Definition and application of a quality estimator for multibeam echosounders

Yoann Ladroit\textsuperscript{a, b, c}, Xavier Lurton\textsuperscript{a}, Christophe Sintès\textsuperscript{b}, Jean-Marie Augustin\textsuperscript{a} and René Garello\textsuperscript{b}

\textsuperscript{a}Ifremer/IMN/NSE/AS, BP70, 29280, Plouzané, France
\textsuperscript{b}Telecom Bretagne/ITI, Technopôle Brest-Iroise, CS 83818, 29238, Brest Cedex 3, France
\textsuperscript{c}DGA/MRIS, 7 rue des Mathurins, 92221, Bagneux, France

*: Corresponding author: Yoann Ladroit, email address: yoann.ladroit@ifremer.fr

Abstract:

For modern multibeam echosounders, the use of a robust and reliable quality estimator associated with each sounding is an absolute necessity. Indeed, due to the large volume of data acquired, a lot of time is lost, both during the survey and the post-processing. This is a costly problem for hydrographers. The definition of a quality estimator based on the characteristics of the beamformed signal gives an answer to this problem. It has been successfully implemented by several sonar manufacturers and its relevance in measuring the quality of each sounding has been demonstrated.

The defined Quality Factor can also be used directly in the sonar measurement process, such as in detection algorithm, in order to improve the performance of the systems. This makes it possible to enhance of existing systems at very little cost with success.

Keywords: Equations; Estimation; Mathematical model; Q factor; Shape; Signal to noise ratio; Uncertainty

I. INTRODUCTION

Modern multibeam echosounders (MBES) collect very wide amounts of data, raising new issues for their validation, management and processing. Indeed, a modern shallow-water system, forming several hundreds of beams and pinging at a repetition frequency of 10 to 20 Hz, records up to several thousands of soundings per second. This obviously illustrates the technical improvements brought recently to MBES, but this also raises a new set of original problems associated with the post-processing of this data flow; the "hand-cleaning" of data which used to be applied on the scarce sounding data obtained by single-beam echosounders and even by first-generation MBES becomes obviously a tedious (if not impossible) task leading to a waste of time and money for survey companies and hydrography services.

To face this situation, the CUBE algorithm method [1] proposes to process statistically the soundings associated to the nodes of a seafloor-gridding mesh, implying the need for each sounding to be associated with a "confidence degree" expressing its measurement quality. The wide use of CUBE
today indeed made it possible to improve tremendously the efficiency of data processing in seafloor bathymetry surveys; however a difficult point of this approach lies in the definition of a relevant sounding quality estimator.

Moreover, the availability of such an estimator is also very useful for the real-time quality control needed for the optimisation of survey operations. Recording data with a poor level of accuracy (due to irrelevant settings of the system, or to an inadequate strategy of area coverage) may impact severely the success of a cruise (implying the necessity of running again lines of unsatisfactory data, and possibly leading to a low quality level of the final product). This emphasizes the need for a control tool providing an objective qualification of the sounding values, and delivering this information in real time making it possible to correct swiftly the acquisition parameters (sonar settings or survey configuration) for a better optimisation of the at-sea operation time.

The Quality Factor defined in previous works [2] [3] is a possible answer to these issues. It provides an estimate of the bathymetric uncertainty of each sounding, based on the acoustic signal received by a multibeam echosounder and processed for seafloor echo detection inside each individual formed beam. Its definition is derived from the variance computation of the estimated time of arrival of the signal, both for the cases of phase and amplitude detection modes [4]. The strength of this approach resides in the fact that only a few signal parameters are needed in the variance estimation, which is done jointly with the sounding computation itself. The Quality Factor is computable independently for each sounding, and it involves very little computation load and limited information about the echosounder characteristics. It is optimally implemented inside the detection process, and recorded to be used as a useful input parameter for further statistical post-processing.

