ACCURATE PROCESSING OF MULTIBEAM ECHOSOUNDER BACKSCATTER BASED ON STATISTICAL ANALYSIS OF SNIPPETS DATA

S Homrani	Lab-STICC UMR CNRS 6285, ENSTA Bretagne, Brest, France
I Mopin	Lab-STICC UMR CNRS 6285, ENSTA Bretagne, Brest, France
O Morio	Service Hydrographique et Océanographique de la Marine, Brest, France
PY Le Moigne	Service Hydrographique et Océanographique de la Marine, Brest, France
H Moreau	Lab-STICC UMR CNRS 6285, ENSTA Bretagne, Brest, France
G Le Chenadec	Lab-STICC UMR CNRS 6285, ENSTA Bretagne, Brest, France

1 INTRODUCTION

Mapping geological resources and benthic habitats may be achieved using the bottom acoustic backscatter strength (BS), i.e. the seafloor response to an incident acoustic wave. This response depends on the type of seafloor, in relationship with its roughness and acoustic impedance[1]. The backscattered acoustic signal can be obtained from measurement of sonar systems such as multibeam echo-sounders (MBES), through the sonar equation. However, some of the terms implicated in the sonar equation need to be compensated for[2], to get an accurate estimation of BS. Terms to compensate include both technical instrument related parameters and environment-dependant parameters. Over the past decade, the backscatter user community has made significant efforts to develop a reproducible correction-calibration processing chain[3,4]. At the same time, MBES manufacturers have made raw time-series of seafloor echoes per beam (snippets) available and reliable[5]. In practice, the application of the correction-calibration chain to BS measurements, and snippet amplitudes especially, is still rarely implemented due to the inherent complexity of the process.

In more details, despite the deterministic approach with which the seafloor backscatter is mainly considered when modeled or processed in real-time during surveys, the physical parameter "acoustic seafloor backscatter" is first and foremost a random variable. It indeed arises from a stochastic process that can be modeled by random locations of acoustic scatterers within the insonified area[6]. The estimator to be used to obtain an accurate backscatter estimate from recorded seafloors echoes, e.g. for seafloor characterization, is then directly dependent on the distribution of the seabed echoes time-samples (snippet) amplitudes[7]. A Rayleigh distribution is commonly expected from raw snippets recordings, whereas the normality of the distribution tends to increase with successive averaging processing steps[8]. Due to their very high resolution, MBES data may contain few independent snippet samples per beam. Performing a statistical analysis on snippet samples is consequently more challenging than e.g. with single beam echosounder measurements[7]. The aim of the present work is to propose an estimation method of backscatter using seabed echo time series (snippets) provided by MBES. Especially, we first check that the statistical hypothesis on which the estimation of BS rely on are verified. The probability distribution of fully corrected independent snippet samples from a supposed homogeneous seafloor is investigated. A focus is realised on oblique incidence angular sectors.

The dataset used in this work arises from a survey conducted by the French Naval Hydrographic and Oceanographic Service (SHOM), on the French Atlantic continental shelf, South-West of Brittany, in depths between -45 and -95 m (Lowest Astronomical Tide), over rocky to muddy seafloors. The dataset includes: raw acquisitions from a Kongsberg EM710 MBES, the resulting SHOM bathymetric map, and a surface sedimentary map derived by the SHOM expert sedimentologists from in situ measurements over the area (including sediment grab samples and a backscatter mosaic computed from MBES surveys). This article first proposes a summary of the theoretical framework that should be followed to estimate seafloor acoustic backscatter. Then, their application to MBES data is addressed, with a dedicated focus on the insonified area and incidence angle computation and correction. In a third part, inferred statistical distributions of processed snippet samples are presented and discussed. Finally, conclusions and perspectives are drawn in order to pursue this work towards an accurate estimation of the seafloor backscatter[3].

