Automatic Data Quality Assessment of Hydrographic Surveys Taking Into Account Experts' Preferences

Julian Le Deunf, Arwa Khannoussi, Laurent Lecornu, Patrick Meyer, John Puentes

Automatic Data Quality Assessment of Hydrographic Surveys Taking Into Account Experts' Preferences

Julian Le Deunf
Development bathymetry department
Shom
Brest, France
orcid.org/0000-0003-3644-5063

Arwa Khannoussi
ERIC EA 3083, Université de Lyon,
Université Lumière Lyon 2
Lyon, France
arwa.khannoussi@univ-lyon2.fr

Laurent Lecornu
IMT Atlantique
Lab-STICC, UMR CNRS 6285
Brest F-29238, France
laurent.lecornu@imt-atlantique.fr

Patrick Meyer
IMT Atlantique
Lab-STICC, UMR CNRS 6285
Brest F-29238, France
patrick.meyer@imt-atlantique.fr

John Puentes
IMT Atlantique
Lab-STICC, UMR CNRS 6285
Brest F-29238, France
john.puentes@imt-atlantique.fr

***Abstract*—Data quality assessment of hydrographic surveys is a complex problem, since context dependent acquisition conditions using multiple sensors contribute to numerous data imperfections, which have different consequences depending on the intended final product. In this work we propose a generic methodology integrating experts' preferences through multi-criteria preference models with data quality techniques to generate explainable overall quality assessments of hydrographic surveys which depend on the expected end-uses.**

Four hydrographic surveys of different geographic locations and contrasting characteristics were studied, according to the preferences of an acoustician, an oceanographer, and a hydrographer. The obtained results indicate that each survey was appropriately evaluated, indicating the reasons that led to the specific assessment.

***Index Terms*—hydrographic survey, data quality assessment, preference modelling, multi-criteria decision aiding.**

I. Introduction

The International Hydrographic Organization (IHO) published recently an updated edition of the Standards for Hydrographic Surveys [5]. This sixth edition shows the importance that the world of hydrography attaches to data quality, in particular by adding a new and stricter order, the "exclusive order", which requires to establish the appropriate survey methodology and the appropriate acquisition sensors to achieve the specified standards (EO, OS, 1A...). In addition, the new standard integrates a matrix (7.6 of [5]) allowing to refine the data quality levels of criteria according to stated hydrographic needs, which means that depending on the final use of a survey, its quality could be perceived differently by users. It follows from this that, evaluating the quality of a survey, while taking into account users' needs, as well as automating this evaluation are real issues. Next to that, data and information quality are usually examined through multiple parameters, dimensions (or criteria). The choice of these parameters, as well as their relative importance, do obviously depend on the user. These parameters and dimensions might also be influenced by the context [8], which in turn might be perceived differently by each user, and therefore modeled in various ways.

These observations constitute the research question of our proposal: How to automate the evaluation of the quality of hydrographical surveys, integrating end-use experts' perceptions, as well as the context of the data, while keeping a high level of traceability of the evaluation process and a good explainability of the result?

To answer this question, we use data quality analysis techniques to automatically determine, from data and metadata, various quality measures of hydrographic surveys, and combine these with preference models from Multi-Criteria Decision Aiding (MCDA) to integrate various expert profiles, into the evaluation process. We show among other things that these multiple quality criteria, combined with different experts' perceptions about these data, can lead to different overall evaluations of the quality of hydrographical surveys. Last but not least, we show how these combined tools can lead to explainable outputs, which can be used to audit existing surveys, and to improve future ones.

The rest of this article is structured as follows. In Section II we briefly introduce the two scientific disciplines which underlie the proposed assessment process. Then, in Section III we detail our proposal and discuss its different steps, before presenting in Section IV a real case study tested at the Shom, the French Naval Hydrographic and Oceanographic Service. We then discuss our findings and the consequences of our proposal in Section V, before drawing in Section VI some conclusions and presenting future perspectives.

II. Methodological background

A. Hydrographic data quality basics

Hydrographic monitoring for analyzing past and current states, as well as longitudinal evolution utilizes different technologies to measure remotely physical variables of a given area, which rely on the principle of sensing specific types of radiations emitted and/or reflected by the studied zones. Sensors that range from radars, radiometers, and sonars, to optical airborne cameras and Light Imaging Detection and Ranging (LIDAR), permit to detect, identify, classify, and map the surfaces of interest. Studies of such hydrographic phenomena produce voluminous, noisy, redundant and/or contradictory heterogeneous data streams, depending on sensor characteristics, operational conditions, and uncontrollable external circumstances. Therefore, data quality assessment is unavoidable before data are processed by different applications, since information extracted from poor or low quality data would produce wrong results.

Yet, quality assessment in hydrographic studies is a complex problem, given that: perfect data acquisition conditions are almost impossible to obtain; generated datasets do have multiple imperfections; data quality evaluation must be adapted to each kind of sensor; workflows vary from one use case to another; and quality depends strongly on the context. Besides individual quality measures at different workflow stages, it is also required to understand the impact of data quality on the different intended products (or end-uses). In general, compliance with the *fitness for use* concept is a widely accepted notion to characterize globally data and information quality. This concept encompasses the definition of related dimensions depending on the analyzed operational context and qualified features. However, given the issues to carry out an automatic data quality control (i.e. unavailability of ground truth values and required infrastructure to do it in real time), part of it is done manually by experts who assign arbitrary global qualitative estimations, based on their knowledge of the expected final product characteristics. This evaluation consists on removing obvious errors and anomalies, deciding to accept or reject remaining data, estimating through sampling if unreliable data may still be part of the dataset, and assigning the estimated data quality levels using a pre-defined reference.

On the other hand, data, defined as streams of bits with no comprehensible sense, may contain sets of imperfect values that can be:

- Erroneous: Data are erroneous when values are different from the true data.
- Incomplete: Data are not fully supplied as expected because of missing values.
- Imprecise: Data inaccuracy does not permit to identify true values but possible approximations.
- Uncertain: Data cannot be specified with absolute confidence.
- Unavailable: The system cannot obtain some sets of values because of its limitations.

Depending on the use case, these imperfections do not have the same impact on the final product, making it necessary to study particular quality profiles depending on user preferences and application requirements.

