in the CDI using Mikado – Manual Mode

4.1.1. Filling the Document Reference information

In the Documentation tab, add an entry and search for the referenced document EMODnet_Quality_Index.pdf.

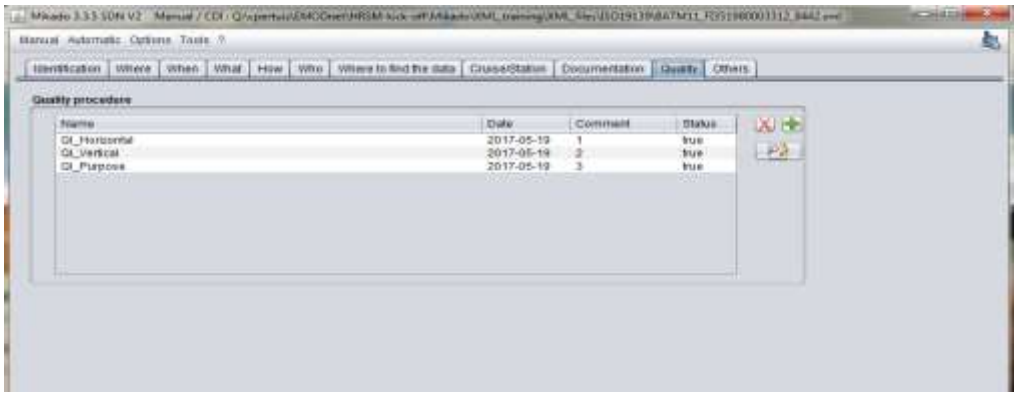


4.1.2. Filling the QI_Horizontal, QI_Vertical, QI_Purpose

In the Quality tab, add an entry and input:

Name: QI_Horizontal
Date: 2017-MM-DD (Publication Date of the referenced EMODnet_Quality_Index.pdf)
Comment: select the index corresponding to the indication given in the EMODnet_Quality_Index.doc.
Note: See page 7 for the values to be used.
Status: true

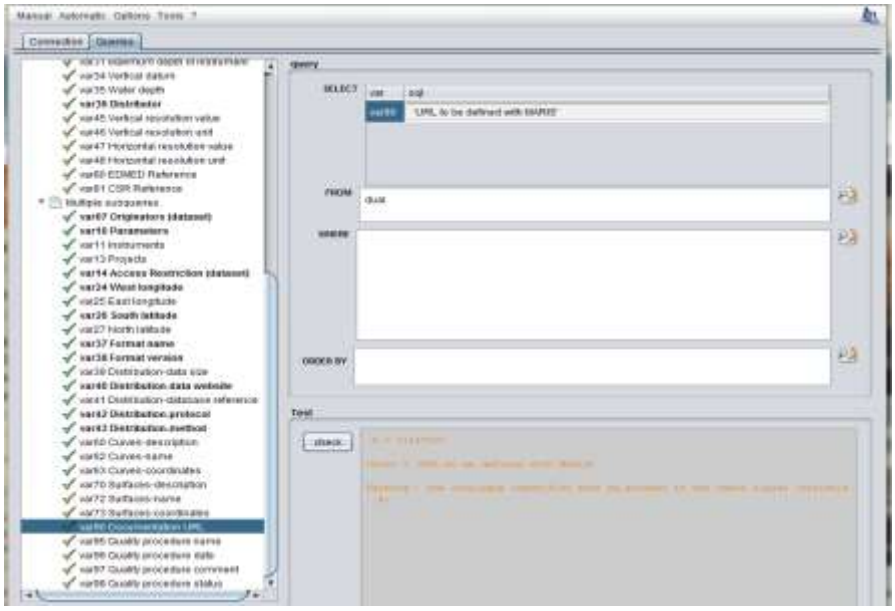
Do the same for QI_Vertical and QI_purpose. The three inputs should look like:



4.2. Implementation in the CDI using Mikado – Automatic mode

4.2.1. Filling the Document Reference information

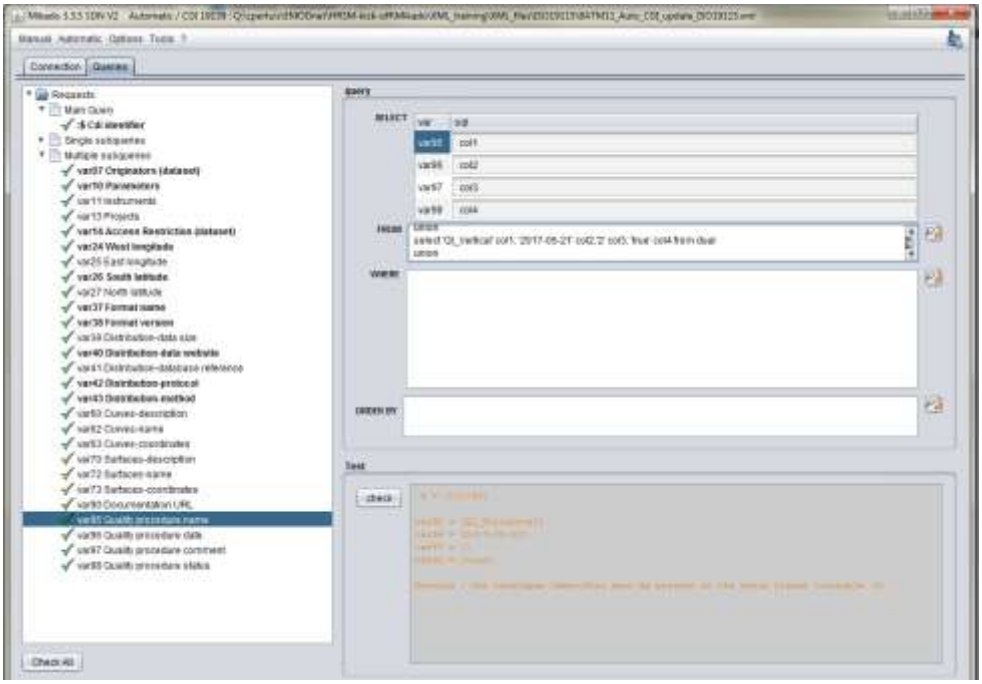
Under the multiple sub-queries folder enter your SQL query under var90: indicate the URL of the validated document EMODnet_Quality_Index.pdf to be given by MARIS.



4.2.2. Filling the QI_Horizontal, QI_Vertical, QI_Purpose

Under the multiple subqueries folder define your SQL queries under var95 to var98 to describe your selected datasets and the corresponding quality indexes given above (§3). None of these variables are part of any SDN list, they have to be written "in hard" in your SQL query. An example of a possible SQL request is:

```
select col1, col2, col3,col4
from (select 'QI_Horizontal' col1, '2017-05-21' col2,'1' col3, 'true' col4 from dual union
      select 'QI_Vertical' col1, '2017-05-21' col2,'2' col3, 'true' col4 from dual union
      select 'QI_Purpose' col1, '2017-05-21' col2,'3' col3, 'true' col4 from dual)
```



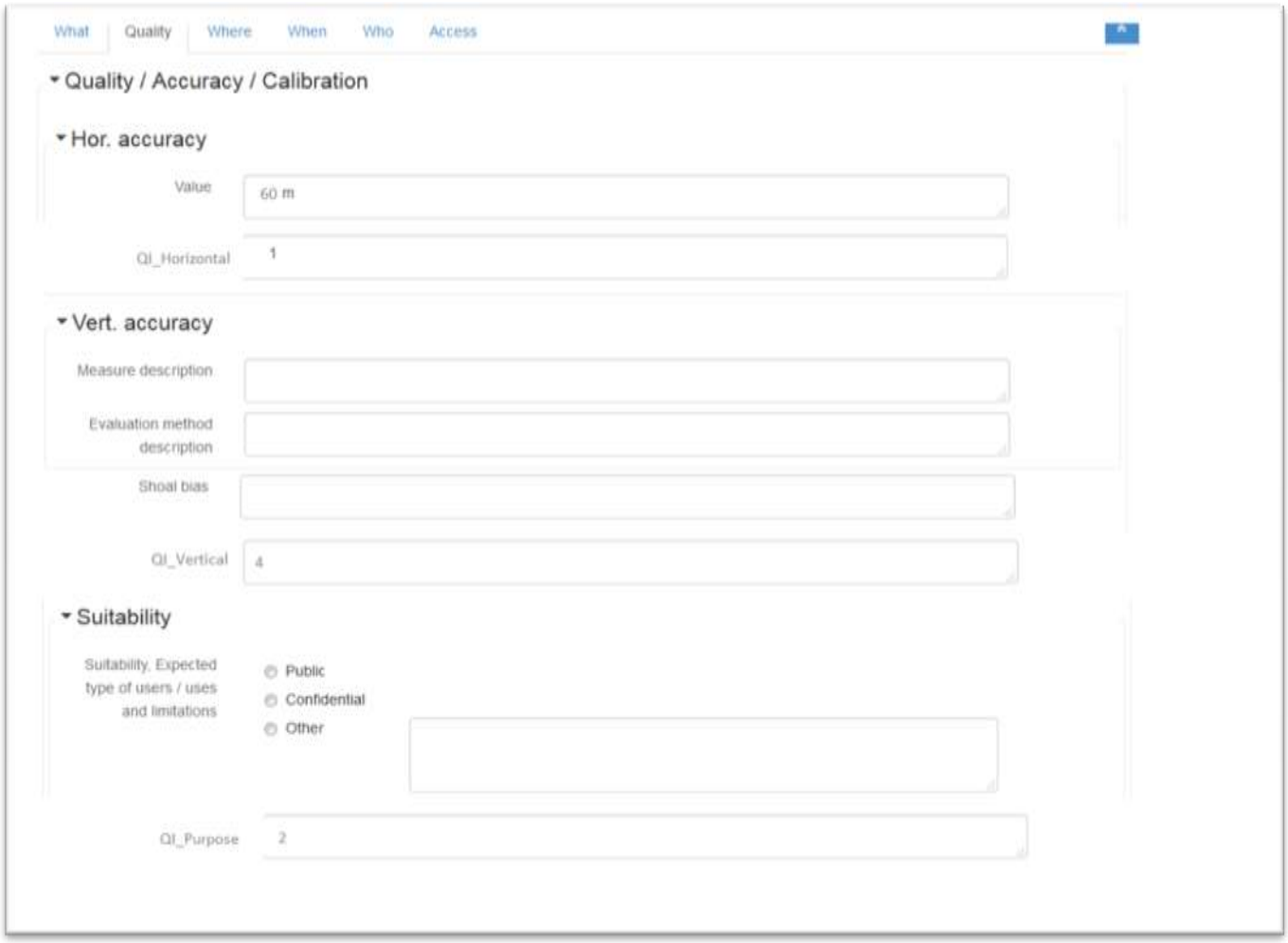
4.3. Implementation in the CPRD using Sextant