2 ESTIMATING BS FROM MEASUREMENTS

In this section, we review the theoretical foundations of seafloor acoustic backscatter estimation to highlight the hypothesis on which the snippet values are derived.

2.1 Definition and Sonar Equation

The seabed acoustic backscattering strength or backscatter corresponds to the seafloor response to an incident acoustic wave i.e it describes the way the seafloor absorbs and or sends back the incoming acoustics intensity. Formally, we consider the ratio between the acoustic backscattered intensity and the acoustic incident intensity in decibels. The backscatter is a function of the incidence angle[9], as distinct physical processes are at stakes depending on how the wave hits the rough interface and the acoustic scattering properties of the seafloor.

The seafloor backscatter can be estimated from MBES recorded data using the sonar equation[3] (cf. figure 1). The backscatter level is then calculated as:

$$BS(\theta_i, t) = EL(t) - [SL + DI_{TX} + 10\log_{10}(\mathcal{D}_{TX}(\theta_p)) + G_{calib}(\theta_p)] + 2TL + 10\log_{10}(\mathcal{A}(\theta_p, \theta_i, t)) \quad (1)$$

where EL is the echo-level measured by the echosounder, SL is the transmitted source level, DI_{TX} is the transmission antenna directivity index, $\mathcal{D}_{TX}(\theta_p)$ is the transmission beam pattern with θ_p the pointing angle, TL embeds the transmission losses including spherical spreading and absorption losses such as $TL = 40\log_{10}(r) + 2\alpha r$ with r the range and α the absorption coefficient, $\mathcal{A}(\theta_p, \theta_i, t)$ is the instantaneous insonified area. $\mathcal{A}$ is shaped by the beam geometry (directivity function), the duration of the transmitted signal τ, the morphology of the seafloor (macroscopic slope), and the incidence angle θ_i. Its computation is detailed in paragraph 3.3. $G_{calib}(\theta_p)$ is the angular calibration curve, which may be obtained from calibration on a reference area[3], applied to BS data to obtain an absolute, system-free, level.

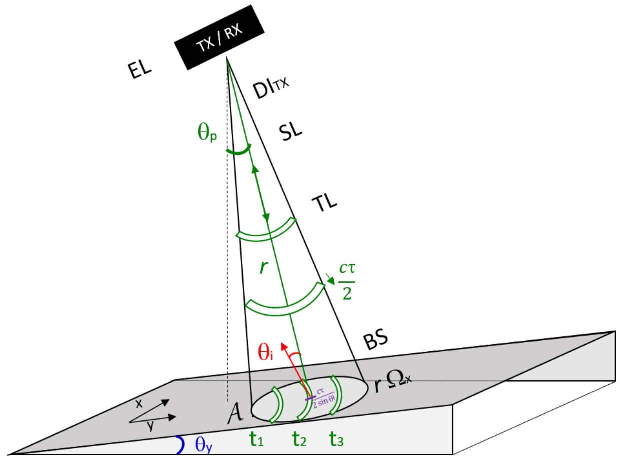

Figure 1: Illustration of the sonar equation.

2.2 Snippets and theoretical hypotheses to estimate BS

For our study, the chronology of the pulse interaction with the seafloor must be considered (cf. figure 1). At each time, a portion of the transmitted pulse reaches a different part of the seafloor. As the pulse spherically propagates, there is a delay dt between the moment when the pulse front and the pulse end reach respectively the seafloor, such as $dt = \frac{c\tau}{2\sin\theta_i}$. This time delay defines the instantaneous insonified area (within the beam footprint). Given the structure of natural seafloors, the pulse successively hits different scatterers with randomly varying backscattering properties (orientation, impedance, size). The scattered intensity is therefore a random variable thus so is $BS = 10\log_{10}(bs)$. $\sqrt{bs}$ as a random variable is expected to follow a Rayleigh distribution[6,7,8] when the insonified area is wide enough to contain a large number of scatterers.