B. Multi-criteria decision aiding

It follows from the above considerations that evaluating the quality of information or data from hydrographic surveys is a problem of a multi-dimensional nature, and potentially depends on the perspective of an expected end-use. It is therefore suitable to use aggregation techniques from the field of Multi-Criteria Decision Aiding (MCDA), which allow to aggregate multiple dimensions, while taking into account the point of view of an expert of the end-use, through what we call his or her preferences.

MCDA is a branch of operational research, and its objective is to help one or several decision-makers (DMs) to prepare and make decisions on a finite set X of n decision objects (or alternatives), when several conflicting consequences (or criteria), represented by the finite set J, must be taken into account. This decision-maker can either be the person who takes the responsibility for the decision act, or he or she can be an expert user whose value system or preferences should be taken into account in the final prescription.

Usually, three types of problems are put forward in this context [13]: the *choice problem* which aims to recommend a subset of alternatives, as restricted as possible, containing the “satisfactory” ones, *sorting problem* which aims to assign each alternative into pre-defined categories or classes, and, the *ranking problem* which aims to order the alternatives by decreasing order of preferences.

The mathematical tools used in MCDA have their origins mainly in two very different methodological trends [3], [6], [13]. On the one hand, the European school of thought has developed around the concept of the outranking relation, where the decision recommendation is constructed from pairwise comparisons of the alternatives. On the other hand, the Anglo-Saxon school is based on the concept of utility or value in the Multi-Attribute Value Theory (MAVT) in order to obtain, by aggregation, a total comparability of the alternatives. Both schools of thought have their advantages and disadvantages, and in the case of practical application, the pros and cons should be weighed before choosing one of the paths.

The main differences between these two methodological schools lie in the way the alternatives are compared and in the type of information which is required from the decision maker. Outranking methods might be preferable if the evaluations of the alternatives on the criteria are heterogeneous, i.e., qualitative and quantitative, and if the DM would like to include some impreciseness about his preferences in the model. Value-based methods can be favored if the criteria are evaluated mostly on numerical scales and if a compensatory behavior of the DM should be modeled.

III. Proposed approach

The proposed quality assessment process is depicted on Figure 1. This process is intended for a user who wishes to evaluate the quality of a hydrographic survey according to a specific use of the survey (e.g., for an acoustic, an oceanographic, or, a hydrographic use, in our case). The process needs to be pre-configured by an expert of the final need (shown on the right of the figure), who is an acoustician, oceanographer or a hydrographer in the present case. The input to the process is the hydrographic survey data, along with related metadata. The outputs are the partial and the overall evaluations of the input hydrographic survey, as well as some recommendations to improve its quality (if necessary).

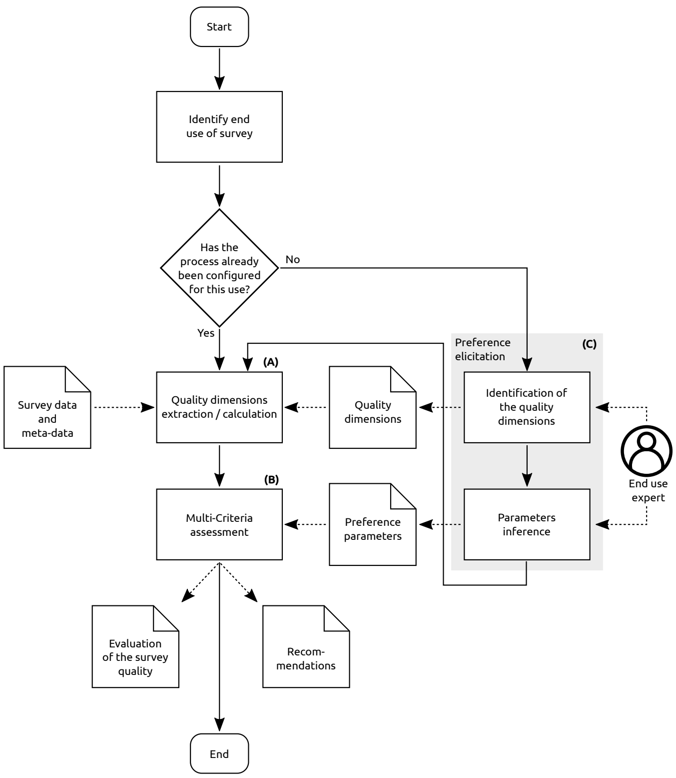

Fig. 1. Quality assessment process for hydrographic surveys

Once the user starts the process of evaluating the quality of a set of surveys, in a first step, he or she needs to identify the end-use of the surveys. If the process has already been configured for this specific use, the quality dimensions that have been specified beforehand by an expert are extracted and calculated (A) from the hydrographic survey data and its corresponding meta-data. Then the Multi-Criteria assessment model (B) is used to aggregate these multiple quality dimensions into an overall quality evaluation, while taking into account the preference model of the expert. In case the process has not been configured yet for this use, an expert of this specific use has to be interviewed in a sub-process called “Preference Elicitation” (C), in order first to configure the “Quality dimensions extraction / calculation” step with the quality dimensions considered by that expert, and second to configure the Multi-Criteria assessment with the preference parameters, representing that expert’s priorities.

In the following subsections we detail each of the main steps of this process (the letters in parentheses on Figure 1 correspond to these subsections).

A. *Quality dimensions extraction*

During a sea survey, the hydrographic survey services acquire a large amount of data at sea for various needs and for different fields or end-uses: hydrography, oceanography, sedimentology, underwater acoustics, nautical cartography, digital terrain model production, defense products, etc. The sensors put in place are thus numerous and depend on the type of campaign carried out.

When interested in bathymetry they use for example a Single Beam Echo Sounder (SBES), a MultiBeam Echo Sounder (MBES), a bathymetric LIDAR or even Satellite Derived Bathymetry (SDB) for the seabed closest to the coast. These data, in the form of point clouds, allow them to have a more or less fine representation of the seabed depending on the coverage achieved during the acquisition phases and the resolution of the sensor technologies.