In the following, we first present the two main sounding estimation methods used in multibeam echosounders and how we modelled their variances. These estimated variances alongside with the actual depth measurement results are used for defining the Quality Factor. Several improvements of the modelling in the QF definition are proposed by comparison with the previous work [2] [3]. Then we present some validation results obtained on both simulated and field data, confirming the relevance of the Quality Factor for depth accuracy estimation. A further part is dedicated to show a practical application of the Quality Factor in improving the sounding detection process. The conclusion discusses the advantages of using a unified quality estimator among the sonar manufacturers and the developments ongoing on this topic.
II. Sounding Detection Methods for Multibeam Echosounders

In the case of multibeam echosounders, soundings are obtained by a single estimation of the oblique range \( r \), corresponding to the estimation of the time of arrival (TOA) \( t \). For each beam, the steering angle \( \theta \) is assumed to be accurately known. Therefore, neglecting any angular uncertainty, the range and time uncertainties are related by: \( \delta r = c \delta t / 2 \) (where \( c \) is the sound velocity).

In MBES the TOA is classically estimated from the amplitude of the signal or from the interferometric phase [4].

A. Amplitude detection

The amplitude detection method is based on the computation of the center of gravity of the signal envelope. The TOA \( t_D \) is thus estimated as:

\[
t_D = \frac{\sum t_i a(t_i)}{\sum a(t_i)}
\]

where \( a \) is the digitized signal and \( t_i \), a time sample. The summation is done over the \( i \) indices.

This is a non-linear estimation of the TOA, which makes the direct estimation of the variance difficult but possible under a certain set of hypothesis.

1) Variance computation: At first, the following hypotheses are done on the received signal \( a(t) \) characteristics:

- square-shaped envelope of width \( T \) containing \( N \) samples
- Rayleigh fluctuating envelope (parameter \( b \))
- independent samples \( \text{Cov}(a(t_i), a(t_j)) = \delta_{i,j} \sigma_a^2 \)

with \( \sigma_a = b \sqrt{(4 - \pi) / 2} \).

With these hypotheses, the derivation of the variance of \( t_D \) from (1) leads to the following result:

\[
\delta t^2 \approx \frac{\sum_{i=1}^{N} (t_i - \bar{t})^2 \sigma_a^2}{N^2 m_a}
\]

\[
\delta t^2 \approx T^2 (4/\pi - 1) \frac{N + 1}{12 N(N - 1)}
\]

(2)

with \( m_a = b \sqrt{\pi / 2} \) [7].

This first expression was validated by Monte-Carlo simulations, in order to confirm that it correctly describes the actual variance even for small values of \( N \), which is actually the usual case. This was essential to the model validation since amplitude detection is usually used near nadir where the envelope width is minimum. From this result we can move on to more general situations, especially concerning the envelope shape since the square shape is far from reality.

2) Effective width and shape factor definition: Real signal envelopes are usually bell-shaped. As a first approximation, it can be said that the envelope is the shape of the across-track beam directivity. For a classical non-shaded linear array, it can be modelled as a squared cardinal sine [4].

To extend the result of (2) to any sort of bell-shaped envelope, it is first necessary to define a robust way to measure the width of such a fluctuating signal. We define this width \( \Delta t \) as twice the second order moment of the envelope, considering the fact that the estimation of \( t_D \) is based on the first order moment of the envelope (which is the center of gravity).

\[
\Delta t = 2 \sqrt{\int_0^{+\infty} s(t) \left( t - \int_0^{+\infty} s(u) du \right)^2 dt}
\]

where \( s(t) = \frac{a(t)}{\int_0^{+\infty} a(t) dt} \)

(3)

This will be considered as the "effective" width of the received signal [7].

Practically, this width is computed from the numerical signal, therefore we have:

\[
\Delta t = 2 / f_s \sqrt{\sum_{i=m}^{n} s_i \left( i - \sum_{j=m}^{n} s_j \right)^2}
\]

\[
N = f_s \Delta t
\]

where \( m \) and \( n \) are the first and the last sample number used for the computation of the center of gravity, where the \( s_i \) are the normalized amplitude samples and \( f_s \) is the sampling rate.

This effective width can be computed for several envelope shape which introduces a scaling factor \( B \). In the case of the square envelope, \( B \) is the ratio between the total width of the signal and the effective width. For other envelope shapes \( B \) factors have been computed using a technical computing software and are presented in Table 1.

