Preparatory Actions for
European Marine Observation and Data Network

High Resolution Seabed Mapping
WP2 : Generate indicators in the DTM

Use of the dataset Quality Index to expand services associated to the EMODnet DTM

Service Contract No. EASME/EMFFM2016/005

12 February 2019

Authors: T. Schmitt, B. Loubrieu, C. Guerin, and C. Monpert

Document history	Date	Contributors
Initial version	12/02/2019	T. Schmitt, B. Loubrieu, C. Guerin, C. Monpert

Table of contents

1. Introduction

EMODnet bathymetry is composed of a large number of datasets from many data providers. Users of the resulting grid and associated datasets need to be able to evaluate the quality of the bathymetric product and associated data they will be using, at the grid node level. For that purpose, it was requested from the data provider to provide a quality assessment of the data they are providing. Tools were developed to help them generating this information as part of the metadata. Using these qualitative information helped the basin coordinators to set the level of priority of a data source over another one in case of overlap . Likewise, these information are being used to generate cartographic layouts enabling the users to assess the confidence he/she can get locally from the bathymetric DEM product.

This document first introduces the methodology used to define four quality indexes (known as QI, in the following sections of this document) for each datasets. Then, in a second part, it describes how the information gathered from each data provider for each dataset is used both in the context of the generation of the bathymetric DEM product and in its qualitative assessment (at the level of the grid cell).

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).

Figure 1, below describes the recent implementation of this ZOC categorization.

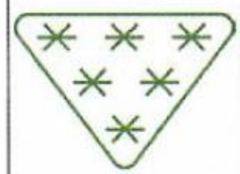
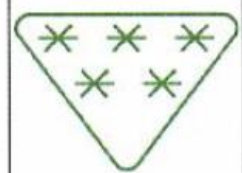
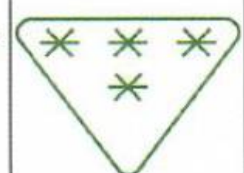
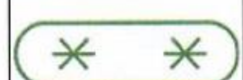
1	2	3	4	5	6
ZOC	Position Accuracy	Depth Accuracy	Seafloor Coverage	Typical Survey Characteristics	Symbol
A1	± 5 m	= 0.50 + 1%d Depth (m) / Accuracy (m) 10 / ± 0.6 30 / ± 0.8 100 / ± 1.5 1000 / ± 10.5	Full area search undertaken. All significant seafloor features detected and depths measured.	Controlled, 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.	
A2	± 20 m	= 1.00 + 2%d Depth (m) / Accuracy (m) 10 / ± 1.2 30 / ± 1.6 100 / ± 3.0 1000 / ± 21.0	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.	
B	± 50 m	= 1.00 + 2%d Depth (m) / Accuracy (m) 10 / ± 1.2 30 / ± 1.6 100 / ± 3.0 1000 / ± 21.0	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.	
C	± 500 m	= 2.00 + 5%d Depth (m) / Accuracy (m) 10 / ± 2.5 30 / ± 3.5 100 / ± 7.0 1000 / ± 52.0	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.	
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				U

Figure 1: Definition of the classes defined as part of the Category of Zone of Confidence used for hydrographic charting quality assessment.

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.

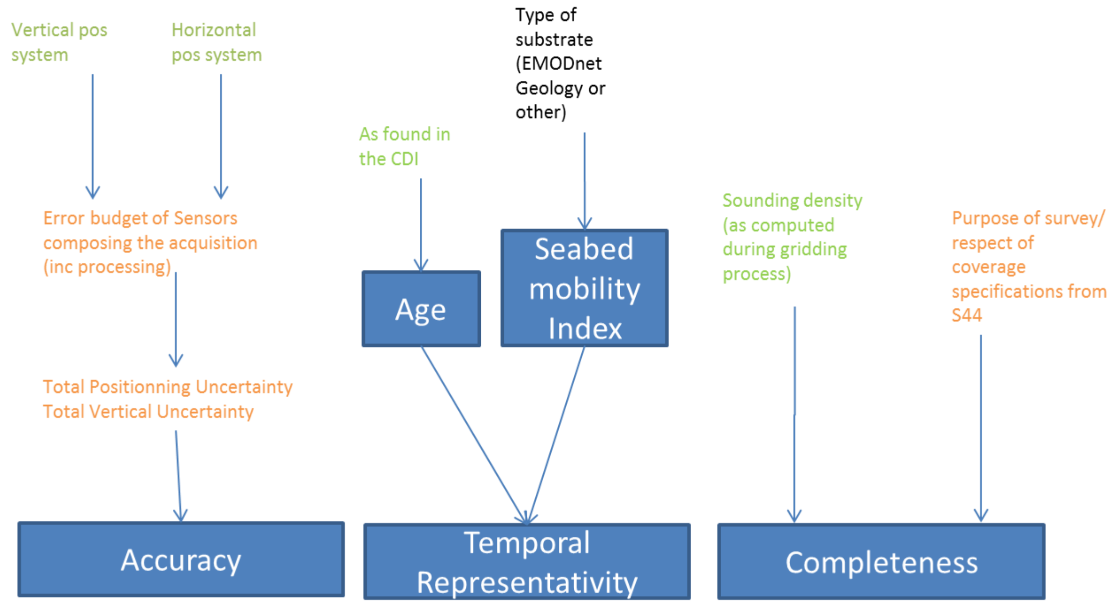


Figure 2: Conceptual description of the elements affecting the quality of bathymetric data.