When studying the physical representation of the ocean, it is more interesting to measure the surface currents and the currents in the entire water column (via bottom current meters or acoustic doppler current profilers for example) or the variations in the water height (using pressure tide gauges or buoys referenced to the ellipsoid) to measure tidal time series at specific locations.

Finally, for sedimentology studies of the seabed, the knowledge of the bathymetry is important in order to detect local and multi-scale morphological breaks, but it is also particularly interesting to have access to a classification of the encountered seabed (mud, sand, rock...). To achieve this type of classification many different sensors are implemented, such as acoustic sensors (SBES and MBES) to have seabed surface information, but also techniques of in-situ sampling with grab sampler or coring, which also provide information on the type of seabed at greater depths (here also at specific location).

The hydrographers who carry out such survey are acquainted with the data acquisition conditions and therefore compute or are able to assign quantitative and qualitative trustable qualifications to the respective datasets. After the interviews with the experts and in the context of the specific case study presented in Section IV, seven different quality parameters of the main survey characteristics were selected among a set of numerous possibilities, namely, CATZOC, POSACC, SOUACC, hydrographic coverage, type of hydrographic sensors, sedimentology, and ocean data. Those quality parameters are described as follows:

- CATZOC, for category of zone of confidence in data or tolerable quality thresholds, as an overall summary of POSACC, SOUACC, and hydrographic coverage, consisting of categorical values assigned from the best to the worst qualification as: A1, A2, B, C, D, and U. The associated quality dimension is *global data uncertainty* because of the resulting complexity of measurements, the integration of data from several sources, or the possibility of human evaluation errors that may generate uncertain evaluations.

- POSACC, for positional accuracy (in agreement with the CATZOC value), which includes four depth intervals from the smallest to the highest as: < 0.50m, < 2m, < 5m + 5%, < 20 + 10%. The related quality dimension is *data precision* given that the qualification intervals represent a variable approximation of the instrument position. The POSACC quality parameter is traditionally obtained by total uncertainty propagation taking into account the entire bathymetric acquisition workflow.
- SOUACC, for sounding accuracy (in agreement with the CATZOC value) that is represented by six intervals from the shortest to the highest as: < 0.25m, < 0.50m, < 1m, < 2m, < 10m, < 20m. The relevant quality dimension is *data precision*, since depending on the accuracy interval of a sounding instrument, resulting values are only known approximately. Like POSACC, SOUACC quality parameters, it is traditionally obtained by total uncertainty propagation taking into account the entire bathymetric acquisition workflow.
- Hydrographic coverage (in agreement with CATZOC value), composed of six values, 300, 200, 150, 100, 80, and 50. The corresponding quality dimensions are *data completeness* and *data precision*, as a result of the estimated degree of coverage and derived unavailability of expected data, respectively.
- Type of hydrographic sensors, represented as four types of sensors, MBES, bathymetric LIDAR, SBES + side-scan sonar, and Satellite Derived Bathymetry (SDB). Sensor types have an impact on the density and pertinence of generated measurements, allowing the user to anticipate some of the information that could be obtained. The applicable quality dimensions are *data uncertainty* depending on the utilized sensors accuracy and *data completeness* determined by the potential complementary nature of sensors.
- Sedimentology, represented by the number of sensors from five to zero (5, 4, 3, 2, 1, 0), used to describe seabed substrates, depending on emitted signals reflections. The inferred quality dimensions are *data precision* linked to the accuracy of different sensors and *data completeness* as a result of potentially missed pertinent data not acquired by a reduced set of sensors.
- Ocean data that identifies measurement of the tide, current, meterology (MTO), water sound celerity (with CTD: Conductivity Temperature Depth), depicted as four sets of grouped evaluations (MTO + CTD + tide + current, MTO + CTD + tide, MTO + CTD, MTO). The identified quality dimensions are *data precision* due to the diverse sensor accuracies and *data completeness* to account for the probable lack of data not recorded from unused sensors.

Quality parameters' values are included as part of surveys' metadata, assigned by engineers responsible for the survey. These values reflect some aspects of a survey quality estimation according to the actual acquisition context. In order to know if the resulting dataset is appropriate to generate the product of a particular profile, it is necessary to determine the conformity of these quality parameters with the end-use expert's preferences.

B. Multi-Criteria assessment

The quality parameters extracted in the previous step become the criteria in the MCDA context. In this section we first motivate the choice of a specific MCDA preference model, before detailing its mathematical formulation.

As it can be seen from Section III-A the evaluation scales of the various quality parameters are quite heterogeneous. Some of them are qualitative or ordinal (CATZOC, type of hydrographic sensors, sedimentology, ocean data), while others are clearly quantitative (POSACC, SOUACC, hydrographic coverage). This speaks in favour of the outranking school of thought mentioned in Section II-B, which intrinsically manages well this diversity of scales. Next to that, to facilitate the adoption of our process by users, we seek to propose a solution in which each step is easily explainable, and where the final assessment can be easily interpreted, in order to generate recommendations for the improvement of future surveys. This interpretability can be achieved by solving the sorting problem from the outranking paradigm. Indeed, sorting alternatives in an MCDA framework comes down to creating an overall qualitative or ordinal assessment scale (based on so-called categories) to aggregate the multiple criteria. This is usually done through category limits or profiles, which can be seen as norms, against which the alternatives are compared, in order to decide to which category they belong. Among all the available sorting algorithms of the outranking paradigm, we propose to use the MR-Sort method, which has been characterized in [1], [2], and, has the advantage of generating highly interpretable results, while being a very expressive model. In the sequel we introduce the formalism of the MR-Sort method.