<table>
<thead>
<tr>
<th>Shape</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate</td>
<td>( \sqrt{12} / 2 \approx 1.32 )</td>
</tr>
<tr>
<td>Sinc</td>
<td>1.04</td>
</tr>
<tr>
<td>Sinc2</td>
<td>0.96</td>
</tr>
<tr>
<td>Cos</td>
<td>1.05</td>
</tr>
<tr>
<td>Cos2</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 1: Results of computation for the corrective factor \( B \) as a function of the envelope shape

As a result, for most bell-shaped envelopes, the scaling factor is close to 1. This is a good thing since the actual
shape of the backscattered echo is only an approximate of these (about the across-track directivity of the beam).

It has been checked numerically that $B$ does not depend on the effective width $\Delta t$ and therefore, can be used as a general scaling factor in the variance equation. We obtain:

$$\delta t^2 = B^2 \Delta t^2 \frac{4/\pi - 1}{12} \frac{N + 1}{N(N - 1)}$$

This expression is valid in the case of a fluctuating envelope with independent samples. This is actually not the case in a real signal since the pulse length introduces a correlation between samples of the signal backscattered by the seafloor.

3) Impact of the pulse length: To model the signal correlation introduced by the pulse duration, as a first approximation, we consider that the pulse interaction with the seafloor can be expressed as a complex filtering of the received signal by a filter of length equal to the transmitted pulse duration.

Therefore, at a given time $t_i$ the received signal amplitude is written:

$$A(t_i) = \frac{1}{N_p} \sum_{k=i}^{i+N_p} a(t_k)e^{j\phi(t_k)}$$

where $N_p = \text{floor}(Tf_s)$, $\phi$ is the signal phase, uniformly distributed in $[-\pi, \pi]$, and $a$ is the amplitude signal, Rayleigh fluctuating (which is considered to be the case for a fully-developed speckle [7]).

A new expression for the TOA variance comes as:

$$\delta t^2 = B^2 \sum_{i,j} (t_i - \bar{t})(t_j - \bar{t}) \sigma_{i,j}$$

where $m_A = m_a/\sqrt{N_p}$ is the mean of the signal $A$ and $\sigma_{i,j} = \text{Cov}(A(t_i), A(t_j))$.

The resulting co-variance of the amplitude signal is shown to be:

$$\sigma_{i,j} = \begin{cases} 
0, & \text{if } |i - j| > N_p \\
\frac{\Delta t^2}{N_p} (N_p - |i - j|)^2, & \text{otherwise}
\end{cases}$$

Finally a new expression of the variance $\delta t^2$ of the TOA:

$$\delta t^2 = B^2 \Delta t^2 \frac{4/\pi - 1}{12} \frac{N + 1}{N(N - 1)} \sigma_{i,j}$$

$$S = \sum_{i=-N_p+1}^{N_p-1} (N - |i|)(N - 1 - |i|)(N + 1 - |i|)(N_p - |i|)^2$$

with $\Sigma$ the covariance matrix and $t$ the time vector.

As shown in Figure 2, the impact of the pulse length on the variance of the TOA is very important and cannot be ignored. It will have to be taken into account in the final Quality Factor formula.

The last parameter to account for in the model is the possible existence of additive noise on the signal.

4) Impact of the signal to noise ratio (SNR): The additive noise can be modelled as a complex Gaussian circular noise of zero mean and variance $\sigma^2$.

Under the hypothesis of sample independence a new expression similar to Equation (5) can be derived:

$$\delta t^2 = B^2 \Delta t^2 \frac{(N + 1)N}{12d^2} \frac{(N - 1)(N_{\text{sig}}/r_{\text{SNR}} + N)^2}{(N_{\text{sig}} + 1)N_{\text{sig}}}$$

$$+ \frac{1}{r_{\text{SNR}}^2} \Delta t_{\text{sig}}^2 \frac{(N_{\text{sig}} - 1)(N_{\text{sig}}/r_{\text{SNR}} + N)^2}{(N_{\text{sig}} + 1)N_{\text{sig}}}$$

with $N_{\text{sig}}$ the number of samples in the envelope, $\Delta t_{\text{sig}} = N_{\text{sig}}/f_s$ ($f_s$ being the sampling frequency), $\Delta t$ the effective width of the envelope (equivalent to $N$ sample) re-evaluated for each SNR value, $r_{\text{SNR}}$ the SNR in linear scale.