Conceptually the quality of a bathymetric dataset can conceptually described as in Figure 2 by the following 3 main parameters:

- The accuracy of the survey, which is a function of both the vertical and horizontal precision of the sounding measurement. This, in turn, can be broke apart in from a description of the vertical positioning system and the horizontal positioning system.
- The temporal representativity or in other terms the consistency between the measurement (at the time when it was acquired) and the actual morphology of the seabed (e.g. bathymetric measurements cannot be considered accurate if they were surveyed, say ten years ago in a highly dynamic area, such as for sand dunes for example)
- The completeness or the sampling of the seabed, which provides some forms of confidence in the sounding measurement (as of a number of soundings by unit of space). This is often related to the survey conditions (speed of the boat, overlapping between adjacent lines, sea states, …)

Four proxy indicators summarizing the above parameters have been defined and described in the EMODnet HRSM document « **Completing metadata elements for the generation of the Quality Index for the EMODnet DTM** ». Table 1 provides the classification values for the providers to describe their dataset. They are known as "Quality Indices":

- Horizontal accuracy (QI_Horizontal)
- Vertical accuracy (QI_Vertical)
- Purpose of the survey (QI_Purpose)
- Age of the survey : the age of the survey is derived from the « start date » field of the metadata. (QI_Age)

QI_horizontal	QI_vertical	QI_age	Purpose of the survey Respect of a standard
-1 : Multisources – unable to assess	- 1 : Multisources – unable to assess		
0: Unknown or > 500m (That is grossly equivalent to TACAN, OMEGA systems or similar)	0: Unknown, plummet, leadline	0 : > 30 y	0 : Purpose of the survey unknown (historical survey with no associated information).
1: between 500m and 50m (That is grossly equivalent to LORAN, DECCA systems or similar)	1: SBES Low Frequency, SDB (similar than 2+5%d)	1 : 10-30 y	1: Transit and/or opportunity
2: between 50m and 20m (That is grossly equivalent to natural GPS systems)	2: MBES low frequency (lower than 100kHz) (similar than 1+2%d)	2 : 5y -10 y	2: Bathymetric/morphologic survey
3: < 20m (GPS with correction) (That is grossly equivalent to aided GPS system DGPS, RTK ...)	3: Lidar, SBES High Frequency	3 : 0y – 5y	3: Hydrographic survey or compatible with hydrographic standards
	4: MBES High frequency (higher that 100kHz) (1+0.5%d)		

Table 1: Description of the classes for each of the Quality Indicator

In the above mentioned document, the data providers were also able to see how to fill Seadatanet and Sextant metadata templates (Figure 3) for the metadata accompanying their dataset.

OTHER INFO

Quality info

Name	Date	Comment
Commission Regulation (EU) No 1089/2010 of 23 November 2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services	2010-12-08	See the referenced specification
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
QI_Horizontal	2017-07-27	3
QI_Purpose	2017-07-27	2
QI_Vertical	2017-07-27	2

Figure 3: Example of QI information provided by a data provider for an individual survey (CDI metadata sheet as it can be found on the Seadatanet portal)

Following the gathering of the metadata from all the providers either through Seadatanet for the CDI catalogue or Sextant for the CPRD catalogue several checks were undertaken in order to see how the information have been filled. For the datasets composing the DTM product, this qualitative assessment essentially pointed out:

- Inconsistent values (non-existing values in the Table 1) or unfilled values
- Errors in the classification

Consequently, data providers were asked individually to revisit the values they provided. In order to help them a cartographic layout of their dataset was provided.
This assessment ended up with an upgraded version of the QI tables, which is the basis for the Quality analysis of the global DTM.

4. Use of the QI as part of the DTM generation

As part of new tools proposed to the basin coordinators for generating the DTM, access from the catalogues to the individuals QI per source of data has been implemented in the Globe software.
Using these enable the coordinators to select data sources, or even better to set a priority level for each of the sources in the case of an overlap. The Globe data explorer has now several options for sorting each of the data sources using one or several QI attributes when generating the DTM or for merging them with similar QI value (Figure 4).

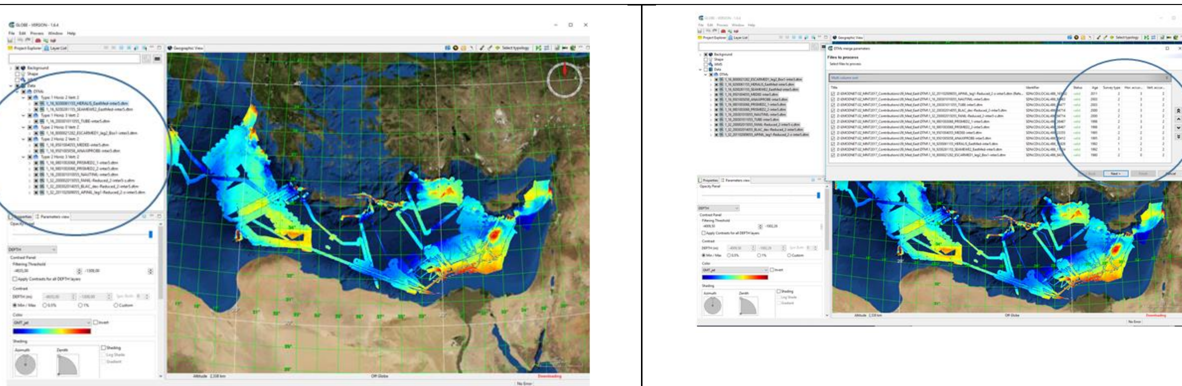


Figure 4: Examples of use (and associated implementation in the GLOBE software) of the QI information as part of the selection/level of priority of superposed surveys as part of the DTM compilation by the basin coordinators.

5. Geographic display of the individual indicators

In the present section, each of the quality indicators is displayed one by one both at the scale of the overall EMODnet DTM product and at a local level in the southern North Sea. The source layer which is generated as part of the DTM creation is used to relate locally a survey with its corresponding quality indicator. This layer is the product of both the selection of surveys composing the DTM, and in case of overlap, the choice of maintaining only the survey with the highest number of soundings. In other words,

the source layer consists of a mosaic of unique sources of bathymetric data, which then provides, using the data/product catalogues a direct relation to the corresponding metadata, including the Quality indicators values.
Using a principal component analysis (PCA) on all the QI for each of the surveys indicated that there was no redundant information, or strong correlation between the QI. In practice, this means that the choice and the definition of the QIs explain the variability of the surveys and no simplification is possible. Hence, the figures below show the spatial distribution of the different classes for each QI. Those are also available on the EMODnet Bathymetry portal as a mean for the users to assess the confidence they can get from the local sources of the DTM.