Let us consider a finite set of alternatives A (the surveys), a finite set of criteria indexes J (corresponding to the quality parameters of the previous step). For each criterion, the possible evaluations are ordered according to the preferences of the expert, which defines the preference directions of the criteria (either a lower value is preferred to a higher value, or vice-versa). The overall evaluation being categorical, let $\{c_1, \ldots, c_k\}$ be the k output categories (representing the k levels of the output ordinal assessment scale), ordered by their desirability, from c_1 being the worst category to c_k being the best one: $c_k \succ \ldots \succ c_1$ ($\succ$ stands for "is preferred to"). In our quality assessment context, these categories could for example be $\{$very high quality $\succ \ldots \succ$ medium quality $\succ$ low quality$\}$. These categories are characterised by a set of separating profiles $B = \{b_1, \ldots b_{k-1}\}$. Each category c_h is thus defined through its upper limit, b_h, and its lower limit, b_{h-1}, with the exception of the worst and best categories, which have only one limit. Each alternative and each category limit can be represented by a vector of evaluations with respect to the criteria. The evaluation with respect to criterion j can be viewed as a function $g_j : A \cup B \rightarrow \mathbb{R}$, where $g_j(a)$ denotes the evaluation of alternative $a \in A$ on criterion j and $g_j(b_h)$ denotes the evaluation of category limit $b_h, \forall h \in \{1, \ldots, k-1\}$, on criterion j. In this presentation of MR-Sort, we assume, without loss of generality, that the performances are supposed to be such that a higher value denotes a better performance. It is obvious that in a real application this is not necessarily the case. Furthermore the performances of the category limits are non-decreasing, i.e. $\forall j \in J, 1 < h < k : g_j(b_{h-1}) \leq g_j(b_h)$.

MR-Sort uses two *assignment rules* for placing the alternatives into categories: the pessimistic and the optimistic assignment rules [3], [13]. The pessimistic rule assigns an alternative a to the highest possible category c_h so that a *outranks* the category's lower frontier b_{h-1}. The optimistic rule assigns a to the lowest possible category c_h so that the category's upper frontier b_h *outranks* a. The pessimistic rule is the most commonly used in practice, as it generates safer recommendations.

An alternative a is said to *outrank* a frontier b_{h-1} if and only if there is a sufficient coalition of criteria supporting the assertion "a is at least as good as b_{h-1}" and no criterion strongly opposes (vetoes) that assertion. To measure if a sufficient coalition of criteria considers that a is as least as good as b_{h-1}, we first define for each criterion j a function $C_j : A \times B \rightarrow \{0, 1\}$ which assesses whether criterion j supports that statement or not:

$$\forall j \in J, a \in A, 1 \leq h \leq k :$$
$$C_j(a, b_{h-1}) = \begin{cases} 1, \text{ if } g_j(a) \geq g_j(b_{h-1}), \\ 0, \text{ otherwise.} \end{cases} \tag{1}$$

To assess whether a coalition of criteria is in favor of the outranking or not, $\forall a \in A, 1 \leq h \leq k$, we first define the overall concordance as:

$$C(a, b_{h-1}) = \sum_{j \in J} w_j C_j(a, b_{h-1}), \tag{2}$$

where w_j is the weight of criterion j. The weights are defined so that they are positive ($w_j \geqslant 0, \forall j \in J$) and sum up to one ($\sum_{j \in J} w_j = 1$). This overall concordance is then compared to a majority threshold $\lambda \in [0.5, 1]$ extracted from the DM's preferences along with the weights (in our case the expert of the end-use).

Even when the coalition is strong enough, a criterion may veto the outranking situation. An alternative a is therefore in a veto relation (denoted with V) with a profile b_{h-1} when:

$$a \,\mathrm{V}\, b_{h-1} \iff \exists j \in J : g_j(a) < g_j(b^v_{h-1}). \tag{3}$$

The veto profile b^v_{h-1} represents the minimum level of performance that an alternative needs to have in order to be allowed into category c_h via the weighted coalition of criteria in favor of this assignment. If on any criterion an alternative has a lower performance than the veto profile of (, then it is forbidden from being assigned to c_h or above.

To summarize, alternative a outranks frontier b_{h-1} (and therefore is assigned to at least the category c_h) if and only if:

$$a \,\mathrm{S}\, b_{h-1} \iff C(a, b_{h-1}) \geqslant \lambda \text{ and not } (a \,\mathrm{V}\, b_{h-1}). \tag{4}$$

Illustrative example: To illustrate the use of the MR-Sort method we consider a decision problem where different surveys have to be assigned to one of two categories "Good" (G) and "Bad" (B) according to an expert's preferences. Let us consider three surveys a_1, a_2 and a_3 which are evaluated on three criteria *CATZOC* (C) ($A_1 \succ_C A_2 \succ_C B \succ_C C \succ_C D \succ_C U$) ($\succ_C$ is the strict preference relation defining the evaluation scale of criterion CATZOC), *Hydrographic coverage* (H) (the higher the better), and, *Sedimentology* (S) (the higher the better). The evaluation of these alternatives on these 3 criteria is presented in Table I, along with the parameters of MR-Sort model (which have been determined beforehand from the preferences of the expert via the procedures presented in Section III-C). The category separating profile b_1 delimits the two categories G and B, through evaluations of B for *CATZOC*, 100 for *Hydrographic coverage* and 2 for *Sedimentology*. The veto profile b^v_1 is defined through the worst evaluation of *CATZOC* (U) and of *Hydrographic coverage* (50) and a number of sensors for *Sedimentology* higher than 2. This implies that only the last criterion will trigger a veto when the number of sensors is equal to zero.

TABLE I
ILLUSTRATIVE EXAMPLE FOR THE MR-SORT ASSIGNMENT PROCEDURE;

Model parameters

	C	H	S	w_C	w_H	w_S	λ
b_1	B	100	2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$
b^v_1	U	50	1				

Assignments

	C	H	S	$C(a, b_1)$	$C(a, b_1) \geqslant \lambda$	$a \,\mathrm{V}\, b_1$	$a \,\mathrm{S}\, b_1$	assignment
a_1	B	150	1	$\frac{1}{3} + \frac{1}{3} + 0 = \frac{2}{3}$	✓		✓	G
a_2	A_1	300	0	$\frac{1}{3} + \frac{1}{3} + 0 = \frac{2}{3}$	✓			B
a_3	C	80	4	$0 + 0 + \frac{1}{3} = \frac{1}{3}$		✓		B

As the first survey a_1 is at least as good as b_1 on the two criteria C and H, it has a sufficient coalition of criteria supporting its assignment into category G. Furthermore, it has 1 sensor so it is not in a veto situation. The second survey a_2 also has a sufficient coalition of criteria in favor of assigning it to category G however it has no sensor. Consequently it raises a veto which invalidates the outranking relation between a_2 and b_1, and therefore it is assigned to category B. The last survey, a_3, is at least as good as b_1 on only one criterion, S, so it does not outrank b_1, and it is consequently assigned to category B.