In order to obtain an accurate estimator of BS as a random variable, for instance a value of the backscatter for an incidence angular sector, it is required to meet the following hypotheses[7]:

1. The sonar equation parameters should be perfectly known and seafloor echo samples corrected from them. Especially, the insonified areas and the incidence angles have to be well taken into account, given the slope of the seafloor topography;
2. The seafloor should be homogeneous at the scale of the group of snippets that are used to derive an estimate of BS;
3. The seafloor echo samples have to be statistically independent to ensure the validity of the mathematical development[7] that ensue to the best estimator of BS, i.e. the successive instantaneous insonified areas have to be non-overlapping;
4. The seafloor backscattering strength as a random variable should follow a Rayleigh distribution i.e. snippet amplitudes corrected from sonar equation $\sqrt{bs}$ should be realisations of a Rayleigh random variable.

These hypotheses define a theoretical framework for an accurate estimation of the backscattering strength. In this article, the three first hypotheses are guaranteed on our dataset by our post-processing method (see section 3), the fourth one is tested and discussed in section 4.

3 APPLICATION TO MULTIBEAM DATA

This work is performed on data from a Kongsberg EM710 multibeam echosounder. Acquisitions were made using a *Very shallow, Dual-swath dynamic* mode which parameters are given in Table 1. Three transmission sectors are generated, whose frequencies are alternated between the two swaths for one ping. Beams were formed with *High Density Equidistant* mode, providing 400 beams for each swath.

Frequencies f	Pulse Length τ	Beam widths (Ω_x,Ω_y)	Swath coverage
swath 1 : 70 kHz, 80 kHz, 70 kHz	300 μs	1°,1°	± 65°
swath 2 : 90 kHz, 100 kHz, 90 kHz	300 μs	1°,1°	± 65°

Table 1: Kongsberg EM710 multibeam acquisition parameters

Within Kongsberg EM710 echosounder raw data, two types of reflectivity data are available: the beam reflectivity (dB) and the snippets backscatter time series (dB). Snippet samples are taken from a restricted portion of the seafloor echo centered on the detection (sounding) of a given beam, meaning that multiple reflectivity measurements are available from each beam. The range of the beam sounding is provided, as well as the ordering of snippet samples and the sampling frequency, so that the user is able to recompute the range of each snippet sample for each beam. The beam reflectivity provided in the raw files is an estimate of BS. It is given in decibels and computed as the sample mean of the snippet backscattering strengths bs_i.

3.1 Sonar equation corrections

Beam reflectivity and snippet backscatter $BS_i = 10\log_{10}(bs_i)$ from MBES raw binary files data are processed similarly. The only distinction lies in the supplementary step preliminary applied to snippets samples to suppress the Lambert law variations against angle and normal incidence high levels[5]. Despite the benefits of this processing in real time, we corrected these compensations in order to retain only the seafloor response $BS(\theta_i)$. In addition, the transmission level and beam pattern are corrected for each beam θ_p (cf. equation 1). The Kongsberg default calibration file is used in $G_{calib}(\theta_p)$.

In both beam reflectivity and snippet backscatter, the insonified areas $\mathcal{A}$ computed in real-time are derived based on the strong assumption that the seafloor is perfectly flat and is perpendicular to the axis of the beam with the shortest measured range[5]. In the following, we want to take into account the seafloor slope therefore we correct the insonified areas using the following approximation[2,10]:

$$\mathcal{A} : \begin{cases} \dfrac{\Omega_x \Omega_y r^2}{\cos\phi_x \cos\theta_{i,y}} & \text{Normal incidence} \quad (2a) \\ \dfrac{c\tau\Omega_x r}{2\cos\theta_x \sin\theta_{i,y}} & \text{Oblique incidence} \quad (2b) \end{cases}$$

with $\theta_{i,y}$ the incidence angle of the beam sounding point in the across-track direction. The angle $\theta_{i,y}$ is corrected for ray-bending at bottom hit. For across track incidence angles at the limit between normal and oblique incidences, the smallest of the two areas is used.