Horizontal QI

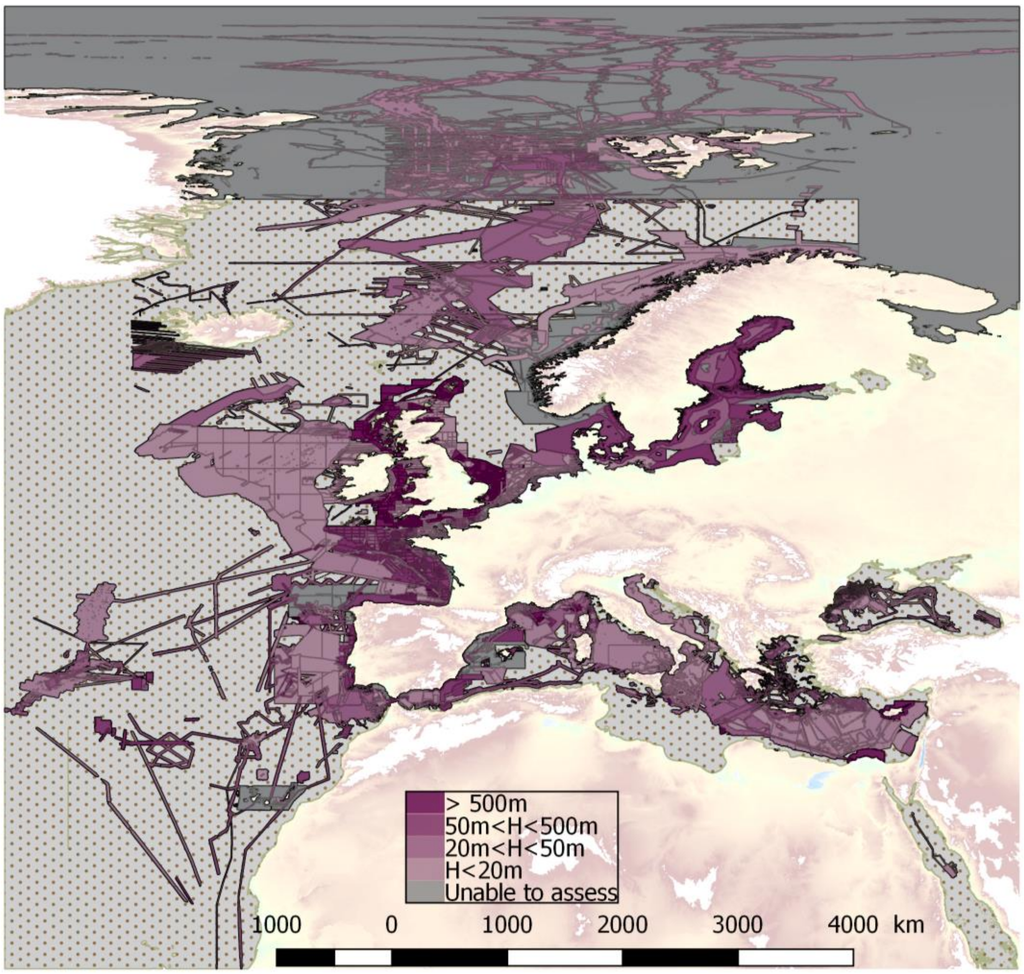


Figure 5: Horizontal QI for the overall EMODnet Bathymetry area of interest

The Horizontal QI is an indicator of the accuracy of the positioning. Globally, we observe a predominance of datasets having a horizontal accuracy equal or better than 50 meters, mostly in coastal and shallow waters. One has to compare the values selected for the classification of the horizontal QI, with the grid size of the DTM (1/16 arc minute, which is roughly 100 meters). With a horizontal accuracy better than 50 meters, we can conclude that most of the datasets used for the compilation of the DTM are suitable with its resolution.
However, we can observe two cases of unknown horizontal accuracy (grey zones on the map):

- In the case of the use of a composite product as the data source (typically the case of the

IBCAO compilation to the North), for which the Horizontal Accuracy indicator may not be unique (multiple surveys with various acquisition devices composing the composite product) or in some case may simply not be known to the data provider. In that case, the horizontal QI value (respectively the following others) was not set.

- The DTM is also filled by previous version of the EMODnet Bathymetry product. For a few numbers of input data of the EMODnet 2016 DTM, (for instance when a partner left the project), the metadata have not been updated with their quality indicators (including the following ones).

Figure 6: Horizontal QI close-up for the southern North Sea.

Figure 6 shows a zoom for the QI Horizontal in the area of the North Sea. This map allows seeing in better details the distribution of the Horizontal QI. Note also that the area with the dotted motif corresponds to the area filled with GEBCO, for which no assessment of the quality indicators was set.