C. Preference elicitation

The preference elicitation sub-process of our proposal involves 2 steps. Recall that it does not concern the general user, but is only intended for the expert user (who is the DM in our work). In the first one, the end-use expert is asked to provide the quality dimensions/parameters which he uses in the overall assessment of the quality of a hydrographic survey. The output of this step is a list of these quality dimensions/parameters which is then used as input by the "Quality dimensions extraction / calculation step". The second step is the inference of the preference parameters of the MR-Sort Multi-Criteria aggregation method. These preference parameters are the preference directions of the criteria scales (whether the criterion has to be minimized or maximized), weights of the criteria (or quality dimensions) ($w_j, \forall j \in J$), the majority threshold λ, the number and the order of the output categories, the category limits ($b_h, \forall h \in \{1, \dots k-1\}$), and the veto profiles ($b_h^v, \forall h \in \{2, \dots k\}$).

These parameters may be both *directly* and *indirectly* elicited. In a direct elicitation, the precise values of these parameters are determined by interviewing the expert, usually in an interactive process, where the effect of the preference parameters on the overall assessment are presented to the expert, in order to tune them as precisely as possible. Usually the category limits and the veto profiles can be seen as some "norms" with respect to the end-use of the survey, whereas the criteria weights are extracted via questions regarding majority coalitions of criteria. In case such a direct elicitation is not possible, an indirect approach can be used, in which the expert is asked to assess the overall quality of some surveys (called "learning" or "assignment" examples). From his answers, mathematical optimization models determine the values of the parameters of the MR-Sort model, which are compatible with these overall assessments (as in [7], [9]–[12], [15], [16]).

IV. Hydrographic case study

To show the relevance of our proposed assessment process, we present in this section a case study that has been conducted at the Shom, the French Naval Hydrographic and Oceanographic Service. We have considered a user who is a Shom engineer, and who wishes to evaluate the overall quality of four hydrographic surveys. He or she starts the process of Figure 1 and has to identify, in a first step, the end-uses of the surveys. He or she wishes to assess the quality of these surveys according to first an acoustic use, second an oceanographic use, and finally, a hydrographic use. Initially the process is not configured for these three specific end-uses, and therefore, in the process of Figure 1, three experts (an acoustician, an oceanographer and a hydrographer) have been interviewed in the "Preference Elicitation" sub-process. Each of these 3 interviews have led to a different configuration of the process (different quality dimensions, different preference parameters).

The three possible end-uses of the surveys are indeed very different as the three experts have very different jobs.

The acoustician models underwater sound propagation. Acoustic sources are of different natures: biological, geophonic, and anthropogenic. The understanding of the underwater soundscape requires the mastery of all these parameters. This is a major challenge for the knowledge of the environment. In order to understand this underwater acoustic landscape it is therefore important to be able to model the entirety of what happens to acoustic waves in the water column but also with the reflection on the sea floor. The quality of the information needs to be good in terms of sedimentology, ocean data and with an MBES sensor with good coverage.

The oceanographer studies all activities related to the understanding and modeling of the physical parameters of the water column (temperature, salinity, transparency, ...) and their evolution. For his work it is important that the oceanographer has a good quality of information for the ocean data and to a lesser extent the sedimentology data. As the oceanographer builds his physics models from a volume of water, it is also important to have a precise knowledge of the bathymetric bottom and in particular: the real presence or absence of artefacts on the bottom that can disturb the physics model (with complex boundary conditions to manage) such as wrecks or underwater obstruction given by the POSACC and SOUACC quality extraction.

The hydrographer is an expert for the measurement of the sea floor, currents and tides. He needs to ensure safe navigation bathymetric data for all nautical products, especially for nautical charts. The most important data to be accurate are those related to bathymetric information (as CATZOC, bathymetric coverage and the hydrographic sensor type) as these are the ones that guarantee a just and safe nautical chart for the navigator. Also depending on the average depth of the survey, the oceanographic information will be more or less relevant (tidal phenomena having more impact at shallow depths than at deep depths).

The four hydrographic surveys, see Figure 2, have also very different characteristics and geographic locations. S2012056 is an airborne LIDAR topo-bathymetric survey of Guadeloupe Island. The average depth is about 15 meters. This survey was designed to establish a land-sea reference with very few oceanographic measurements (no current meters or tide gauges). S2015009 is an SBES exploration survey carried out on the Clipperton islet under very complex conditions. The average surveyed depth is 20 meters. The purpose of this survey was to measure the Cliperton channel for safe landing and boarding on the island. S2017026 is a classical MBES hydrographic survey which aims to improve the bathymetric knowledge of the traffic separation scheme off Ouessant Island. With an average depth of 100 meters, this survey allows for the updating of nautical works and more particularly nautical charts. S201902600 is an offshore MBES survey which has been performed in the Atlantic Ocean off the coast of the Faroe Islands. Its average depth is 1400 meters. This survey was carried out in order to know the acoustic environment as precisely as possible with a lot of oceanographic and sedimentological measurements.

A. Preference elicitation

As said above, the three experts have different requirements and priorities on hydrographic surveys, which lead to 3 different configurations of the proposed process. In a first step, for