The seafloor slope modifies the insonified area but also the incidence angle (cf. figure 1). It is consequently recomputed[10] taking into account seafloor slope. Slopes were obtained using a Digital Terrain Model (DTM) derived from the same echosounder data by expert hydrographers of SHOM following the OHI norm requirements. Slopes were computed in the North and East directions in RGF93 - Lambert 93 geodetic system, with the *moving slope method*[11] applied with three support points and a model order of two. They were then projected on the along and across directions using the heading stored in raw data. During the DTM processing, some soundings have been flagged as incorrect by the hydrographer. In the following backscatter calculations, these soundings are identified and suppressed from the data.

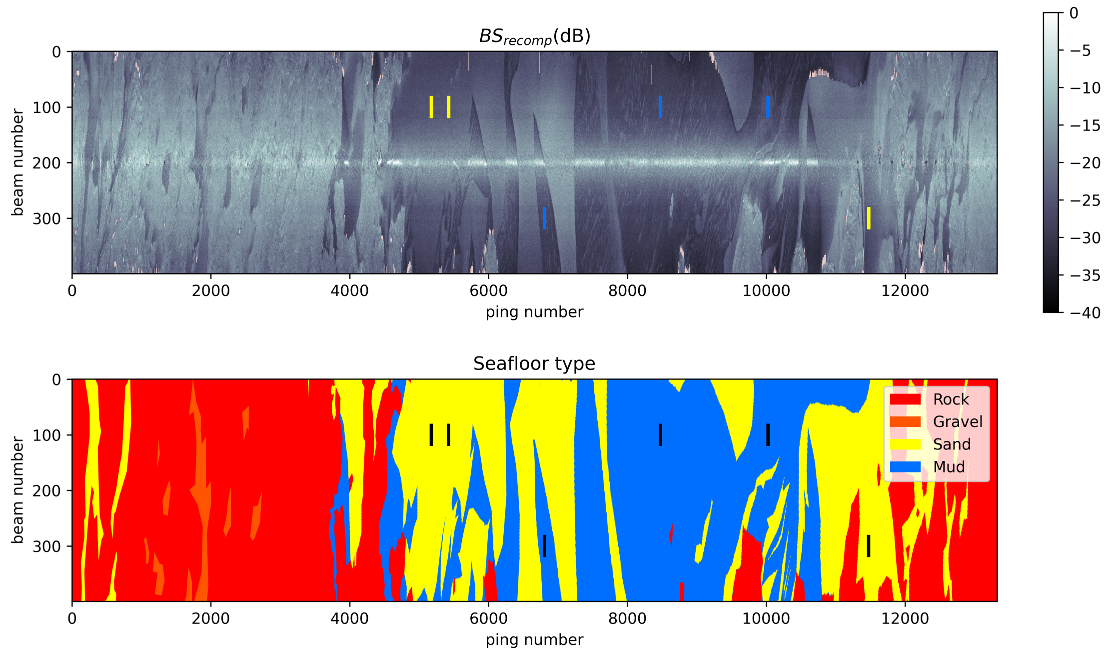

Figure 2: Homogeneous patches (rectangles) on which test of the Rayleighness of the snippet samples amplitudes x were made. Up: Corrected beam backscatter, (suppressed soundings are coloured in light pink) ; Down: Sedimentary map from SHOM experts projected in the MBES raw data frame.