Vertical QI

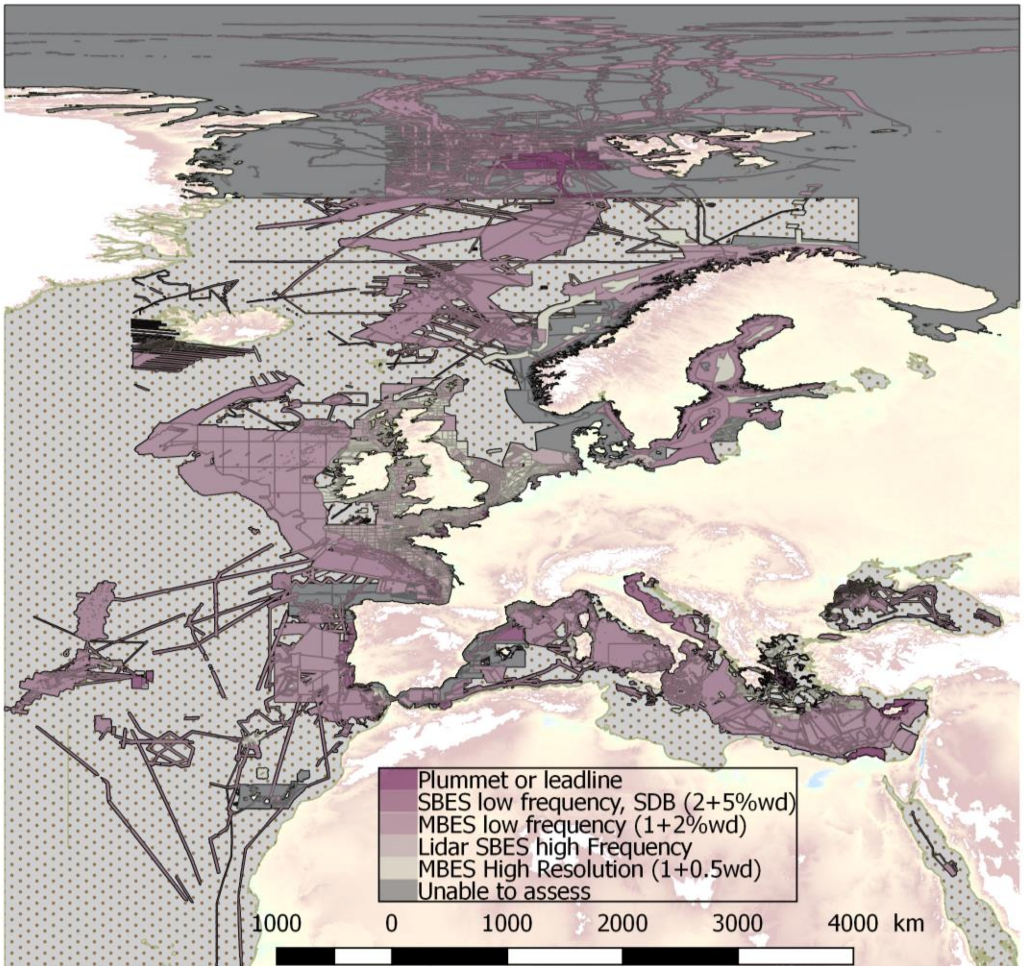


Figure 7: Vertical QI for the overall EMODnet Bathymetry area of interest

Vertical QI values are the result of a classification of historical and current sounding techniques. Most of them are dependent on the waterdepth (see Table 1 above). On Figure 7, after excluding the GEBCO and the not assessable areas (essentially composite product), we can observe a large predominance, over 90% of the remaining polygons, with vertical QI values above 2, which numerically corresponds to a vertical precision better than 1+2% of the waterdepth. This clearly illustrates the wide coverage of multibeam data, lidar and high resolution singlebeam data on the continental shelf and also the preference from the basin coordinators to exclude single beam data in the deep basins, where multibeam data provide a better coverage.

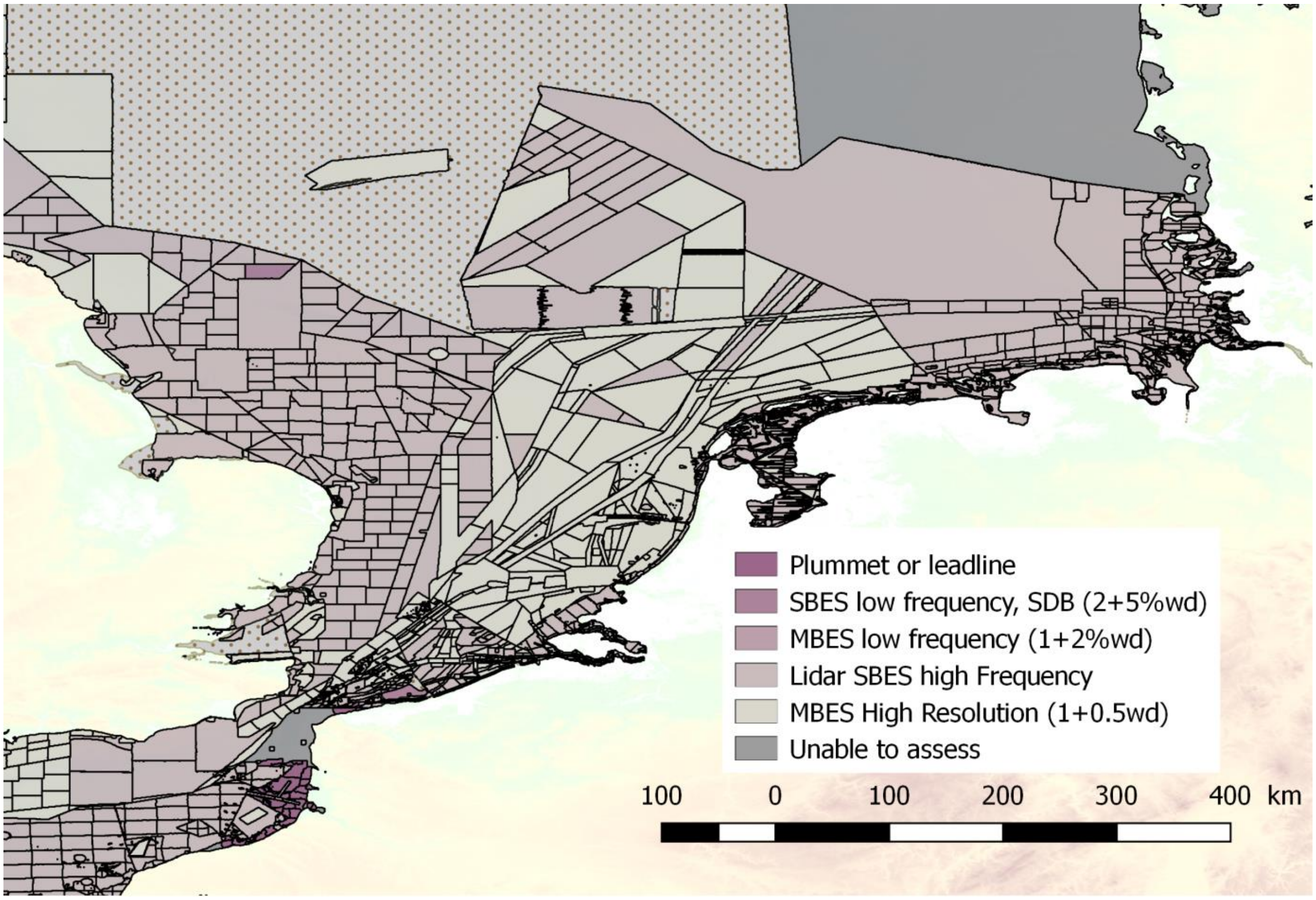


Figure 8: Vertical QI close-up for the southern North Sea.

Purpose QI

The Purpose QI provides information about the objectives in terms of the level of precision, resolution and accuracy sought by the data owner during the acquisition and processing of the bathymetric information. These objectives are often defined as a compromise between the level of precision needed for a specific usage and the logistic resources that can be provided to reach it. For example, less care is generally taken when data are acquired as part of a transit than for survey dedicated for scientific research or for the safety of navigation (Hydrographic survey). Using this indicator during the selection of overlapping surveys can help defining the level of priority in which the data must be considered (or filtered out). It also provides a qualitative appreciation of the efforts put during the processing phase.