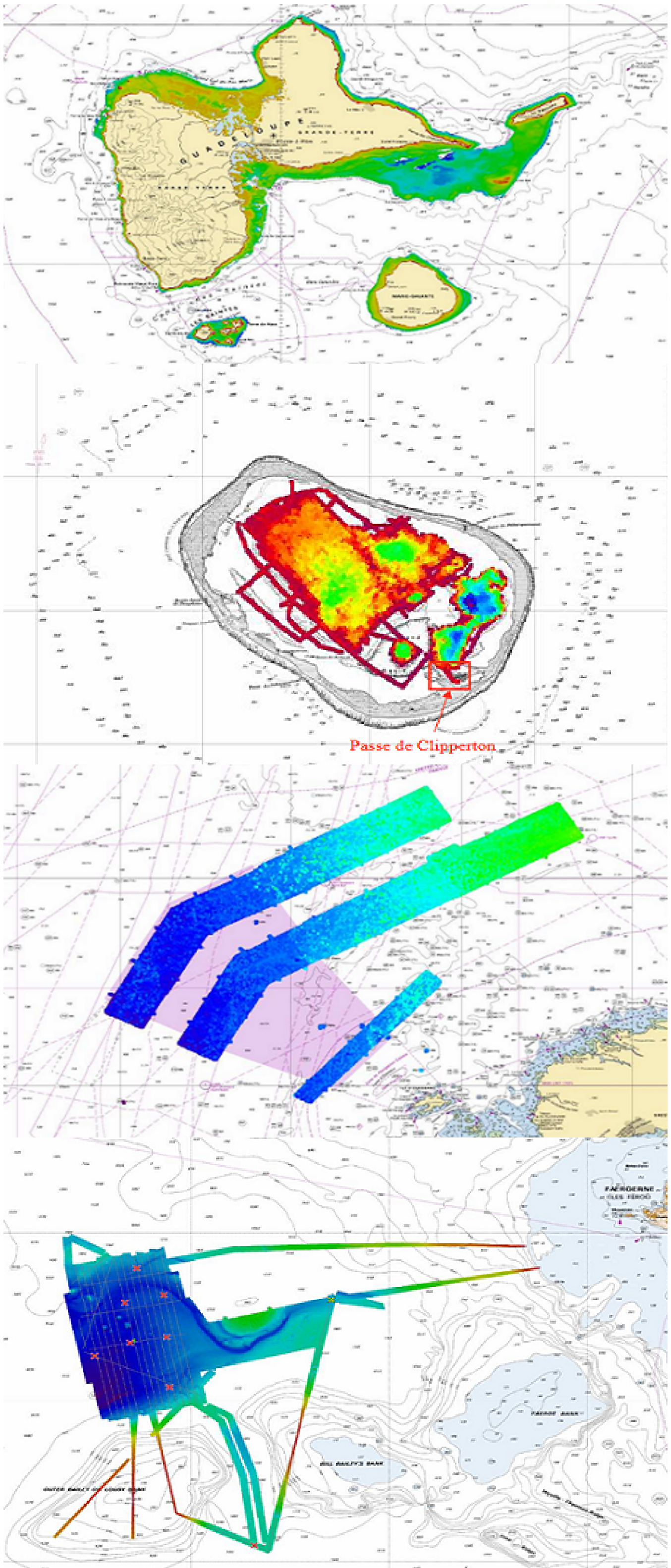

Fig. 2. Geographic extent of the four studied hydrographic surveys. From top to bottom: S2012056, S2015009, S2017026 and S2019026.

each expert, we identify during the interview which quality parameters should be considered in the final assessment. The hydrographer considers all 7 parameters presented in Section III-A in the evaluation of the quality of a survey, whereas the acoustician and the oceanographer do not consider the CATZOC, and therefore only require 6 of them.

In a second step, we determine for each expert the granularity of the output assessment scale. For the acoustician and the hydrographer, the two categories “Good (G)” and “Bad (B)” are sufficient (obviously preferentially ordered Good (G) $\succ$ Bad (B), whereas for the oceanographer a supplementary intermediate “Acceptable (A)” category is needed (which leads to the preference order Good (G) $\succ$ Acceptable (A) $\succ$ Bad (B)).

Then, during the interview of the hydrographer, we have identified that depending on the average depth of the survey, the preference parameters of the hydrographer vary. This influence of the context of the survey on the preference parameters therefore leads us to consider 3 different parameter profiles for the hydrographer. As a consequence, once a survey has to be evaluated by the process for a hydrographic use, the average depth of the survey has to be checked first, in order to determine which preference profile of the expert hydrographer has to be used to configure the process.

Regarding the evaluation scales of the criteria, the three experts have unanimously defined them as follows:

- CATZOC (A1 $\succ_{\mathrm{C}}$ A2 $\succ_{\mathrm{C}}$ B $\succ_{\mathrm{C}}$ C $\succ_{\mathrm{C}}$ D $\succ_{\mathrm{C}}$ U)
- POSACC ($< 0.5m \succ_{\mathrm{P}} < 2m \succ_{\mathrm{P}} < 5m + 5\%$ depth $\succ_{\mathrm{P}}$ $< 20m + 10\%$ depth)
- SOUACC ($< 0.25m \succ_{\mathrm{SA}} < 0.5m \succ_{\mathrm{SA}} < 1m \succ_{\mathrm{SA}} < 2m \succ_{\mathrm{SA}} < 10m \succ_{\mathrm{SA}} < 20m$)
- Hydrographic coverage ($300\% \succ_{\mathrm{H}} 200\% \succ_{\mathrm{H}} 150\% \succ_{\mathrm{H}} 100\% \succ_{\mathrm{H}} 80\% \succ_{\mathrm{H}} 50\%$)
- Hydrographic sensors (MBES $\succ_{\mathrm{HS}}$ Lidar $\succ_{\mathrm{HS}}$ SBES + sonal $\succ_{\mathrm{HS}}$ SDB)
- Sedimentology (number of sensors) ($5 \succ_{\mathrm{S}} 4 \succ_{\mathrm{S}} 3 \succ_{\mathrm{S}} 2 \succ_{\mathrm{S}} 1 \succ_{\mathrm{S}} 0$)
- Ocean data (MTO + CTD + tide + current $\succ_{\mathrm{OD}}$ MTO + CTD + tide $\succ_{\mathrm{OD}}$ MTO + CTD $\succ_{\mathrm{OD}}$ MTO)

The remaining preference parameters identified during these three interviews are given in Tables II to IV. These preference parameters have been determined through a direct elicitation approach. The main reason for this choice is the poor availability of enough learning or assignment examples. Next to that, the experts' advanced knowledge eased the task of providing the separation profiles, the veto profiles, and the weights directly. For each expert we first present the separation profile(s) and the veto profile, before showing the criteria weights and the majority threshold.

B. *Quality parameters extraction*

Each of the three experts has provided the list of quality parameters that need to be taken into account in the overall assessment of the surveys. These information are given as inputs for the “Quality dimensions extraction / calculation” step of the process, together with the survey data and metadata. This step then generates the data of Table V.