3.2 Independence of snippet samples and choice of homogeneous seafloor patches

Two criteria are used two ensure that snippets samples are independent i.e. they come from non-overlapping insonified areas of the seafloor[7] (cf. figure 1). The first criterion relates to the instantaneous insonified area in the across track direction. In this direction, this area is bounded by the pulse length (cf. figure 1). Accordingly, the snippet samples used for statistical analyses must at least be separated by this temporal distance. For the incidences studied in this work the oblique incidences, snippets contain around 15 samples for oblique incidences. Given the pulse length and the snippet sampling frequency, each beam snippet time series yield three independent samples. In order to perform statistical analyses, a large number of samples are required. This leads to the introduction of the second criterion of samples independence, in the along-track direction. In this direction, the insonified area is bounded by the beam aperture. To be independent, the snippets samples must therefore come from pings whose beam footprints do not overlap. Given the shallow depths of the surveyed area, three successive pings overlap. To meet the criterion, we had to retain only one ping out of four which drastically sub-sampled the dataset.

The survey line chosen as example in this paper includes variety of seafloor types and topography (cf. figure 2): it features two rocky platform zones, with small patches of gravels, and an increased, lower basin, mapped with alternated bands of sand and mud. The rocky platforms show high roughness from rocky outcrops, while the lower basin depicts some smooth large scale seafloor waves on the sandy areas. The sediment labels given in figure 2 are provided by expert sedimentologists of SHOM, based on diverse acoustic sensors and terrain samples.

Six small patches of homogeneous zones are extracted from the survey line, on sandy and muddy terrains. The homogeneity of each small area is assessed from the homogeneity of the seafloor nature indicated on the sedimentary map (cf. figure 2) .