Globally, Figure 9 clearly indicates that most of the coastal surveys were acquired with hydrographic related objectives, essentially related to the safety of navigation. Likewise, in deep sea areas, the purpose of the surveys is essentially morphological and research oriented. This is also the case of some of the composite products, such as the IBCAO compilation in the Arctic areas. In the southern North Sea close up, the need to provide bathymetric information for the safety of navigation is evident and is clearly summarized on Figure 10.

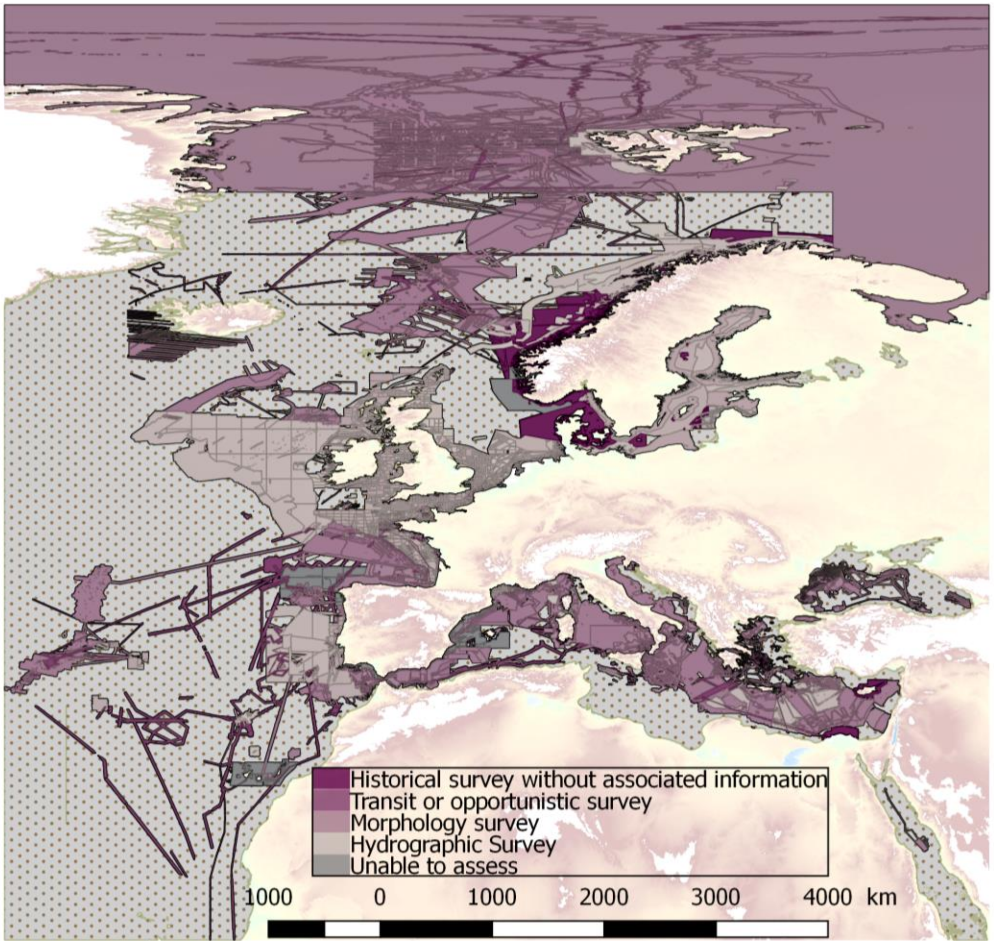


Figure 9: Purpose QI for the overall EMODnet Bathymetry area of interest

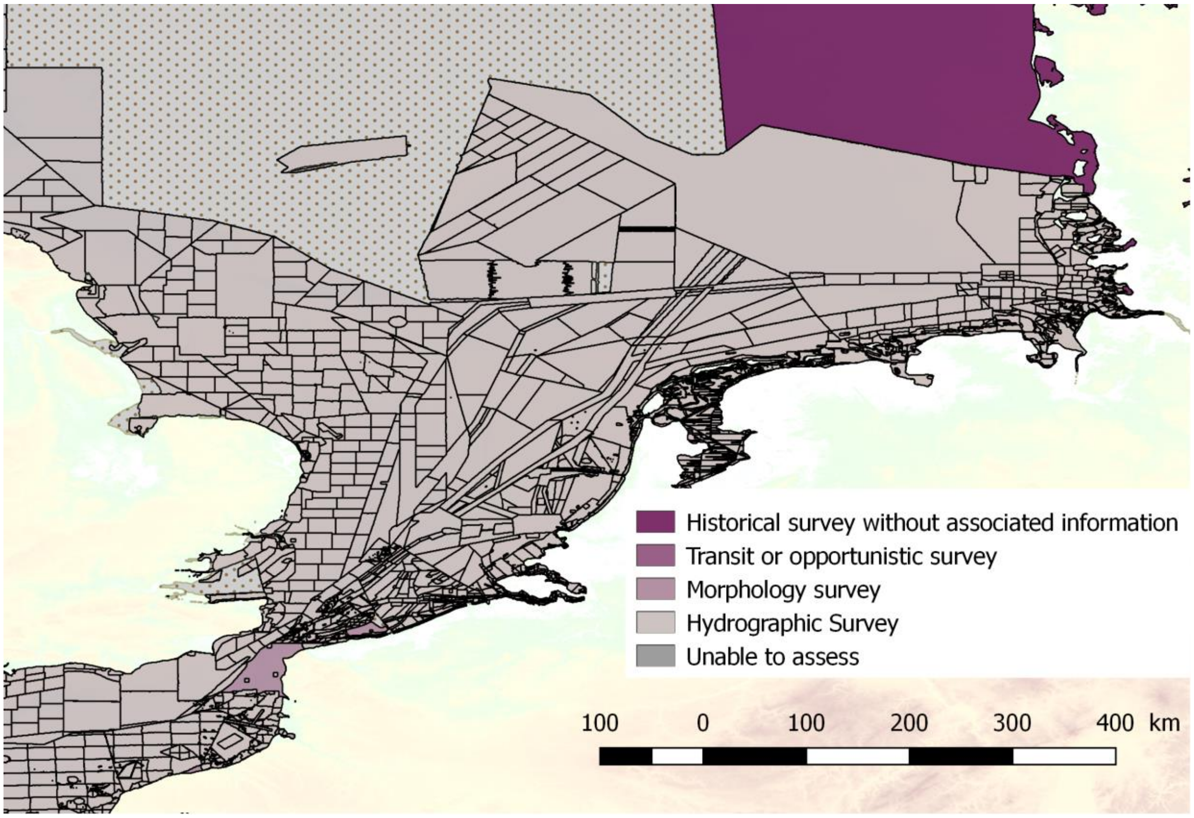


Figure 10: Purpose QI close-up for the southern North Sea.