TABLE II
PREFERENCE PARAMETERS FOR THE ACOUSTIC END-USE

Acoustician	POSACC	SOUACC	Cover.	Sensors	Sed.	Ocean data
G / B (b_1)	$< 5m + 5\%$	$< 1m$	150%	MBES	4	MTO + CTD + tide + current
Veto (G) (b_1^v)	·	·	·	·	1	·
Weights (w_j)	$2/22$	$1/22$	$4/22$	$6/22$	$6/22$	$3/22$
Majority threshold (λ)	$11/22$					

TABLE III
PREFERENCE PARAMETERS FOR THE HYDROGRAPHIC END-USE

Hydrographer	CATZOC	POSACC	SOUACC	Cover.	Sensors	Sed.	Ocean data
G / B (10m) ($b_{1,10m}$)	A1	$< 0.5m$	$< 0.25m$	200%	MBES	2	MTO + CTD + tide
G / B (100m) ($b_{1,100m}$)	A2	$< 5m + 5\%$	$< 1m$	100%	SBES + sonal	2	MTO + CTD + tide
G / B (1000m) ($b_{1,1000m}$)	B	$< 20m +$ 10%	$< 20m$	100%	SBES + sonal	2	MTO + CTD
Veto (G) (b_1^v)	·	·	·	100%	·	·	·
Weights (w_j)	$7/28$	$3/28$	$4/28$	$6/28$	$5/28$	$1/28$	$2/28$
Majority threshold (λ)	$14/28$						

TABLE IV
PREFERENCE PARAMETERS FOR THE OCEANOGRAPHIC END-USE

Oceano-grapher	POSACC	SOUACC	Cover.	Sensors	Sed.	Ocean data
G / A (b_2)	$< 2m$	$< 0.5m$	200%	MBES	4	MTO + CTD + tide + current
A / B (b_1)	$< 5m + 5\%$	$< 2m$	100%	SBES + sonal	2	MTO + CTD + tide
Veto (G & A) (b_1^v & b_2^v)	·	·	·	·	·	MTO + CTD
Weights (w_j)	$6/42$	$4/42$	$8/42$	$2/42$	$10/42$	$12/42$
Majority threshold (λ)	$21/42$					

TABLE V
PERFORMANCE TABLE: OUTPUT OF THE "QUALITY EXTRACTION / CALCULATION" STEP

Survey (depth)	CATZOC	POSACC	SOUACC	Cover.	Sensors	Sed.	Ocean data
S2019026 (1000m)	B/C	$< 20m + 10\%$	$20m$	200%	MBES	5	MTO + CTD + tide
S2015009 (10m)	C	2.5	$0.55m$	50%	SBES + sonal	2	MTO + CTD
S2012056 (20m)	B/D	5	$0.5m$	200%	Lidar	0	MTO
S2017026 (100m)	A1/B	3.6	$0.85m$	130%	MBES	2	MTO + CTD + tide

On this table we can observe that the quality of the CATZOC for the experts is not necessarily associated with a unique value (for the survey S2019026 we have for example two values of CATZOC: B and C). This CATZOC value can evolve depending on the area and the range of local depths acquired. During the "multi-criteria assesment" step, the most pessimistic value of the CATZOC has been kept (in order to be as conservative as possible for the safety of navigation).

In addition, for the measurement of POSACC and SOUACC the Shom relies on a calculation of total propagation of uncertainty (TPU) taking into account the entire bathymetric acquisition chain described in [4]. Here again, a single worst-case TPU value (one for vertical and one for horizontal TPU) is kept for all the bathymetric survey soundings.

C. Multi-Criteria assessment

Table V serves as input for the "Multi-Criteria assessment" step, together with the previously identified preference parameters of the three experts. The MR-Sort overall assessment method generates for each end-use an overall quality evaluation, which depends on the configuration of the process. The output of this assessment step is shown in Table VI

As we already said, the use of an MR-Sort model allows us to explain the results of the assessment. For example, for the hydrographic use, survey S2017026 at 100m is at least as good as the separation profile $b_{1,100m}$ (Table III) on all criteria except the CATZOC. Its concordance is therefore equal to $21/28$ which is greater than the threshold $\lambda = 14/28$. As a consequence there is a sufficient coalition of criteria supporting its assignment into category "Good".

Similarly, survey S2012056 for the oceanographic use is at least as good as the separation profile b_1 (Table IV) on four criteria, which leads to a concordance greater than the threshold λ. However its evaluation on the "Ocean data" criterion is MTO, which raises a veto and thus invalidates the outranking relation between S2012056 and b_1. Consequently this survey is evaluated as a bad survey instead of an acceptable one.

V. DISCUSSION

The process we propose has at least 3 advantages over a method that does not involve maritime expertise via preference models. First of all, as it can be seen in Table VI, the overall assessment of a survey is not necessarily the same for the three end-uses, and depends on the preferences of the expert and the context of the survey (here the average depth). More specifically this can be observed for survey S2017026 which is evaluated differently for each of the three end-uses.

Second, even if a survey is equally evaluated for two end-uses, there might be very different reasons which have led to this overall assessment. As an example, consider survey S2015009, which is evaluated as "Bad" both for the acoustic and the hydrographic end-uses. The reason of this bad evaluation for the acoustic end-use lies in the fact that, even if this survey is sufficiently good on the POSACC and the SOUACC quality parameters, the added weights of these parameters ($3/22$) is far from a majority coalition ($11/22$). For the hydrographic end-use, the bad overall assessment simply comes from the fact that all the evaluations of this survey are

TABLE VI
OVERALL ASSESSMENTS: OUTPUT OF THE "MULTI-CRITERIA ASSESSMENT" STEP

Survey (depth)	Acoustic use	Hydrographic use	Oceanographic use
S2019026 (1000m)	Good	Good	Good
S2015009 (10m)	Bad	Bad	Bad
S2012056 (20m)	Bad	Bad	Bad
S2017026 (100m)	Bad	Good	Acceptable

unanimously lower than the minimal requirements for a good survey.