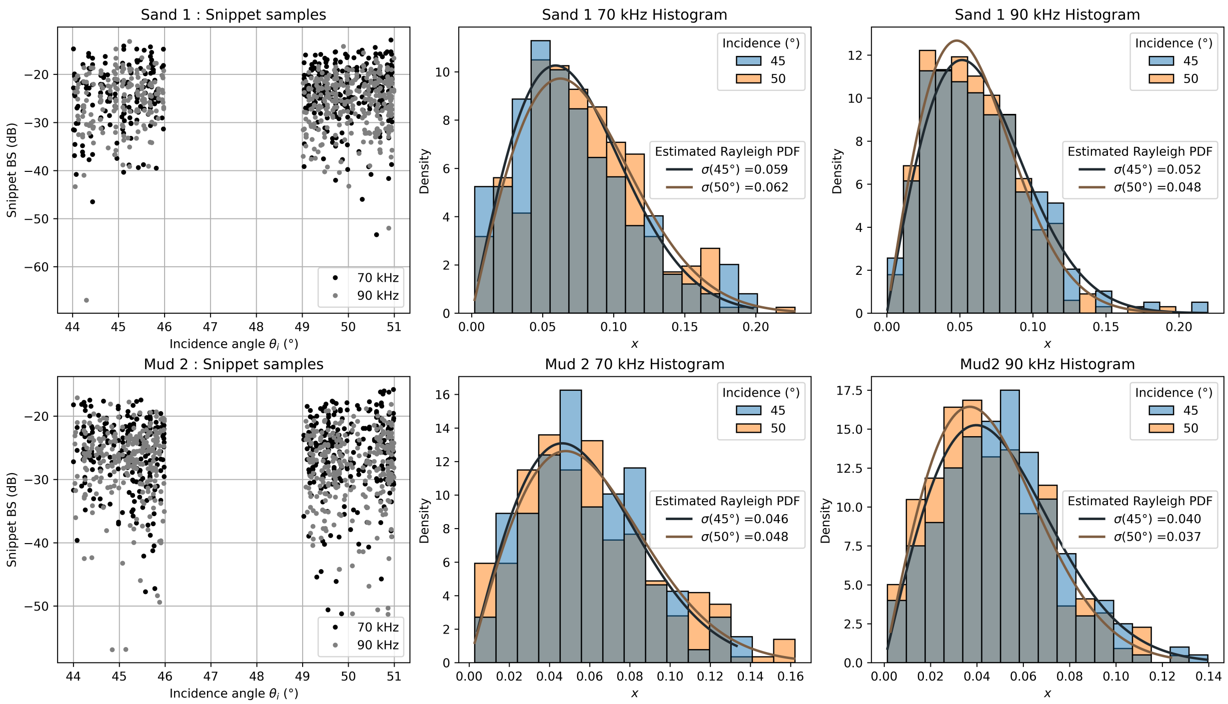

Figure 3: Examples of the snippet sample amplitudes analysis for two homogeneous of the six studied seafloor patches (figure 2). The two external transmission sectors of the Kongsberg EM710 (70 and 90 kHz), and two oblique incidence angles (45 ° and 50 °) are considered. Left: Snippet backscattering strength BS. Right: histograms of snippets amplitude x and inferred Rayleigh distributions of σ parameters. Grays histogram regions correspond to the overlap of the two 45 ° and 50 °histograms.

4 TESTING THE RAYLEIGH DISTRIBUTION HYPOTHESIS

Statistical analyses are performed on snippet samples coming from homogeneous seafloor patches (cf. figure 2). Within these patches, samples are grouped by transmitted frequency and incidence angles on the seafloor. Two particular oblique incidence angles are used as examples (cf. figure 3): 45° $\pm$ 1° and 50° $\pm$ 1°. The Rayleighness of the incidence group is assessed through a Kolmogorov-Smirnov test. This test is based on a comparison of the cumulative distribution function (CDF) of the theoretical distribution and the CDF of the snippet incidence group. Two parameters are estimated from the analysis: the parameter of the Rayleigh distribution σ and the p-value of the Kolmogorov-Smirnov test p. The parameter σ is estimated from snippet amplitudes (i.e. $\sqrt{bs_i}$) of groups as the Maximum Likelihood Estimate. The p-value indicates the probability that a sample, coming from a random variable following the same distribution as the tested sample is supposed to follow, has a more extreme Kolmogorov-Smirnov test statistic value than the tested sample. If the p-value is above 0.05, i.e. for a 95 $\%$ confidence interval, the Rayleigh distribution hypothesis cannot be rejected and the test is passed.

Twenty-four snippet groups were obtained from the seafloor patches-emission frequency-incidence angle possible combinations (cf. Table 2). Each group contains between 150 and 300 snippet samples, this number depending on the seafloor slope. For the 24 groups, the test was passed with 87.5 $\%$ of success. For the four tests performed on each seafloor patch, the success rate varies from 75 to 100 $\%$, from a seafloor patch to another. The amplitudes of the snippets whose incidence groups have successfully passed the test can be used to derive an estimation of backscatter BS of the beam they are taken from, using the square sample mean. This estimator is demonstrated[7] to be an unbiased and efficient estimator of the backscatter under the Rayleigh distribution hypothesis.

No clear relationship was found between success or failure at the test, versus the seafloor nature, nor the frequency or the incidence angle. By comparing the Rayleigh Probability Density Function (PDF) and the snippet group histograms (figure 3) for failed tests or successful test with low p-values, two non Rayleigh-typical behaviours were observed. Either the histogram is more symmetrical than the inferred theoretical Rayleigh PDF, or histogram bins with a higher density than the theoretical PDF occur at the tail of the distribution. The second behaviour rather suggests a bimodal distribution. This kind of distribution might be associated with heterogeneous patches of seafloor within the analysed seabed portion[12]. The homogeneity of the seafloor was assessed here using the sedimentary map. However, the map was drawn at a large scale. Thus, studied patches might actually contain sediment heterogeneous sub-patches at a finer scaler.