Age QI

The Age QI provide a temporal indicator. The seabed is an highly dynamic environment where the interplay of the nature of the seabed and the physical forcings generate and constantly modify its shape. Sandy areas associated with strong currents such as for the area of the Straight of Dover is an example of such highly dynamic environment. At the opposite, rocky shores Scotland of Ireland, for example, are less subject to morphological changes. Classification of the Age of the survey has been broadly defined so that significant changes of the morphology spaning from one grid cell to the other (aprox 100m horizontally) are potentially observed.
Figure 11 does not really indicate a general trend in terms of distribution of the surveys with respect to their date of acquisition. One can see that the Baltic Sea, UK, France and Portugal approaches have been generally surveyed more than thirty years ago. Even if, locally recent high resolution surveys have been done to update the bathymetric information, the main explication that can be found is that this observation is somewhat biaised by the fact that those countries hold the oldest hydrographic organisations.
Figure 12 also clearly shows effort put by countries like Belgium and the Netherlands to continuously update their bathymetric information along the main navigation routes (amongst the highest maritime traffic area), knowing also that the seafloor of both countries is subject to constant changes due to the soft sediment nature of the seabed (input from the Rhine, amongst other) and the strong tidal currents.

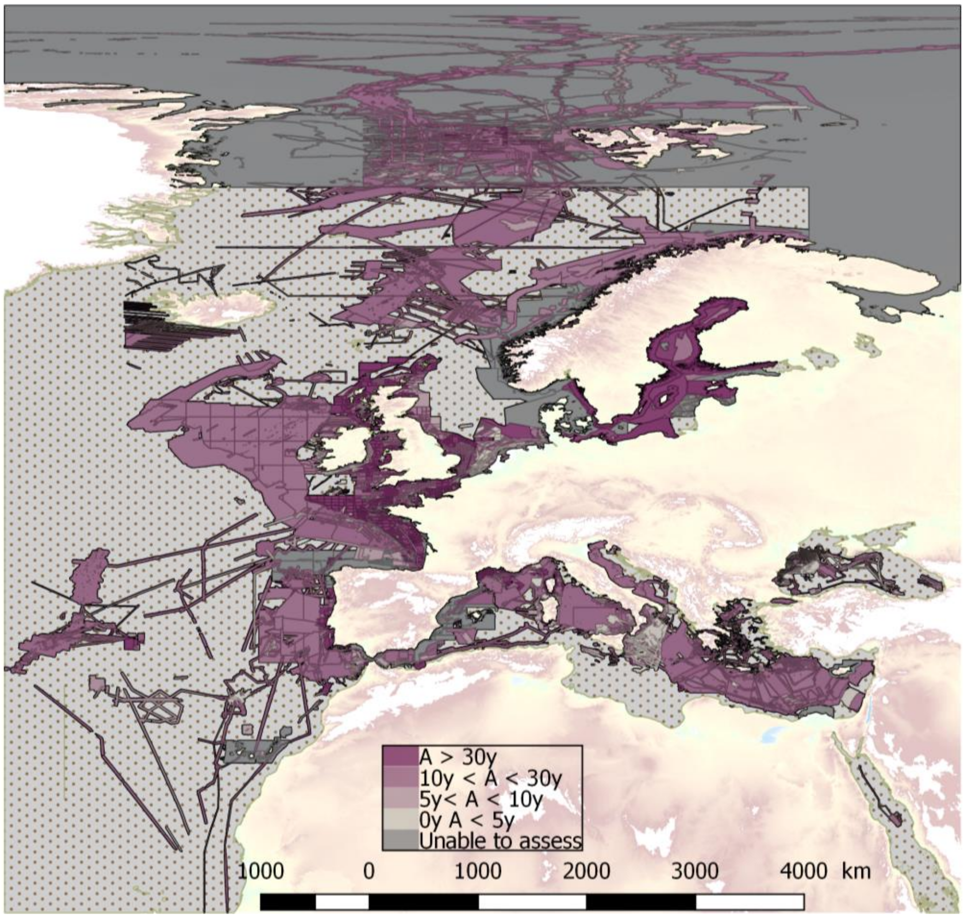


Figure 11: Age QI for the overall EMODnet Bathymetry area of interest

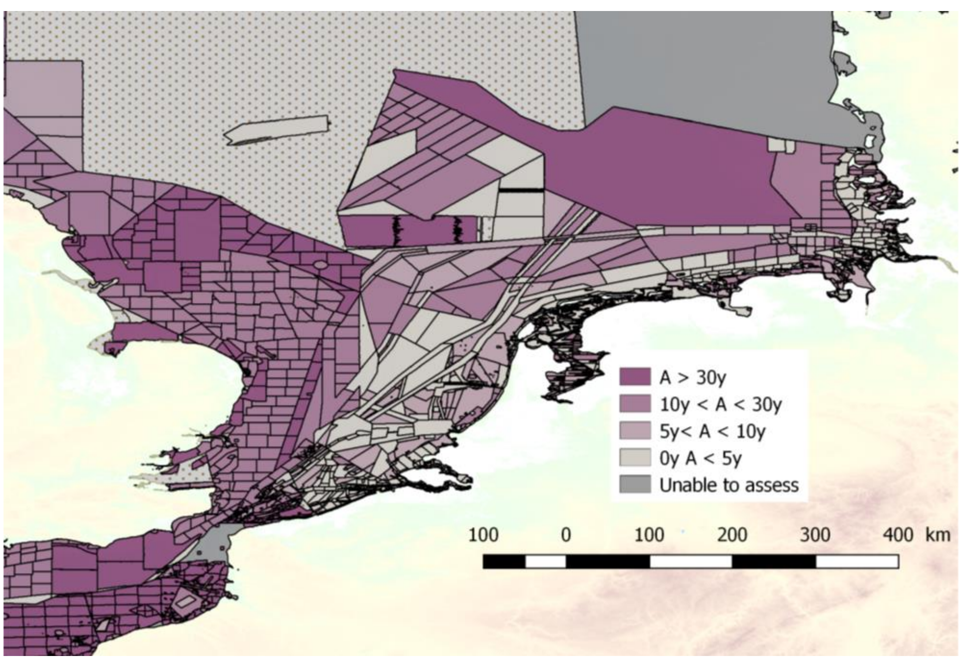


Figure 12: Age QI close-up for the southern North Sea.