Third, the Multi-Criteria assessment step allows to give recommendations for future surveys by providing, if necessary, clear explanations to the user. Let us consider as an example survey S2017026, which is considered as a bad survey for the acoustician, because the 3 criteria "POSACC", "SOUACC" and "Sensors" for which it is considered as good, are not enough for an overall "Good" evaluation. If however, for a future survey, one would improve the "Ocean data" or the "Coverage" criterion by only 1 level, this would yield a good assessment of this survey.

The identified data quality parameters of a survey and corresponding dimensions represent a particular characterization choice, compatible with the examined users' preferences profiles. Depending on the available data and metadata it is possible to further refine the quality representation, permitting to include more specific preferences of a user profile, without modifying the proposed quality assessment process for hydrographic surveys. Moreover, it is also possible to calculate data quality dimensions directly from raw data, like estimations of covered surface, missing data, depth variability, and data density depending on the size of analyzed surface cells. This implies however, that whenever new data quality characterization values are to be taken into account, the respective preferences elicitation must be adjusted accordingly.

VI. CONCLUSION AND PERSPECTIVES

The problem of hydrographic data quality assertion depends on multiple factors that make complex to verify if a given dataset is compatible with a specific user profile, defined by its preferences. Our work has defined and validated an approach combining the use of data quality parameters and multi-criteria decision aiding to solve this problem. Although, bathymetric metadata contain several quality parameters (in key-value form), these values do not give information about the use that could be made of a dataset and are not useful to all users. The main interest of the method presented here is to make the quality analysis explicable through an MCDA approach with respect to the user's criteria, by selecting the quality parameters and dimensions of the hydrographic survey.

Additional calculated data quality dimensions could be examined and compared to the estimations done by experts, in order to define a raw data indicator of confidence. On the other hand, knowing which user profile will exploit a given dataset, it would also be suitable to take account of information quality dimensions, as volume to verify if the available dataset can be used for the intended task, redundancy to confirm the interest of coping with missing data, and coherence to identify contradictions, among others. Depending on user preferences, more advanced data and information quality analysis could be included, matching essential and complementary parameters and dimensions.

To further show the interest of the proposed approach, we also intend on deploying it on more surveys in a real-life context. Next to that, instead of using a direct preference elicitation approach to determine the preference parameters of each expert, an indirect approach could be used, in which the expert is confronted with past surveys, which he has to assess. This information is then used by a learning algorithms which determines automatically the parameters of the preference model. Last but not least, currently the explanation of the output is done manually. In the future we plan to automatize this task through rule generation algorithms.

Our work clearly shows the interest of the proposed method for the use of a national hydrographic service. Indeed, from a survey and its associated metadata we can determine its overall level of quality, but also the usefulness of this dataset for different end-uses, as hydrographic, acoustic, or oceanographic purposes. It can therefore be used *a posteriori* assess the interest of a survey after its completion.

But our proposal could also be used *a priori* to plan future hydrographic surveys according to the needs expressed by the prescriber of the survey, allowing the survey to be planned to meet the exact requirements of the national hydrography programme [14] and avoiding over-quality.

VII. ACKNOWLEDGMENT

Thanks to Shom for providing access to the data in this research project as well as access to experts to model users' preferences.

REFERENCES

[1] D. Bouyssou and T. Marchant. An axiomatic approach to noncompensatory sorting methods in MCDM, I: the case of two categories. *European Journal of Operational Research*, 178(1):217–245, April 2007.

[2] D. Bouyssou and T. Marchant. An axiomatic approach to noncompensatory sorting methods in MCDM, II: more than two categories. *European Journal of Operational Research*, 178(1), April 2007.

[3] Denis Bouyssou, Thierry Marchant, Marc Pirlot, Alexis Tsoukiàs, and Philippe Vincke. *Evaluation and decision models with multiple criteria: Stepping stones for the analyst*. International Series in Operations Research and Management Science, Volume 86. Springer, Boston, 1st edition, 2006.

[4] Rob Hare. Depth and position error budgets for mulitbeam echosounding. *The International Hydrographic Review*, 72(2), May 2015.

[5] IHO. Iho standards for hydrographic surveys. Technical report, IHO, 2020. Special Publication N°44, 6 edition.

[6] R.L. Keeney and H. Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs*. J. Wiley, New York, 1976.

[7] A. Leroy, V. Mousseau, and M. Pirlot. Learning the parameters of a multiple criteria sorting method. In R. Brafman, F. Roberts, and A. Tsoukiàs, editors, *Algorithmic Decision Theory*, volume 6992, pages 219–233. Springer, 2011.

[8] Pedro Merino Laso, David Brosset, and John Puentes. Monitoring approach of cyber-physical systems by quality measures. pages 105–117, 07 2017.

[9] V. Mousseau, J. Figueira, and J.P. Naux. Using assignment examples to infer weights for ELECTRE TRI method: Some experimental results. *European Journal of Operational Research*, 130(2):263–275, April 2001.

[10] V. Mousseau and R. Słowiński. Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, 12(2):157–174, 1998.

[11] A. Ngo The and V. Mousseau. Using assignment examples to infer category limits for the ELECTRE TRI method. *JMCDA*, 11(1):29–43, November 2002.

[12] A-L. Olteanu and P Meyer. Inferring the parameters of a majority rule sorting model with vetoes on large datasets. In *DA2PL 2014: From multiple criteria Decision Aid to Preference Learning*, pages 87 – 94. Ecole Centrale Paris and Université de Mons, 2014.

[13] B. Roy. *Multicriteria Methodology for Decision Aiding*. Kluwer Academic, Dordrecht, 1996.

[14] Shom. Programme national d'hydrographie 2021-2024. Technical report, Shom, 2021. Available online: https://www.shom.fr/fr/qui-sommes-nous/programme-national-dhydrographie-pnh.

[15] O. Sobrie, V. Mousseau, and M. Pirlot. Learning a majority rule model from large sets of assignment examples. In *Algorithmic Decision Theory*, pages 336–350. Springer Berlin Heidelberg, 2013.

[16] O. Sobrie, V. Mousseau, and M. Pirlot. A population-based algorithm for learning a majority rule sorting model with coalitional veto. In H. Trautmann, G. Rudolph, K. Klamroth, O. Schütze, M. Wiecek, Y. Jin, and C. Grimme, editors, *Evolutionary Multi-Criterion Optimization*, pages 575–589, Cham, 2017. Springer International Publishing.