In Figure 3, this new expression 10 is compared to numerical simulations and to Equation (9) without the SNR taken into account.
This shows that the previous model (9) describes correctly the computed variance for SNR above typically 10 dB. For lower SNR values it is no longer the case due to the fact that the scaling factor \( B \) does not correctly describe the shape of the envelope. The new model (10) fits correctly the computed variance but it involves an evaluation of the SNR on the signal, which is what we wanted to avoid.

In fact this is not a problem in real data processing. Indeed, soundings that will be considered for amplitude measurement will be coming from steep angles, where the SNR is well above 10 dB. In addition, by a simple test on the shape of the echo inside the considered part of the received signal, it is easy to discard amplitude measurement where the bell-shape is no longer a good approximation.

**B. Phase detection**

The phase signal is the result of the interferometry between two subsets of the receiving array. In each beam, we assume that the phase evolves as a noised linear ramp \( \phi(t) = \alpha t + \beta \).

The general method to estimate the TOA, is to estimate the parameters \( \alpha \) and \( \beta \) of the ramp either by classical linear regression or by weighted linear regression.

1) Variance computation: We consider a phase ramp containing \( N \) samples. To compute the TOA variance, we make those hypothesis about the phase samples \( \phi(t_i) \):

- \( \phi(t_i) = \alpha t_i + \beta + \epsilon_i \)
- stationary noise \( \epsilon \) on the time window considered (homoscedasticity)
- gaussian noise, zero-mean, variance \( \sigma^2 \)
- independant phase samples

Under those hypotheses, the classical linear regression gives an unbiased estimation of \( \alpha \) and \( \beta \), and therefore of the "true" phase ramp. The variance of the phase estimation \( \hat{\phi}(t_i) \) can be computed and results in:

\[
Var(\hat{\phi}(t_i)) = \sigma^2 \left( \frac{(t_i - \bar{t})^2}{\sum_{j=1}^{N} (t_j - \bar{t})^2} + \frac{1}{N} \right) \quad (11)
\]

This involves an estimation of the parameter \( \sigma \), which is the unbiased estimator:

\[
\sigma^2 = \frac{1}{N - 2} \sum_{i=1}^{N} e_i^2
\]

where \( e_i = \hat{\phi}_i - \phi_i \) \quad (12)

Finally, as a first approximation, we estimate the TOA variance \( \delta_t^2 \) as:

\[
\delta_t^2 = Var(t_D) = \frac{Var(\hat{\phi}(t_i))}{a^2} = \sigma^2 \left( \frac{(t_D - \bar{t})^2}{\sum_{i=1}^{N} (t_i - \bar{t})^2} + \frac{1}{N} \right) \quad (13)
\]

where \( t_D \) is the estimated TOA, and \( a \) the estimation of the slope \( \alpha \) given by the linear regression.

This is only a first-order approximation of the variance. Figure 4 shows the confidence interval of the estimation on a phase ramp of dynamic \( \pi \), with a strong noise of standard deviation \( \sigma = \pi/3 \).

![Figure 4](image_url)

We see that the estimation made is quite robust. Indeed, when the window is centered (\( t = 0 \)) 95% of the estimated variance values will be at most 20% from reality, which means that the estimated standard deviation will be less than 5% from its true value.

This figure also emphasizes the importance of window centering.
2) Impact of the window centering: Figure 4 shows that the TOA variance increases noticeably as the zero-crossing happens far from the center of the time-window. This reminds that due to the property of the TOA estimator chosen, it is better to compute a second estimation with a centered time-window around the first estimation to reduce the sounding variance.

The weighted linear regression gives the same result at the center of the time window, giving a single expression for the TOA variance, with the recommendation that the final phase measurement has to be done on a centered time-window.

Also, since the previous work was done with the hypothesis of independent samples, the number of samples N in the phase ramp will be replaced by the number of independent samples N_A, that is to say: N_A = N/N_p.

Under these conditions, the time-standard deviation is:

\[ \delta t = \frac{a}{\alpha \sqrt{N_A}} \]  

where \( a \) is the estimated slope of the phase ramp, \( N_A \) the number of independent samples used in the phase ramp fitting and \( \sigma \) the estimated standard deviation of the phase.

From the two definitions of \( \delta t \) determined from amplitude or from phase (Equation (9) and (14)), the TOA standard deviation for each sounding can be computed and gives an information about the acoustic component of its uncertainty of it. This can be considered as an estimation of the sounding quality.