6. Combined quality indicator

In order to further help the data user in assessing the local confidence he/she can get from the DTM, the combination of the four indicators into one synthetic indicator was pursued.
A first approach evaluated was to combine the indicator for generating groups of datasets.
The preferred approach which is proposed here consists of generating a score (composite QI value referred as CQI in the following lines) as a simple sum of each of the indicator value (see Table 1) reported to the best score values (13 sum of the 4 best individual score). This value is then reported to 100, to provide a percentage value. In other words the formula is as follows:

$$CQI = 100 * \left(\sum_{i=1}^{4} QI\ values \right) / 13$$

This approach provides a relative ranking of each of the dataset, in reference to what could be the best conditions in terms of data acquisition.

Figure 13 and Figure 14, provide the geographic distribution of the CQI. They have been generated based on 3 classes (plus one class related to the non-assessable data sources) as one can see from the values given in the legend. Those are separated using the Jenks natural breaks method (which breaks the occurrences with respects to the variability observed in the frequency histogram).

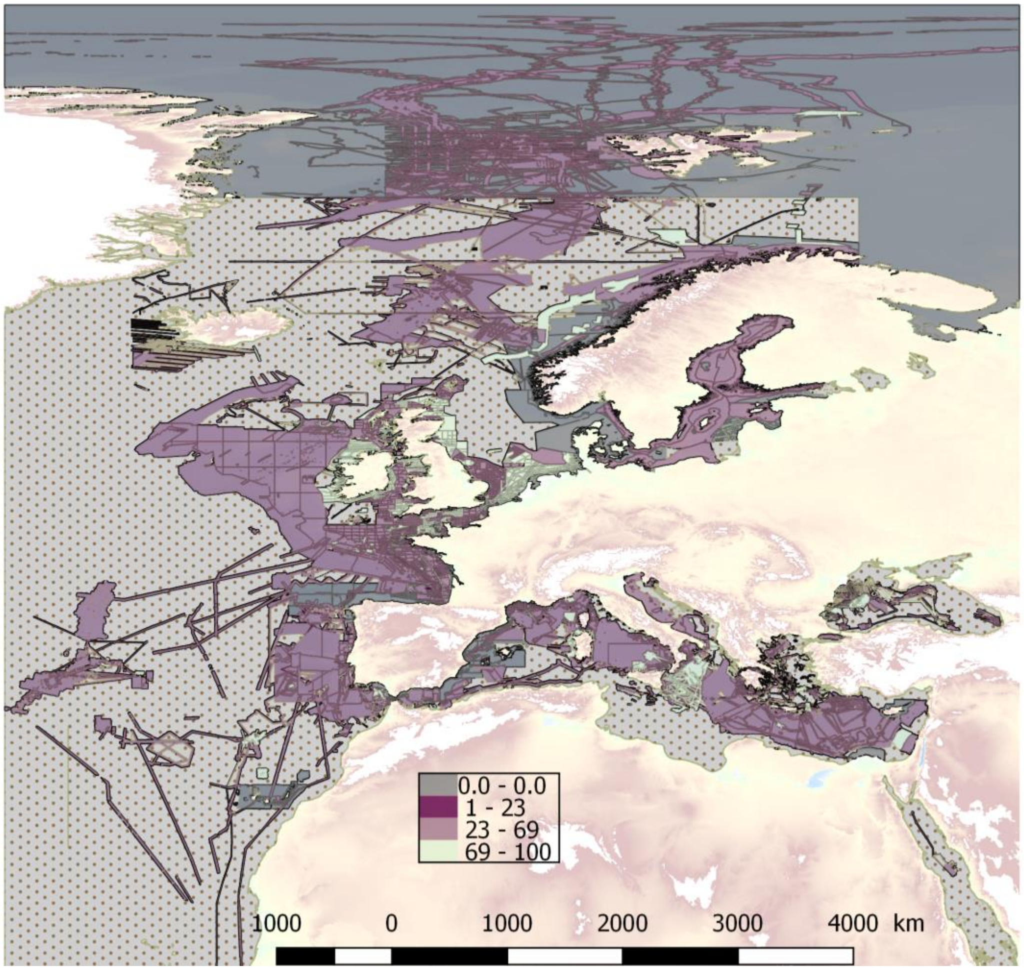


Figure 13: Combined Quality Indicator (CQI) for the overall EMODnet Bathymetry area of interest

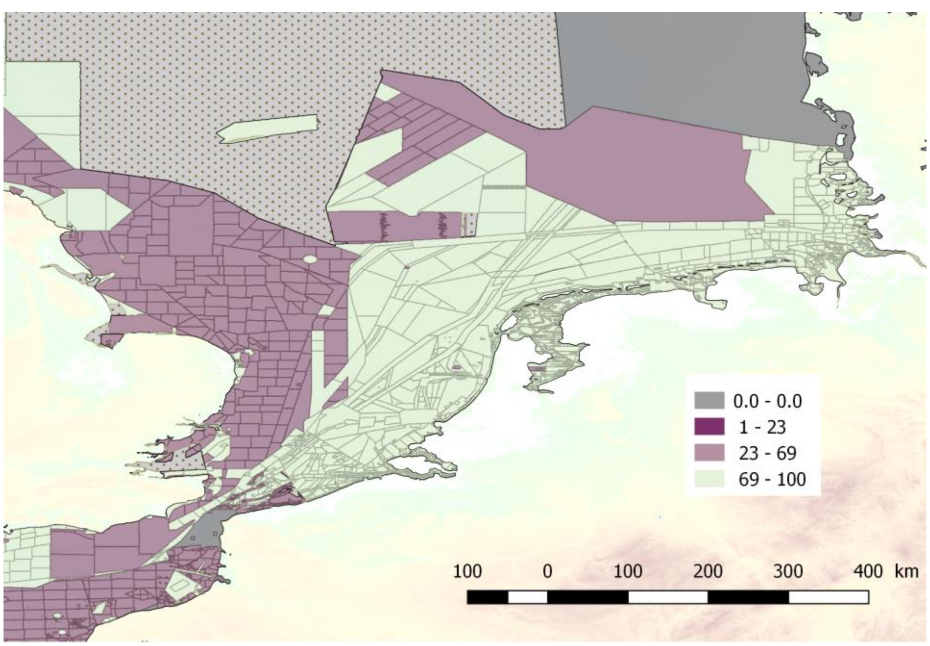


Figure 14: CQI close-up for the southern North Sea.