4.3.1. Filling the Document Reference information

The data provider will not have to indicate this information: the document reference will be filled in by default when editing a new entry in the CPRD catalogue and will be defined as an associated resource (like the data provider already does when indicating a website for example).

4.3.2. Filling the QI_Horizontal, QI_Vertical, QI_Purpose

Under the Quality tab, choose in the proposed list the correct value for each of the corresponding Quality Index. The data producer will have to consider giving each of the quality indicator based on the contribution with the lowest quality.


What
Quality
Where
When
Who
Access
Quality / Accuracy / Calibration
Hor. accuracy
Value
60 m
QI_Horizontal
1
Vert. accuracy
Measure description
Evaluation method description
Shoal bias
QI_Vertical
4
Suitability
Suitability. Expected type of users / uses and limitations
Public
Confidential
Other
QI_Purpose
2

ANNEX

XML layout

For those interested to know how the CDI XML file is impacted, the only changes that will be expected are related to the accuracy and the purpose parameters. Those will be added up into the Quality info section. A proposal for the corresponding section for the CDI xml file is given below for the QI_Horizontal. Likewise similar sections for the QI_Vertical and QI_Purpose_are expected.

IMPORTANT NOTE: The section "documentation" of the xml must reference a valid and registered document. When the present document will be approved, we will reference it.

```
<gmd:report>
        <gmd:DQ_DomainConsistency>
          <gmd:result>
            <gmd:DQ_ConformanceResult>
              <gmd:specification>
                <gmd:CI_Citation>
                  <gmd:title>
                    <gco:CharacterString>QI_Horizontal</gco:CharacterString>
                  </gmd:title>
                  <gmd:date>
                    <gmd:CI_Date>
                      <gmd:date>
                        <gco:Date>2017-05-19</gco:Date>
                      </gmd:date>
                      <gmd:dateType>
                        <gmd:CI_DateTypeCode
codeList="http://vocab.nerc.ac.uk/isoCodelists/sdnCodelists/gmxCodeLists.xml#CI_DateTypeCode"
codeListValue="publication"  codeSpace="ISOTC211/19115" >publication</gmd:CI_DateTypeCode>
                      </gmd:dateType>
                    </gmd:CI_Date>
                  </gmd:date>
                </gmd:CI_Citation>
              </gmd:specification>
              <gmd:explanation>
                <gco:CharacterString>1</gco:CharacterString>
              </gmd:explanation>
              <gmd:pass>
                <gco:Boolean>true</gco:Boolean>
              </gmd:pass>
            </gmd:DQ_ConformanceResult>
          </gmd:result>
        </gmd:DQ_DomainConsistency>
      </gmd:report>
```