III. DEFINITION OF THE QUALITY FACTOR

The Quality Factor for multibeam echosounders is defined [2] as the relative depth (z) uncertainty:

\[ QF = \log_{10}(q_F) = \log_{10}\left(\frac{z}{\delta z}\right) \]

A bathymetry measurement generally features a joint estimation of angle \( \theta \) and TOA \( t \) so that:

\[ \frac{\delta z}{z} = \left(\frac{\delta t}{t} + \tan(\theta) \delta \theta\right) \]  

In MBES measurement \( \delta \theta \approx 0 \), since the angle is fixed by the beam steering. Therefore:

\[ \frac{\delta z}{z} = \frac{\delta t}{t} \]  

\[ QF = \log_{10}\left(\frac{t}{\delta t}\right) \]

High QF values correspond to accurate measurements. Typically, field values are spread from 2 to 3, corresponding to relative depth errors from 1% to 0.1%.

A. Amplitude detection

During the amplitude detection process, the QF associated with sounding can be easily computed alongside the sounding value itself. The relative time uncertainty can be obtained from Equation (9) and transferred into the quality factor:

\[ QF_A = \log_{10}(q_{f_A}) \]

\[ q_{f_A} = \frac{t_D}{\delta t} \]

\[ q_{f_A} = \frac{\sqrt{12}}{B\sqrt{4/\pi - 1}} \frac{t_DNN_p}{\Delta t\sqrt{S}} \]

With \( t_D \) the TOA estimation and:

\[ S = \sum_{i=-N_p+1}^{N_p-1} (N-|i|)(N-1-|i|)(N+1-|i|)(N_p-|i|)^2 \]

\[ S = \frac{N_p}{30} (20N^3N_p^2 + 10N^3 - 15N^2N_p^3 + 15N^2N_pN_p^4 + 6N - 20N^2N_p^2 - 16N - N^5 + 5N^3 - 4N_p) \]  

It is important here to notice that at a fixed range, the wider the envelope, the lower the QF, and the shorter the transmitted pulse, the higher the QF. This expression will be validated on real and simulated data in the next part.

B. Phase detection

The expression for the Quality Factor of the phase detection is given in Equation (19).

\[ QF_\phi = \log_{10}(q_{f_\phi}) \]

\[ q_{f_\phi} = \frac{a t_D \sqrt{N}}{\delta \Delta \Phi \sqrt{N_p}} \]  

where \( \delta \Delta \Phi \) is the estimated variance and \( t_D \) the TOA estimation.

This expression is valid only if the phase ramp is approximatively centered around the final TOA estimation. Otherwise it is possible to obtain more precise results with :

\[ QF_\phi = \log_{10}(q_{f_\phi}) \]

\[ q_{f_\phi} = \frac{a t_D}{\delta \Delta \Phi} \frac{1}{\sqrt{\left(\sum_{i=1}^n (t_i-t)^2 + \frac{1}{N/N_p}\right)}} \]

Expressions (18) and (20) make it possible to estimate the vertical uncertainty of soundings according to the computation method used. They have been tested on simulated data and Implemented on existing systems in order to evaluate their relevance.
IV. VALIDATION ON SIMULATED AND FIELD DATA

A. Methodology

The principle of the QF validation over real data is to estimate the sounding variance on several pings over a homogeneous area. First this was done on simulated MBES data, using a signal simulator based on the extended scatterers model. Then, we used real data recorded on two echo-sounders installed on RV Pourquoi Pas?.

The soundings obtained from phase and amplitude detections are compared to a reference digital terrain model (DTM), in order to compute an estimate of their actual variance. On simulated data, the reference DTM is directly known, since it is part of the simulation input. On real data, if a reference DTM is not available for the considered area, it is created from the data by smoothing the detection result at a relevant scale. This is classically what is done to assess the performance of a bathymetric system during its sea acceptance test.

B. Validation on simulated data

The simulated data used here have been created with the following characteristics:

- flat seafloor
- 20m depth
- 300kHz system
- beam aperture along-track: 1°
- beam aperture across-track: 1°
- pulse length: 0.3ms
- total aperture (across-track): 120°

The simulator gives the complete beamformed signal for each ping; a classic bottom detection algorithm is then applied on the data. The algorithm gives a TOA estimation inside each beam alongside with its computed Quality Factor.