7. Use of the quality information and concluding remarks

Quality indicators have been defined based on existing approaches (see chapter 2 of this document) with the aim to provide directly expert knowledge from the data providers on their datasets. Bearing in mind that the users of bathymetric models might have specific needs, it has been decided to use four independent indicators on the vertical, horizontal, age and purpose of the survey. The distribution of each of these descriptors can be displayed geographically individually or as a synthesis (CQI) in GIS solutions and directly on the portal (see example from the Figure 15).

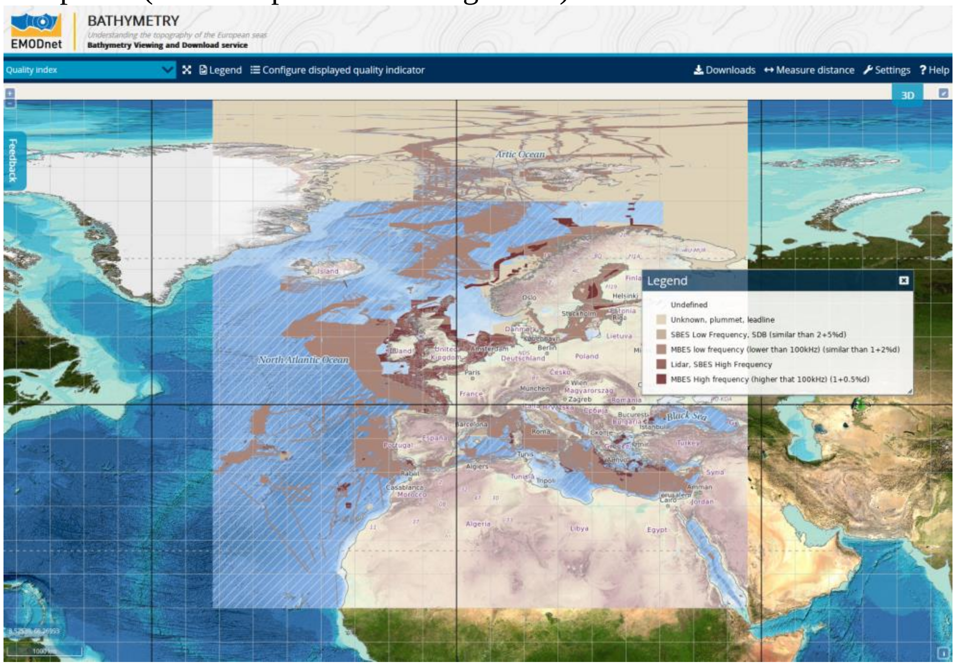


Figure 15: Implementation of the QI display on the EMODnet Bathymetry portal. Example of the Vertical QI.

Before concluding, several points are herein discussed in order that the users fully apprehend the meaning and the limitations of the followed methodology.

- Each of the quality indicators describes the acquisition system and some elements of processing. Those are globally known either by the specification of the survey, or the theoretical characteristics of the elements composing the acquisition system. In this way we can say that the qualification is assessed a-priori by most of the data providers and not a-posteriori (as a measure of performance of the survey).
- In the same line as the previous argument, the quality indicators are defined globally for the entire survey. It does not take into account, for example temporal variability of the precision of the devices composing the acquisition system (e.g. adverse conditions affecting the accuracy of the measurements of the motion of the vessel).
- Efforts have been made to explain in details the rationale of the methodology and the different classes of the Table 1. However, it has been shown through this work that further interactions were needed with the data owners in order to properly fill these indicators. Some inconsistencies were found (see section 3 of this report), but more importantly, despite the good will of the data providers describing the data using these quality indicators can involve some level of subjectivity, or inhomogeneous understandings of the meaning of the parameters. Efforts will have to be taken

in the future to better integrate the generation of the QI indicators with the global generation of the DTM in order to provide more robust indicators.

- Concerning the generation of the CQI, the choice has been made to consider each of the quality indicators with the same importance in the calculation. An alternative could be to weight, for instance one of the indicators to reinforce its importance. For example, emphasizing the component on the measurement precision could be materialized by a weighting the Horizontal and Vertical indices over the age and the purpose ones (e.g. weighting by 2, 2, 1, 1 respectively). At the present stage, we motivated our choice not to use weights, in order to remain neutral and not anticipate some usages of the DTM which should request to emphasize some parameters over the others. In any ways, if the user wants to focus on one or several of these indicators, he/she is able to interrogate the individual corresponding QI layers.
- With respect to the meaning of the CQI score, one has to understand that the combined quality indicator provides a relative ranking of the datasets. Following such a vision, it is suggested that the user refers to either a “low score” (0-40%), or a “mid score” (40-70%) or a “high score” (70-100%). It must be noted, that used alone, the CQI does not provide a clue on the reasons of a low or mid score.
- Finally, while the methodology developed here is essentially based on the description of each of the dataset by experts who have acquired/processed/managed the datasets, it is clear that other elements could be of interest to assess in further details our product at the grid node level. Amongst those parameters, one can anticipate that the number of soundings or the standard deviation per grid node, or the slope or some other morphological indicators could valuably be aggregated with the actual methodology.

At the present stage, the EMODnet Bathymetry product is one of the only publically available bathymetric products which provides a detailed description of its quality at the geographical level. Most of the other bathymetric products either provide a statistical value of the adequacy between reference data samples and the product (typically Root Mean Square Error); or simply indicates the origin of the soundings (Source Identifier concept of GEBCO) with no relations with associated qualitative descriptors. For the EMODnet HRSM consortium this topic of confidence assessment of bathymetric compilation products is of high importance and will be pursued in the future.