One ping is presented in Figure 5:

![Figure 5. Simulated backscattered signal after beamforming as a function of beam angle and time, in amplitude (left) and phase (right)](image)

110 pings were generated on this simulated seafloor giving the depth uncertainty results presented in Figure 6.

The agreement is very good between the fluctuation of the simulated soundings and the QF estimation. This shows that the QF is able to predict reliably the sounding quality.

C. Validation on real data

The QF computation has been implemented in the bottom detector module of a Reson Seabat 7111 and a Reson Seabat 7150 aboard RV Pourquoi Pas?. It was also made part of the bottom detection algorithm of the sounder as explained in part V.

The data used for the QF validation was recorded under the following conditions for the Seabat 7111:

- flat seafloor
- 25m depth
- 100kHz system
- beam aperture along-track: 1.8°
- beam aperture across-track: 1.8°
- pulse length: 0.4ms
- total aperture (across-track): 152°

We had full access to the complex beamformed signals, so that we were able to control every step of the bottom detection (see part V) and Quality Factor computation.

The variance was estimated on a smoothed DTM obtained from the original data, and using several hundreds of pings.

Again the comparison shows that the QF describes correctly the observed quality of the data. The same type of analysis was
conducted using the Seabat 7150 at 12 kHz in 2200 meters of water depth with the same good results.

These validations on simulated and real data illustrate the relevance of the Quality Factor and its capacity to predict correctly the sounding uncertainty.

This sounding accuracy estimation available through the Quality Factor has been, and is still currently tested in several applications.

V. QF-ASSOCIATED APPLICATION: BOTTOM DETECTION ALGORITHM

The development of the QF was done in order to improve efficiency of both the survey operation control and post-processing tasks.

It can also be made a part of the sounding computation itself. As mentioned in the previous part, the QF computation has been implemented in the bottom detection modules of two MBES. This made it possible to improve the detection algorithm by using the QF all along the TOA estimation process.

This process begins by selecting a wide time-window on the signal time series. This signal section is then used to compute the first TOA estimations (both from phase and amplitude processing). A first QF value is computed alongside this rough estimation, and its associated \( \delta t \) is used to define a reduced time-window around it. This iterative process goes on until the selected part of the signal (defining the spatial resolution of the measurement) has been decreased to a level providing a satisfactory trade-off with the resulting sounding variance.

Finally, on the last step of the algorithm, the QF is used as a criterion of choice between each pair of sounding solutions (coming from phase and amplitude processing) computed inside each beam.

Figure 8 presents a line run with the Seabat 7111 and processed with two different bottom detection algorithms. In the first version (left), the QF is not used. In the second version (right), the QF is used along the detection process as described above, which results in a lower sounding variance and a higher number of sounding detections in the final result. On this data set, the use of the QF made it to detect correctly 9% more soundings, and to reduce the possible fluctuation of the data.

This makes clear the interest of using the QF for increasing the number of valid sounding detections, and for decreasing the sounding fluctuation rate (see inside the white circles).

This capability comes in complement to the quality estimation of the resulting sounding values, and participates in the improvement of the global efficiency and quality of the seafloor survey operations.

VI. CONCLUSION AND PERSPECTIVES

The Quality Factor presented above provides very promising perspectives for improving the efficiency of real-time and post-processing operations in seafloor survey using multibeam echosounders.

It has been tested and validated on several systems, and has showed good results when compared with the actual variability of the collected data. Its implementation inside a MBES can improve the system performance at nearly no cost, and it gives a reliable quality information for each resulting sounding. This quality information may then be used either in real time for the survey control, or in post-processing applications.

Its use in the new version of the algorithm CUBE [1] is presently under study. The idea is to use the QF as the sounder-dependent component in the Hare-Godin-Mayer [6] uncertainty model whose results are used in CUBE as a confidence level attributed to the sounding values. This should improve the relevance of the sounding hypothesis provided by CUBE by giving a more appropriate acoustical uncertainty value to each computed sounding.

Moreover, since it is now established that the QF provides a good estimator of the acoustical measurement uncertainty, its computation is being implemented by several manufacturers in order to be applied systematically during the sounding processing, and finally provided in bathymetry datagrams [5].

REFERENCES