<table>
<tr><th>f</th><th>θ_i</th><th></th><th>Sand 1</th><th>Sand 2</th><th>Sand 3</th><th>Mud 1</th><th>Mud 2</th><th>Mud 3</th></tr>
<tr><td rowspan="4">**70**</td><td rowspan="2">**45°**</td><td>σ</td><td>0.059</td><td>0.060</td><td>0.053</td><td>0.042</td><td>0.046</td><td>0.050</td></tr>
<tr><td>p</td><td>0.02</td><td>0.09</td><td>0.30</td><td>0.70</td><td>0.16</td><td>0.32</td></tr>
<tr><td rowspan="2">**50°**</td><td>σ</td><td>0.062</td><td>0.057</td><td>0.053</td><td>0.044</td><td>0.048</td><td>0.047</td></tr>
<tr><td>p</td><td>0.64</td><td>0.28</td><td>0.74</td><td>0.06</td><td>0.31</td><td>0.82</td></tr>
<tr><td rowspan="4">**90**</td><td rowspan="2">**45°**</td><td>σ</td><td>0.052</td><td>0.058</td><td>0.065</td><td>0.045</td><td>0.040</td><td>0.046</td></tr>
<tr><td>p</td><td>0.78</td><td>0.67</td><td>0.22</td><td>0.34</td><td>0.04</td><td>0.15</td></tr>
<tr><td rowspan="2">**50°**</td><td>σ</td><td>0.048</td><td>0.049</td><td>0.054</td><td>0.041</td><td>0.037</td><td>0.042</td></tr>
<tr><td>p</td><td>0.92</td><td>0.88</td><td>0.06</td><td>0.01</td><td>0.90</td><td>0.84</td></tr>
<tr><td colspan="3">**Test Result**</td><td>**Passed, except for (70 kHz, 45°)**</td><td>**Passed**</td><td>**Passed**</td><td>**Passed, except for (90 kHz, 50°)**</td><td>**Passed, except for (90 kHz, 45°)**</td><td>**Passed**</td></tr>
</table>

Table 2: Statistical analysis of snippet sample amplitudes ($\sqrt{bs_i}$) of the six homogeneous seafloor patches, for the two external transmission sectors of Kongsberg EM710 MBES (70 kHz and 90 kHz) and at two oblique incidence angles. Parameters σ and p respectively refer to the Rayleigh distribution parameter, and to the p-value of Kolmogorov-Smirnov test.

5 CONCLUSION

Seabed echo samples backscattering strength (snippets) derived from a multibeam echosounder (Kongsberg EM710) measurements acquired on the Atlantic shelf, South west of Brittany, were used to assess the Rayleighness of MBES data. Pre-processing of the data included the computation of the insonified areas and the actual incidence angle on the real seafloor using a high resolution bathymetric model, and the sub-sampling of snippets to retain only independent samples. Distributions of snippets amplitudes were examined for 45° and 50° incidence angles, from six sandy and muddy homogeneous seafloor patches. Over the 24 performed test, 87.5 % concluded to a Rayleigh distribution. In those cases, the backscatter could therefore be estimated accurately[7] using the square sample mean of snippets amplitudes $\sqrt{bs_i}$ i.e. $BS = 10\log_{10}\left(\frac{1}{n}\sum_n bs_i\right)$ with n the number of samples.

Future work will include the assessment of the type of distribution (Rayleigh or other) of snippet that have failed the Rayleigh test, and the investigation of a method to derive the homogeneity of the seafloor inside of the analysed data.

ACKNOWLEDGMENTS

The authors would like to thank the French Service d'Hydrographie et d'Océanograhie de la Marine (SHOM) for collecting the data, with the support of the French General Direction of the Energy and Climate (Direction Générale de l'Énergie et du Climat, DGEC).

REFERENCES

1. D. Jackson and M. Richardson, *High-frequency seafloor acoustics*, Springer Science & Business Media. (2007).
2. M. Malik, X. Lurton, and M. Larry,"A framework to quantify uncertainties of seafloor backscatter from swath mapping echosounders" *Marine Geophysical Research.***39**, 151-168, Springer (2018)
3. X. Lurton and G. Lamarche (Eds), "Backscatter measurements by seafloor-mapping sonars - Guidelines and recommendations", **200** (2015).
4. D. Eleftherakis, L. Berger, N. Le Bouffant, A. Pacault, J-M. Augustin, and X. Lurton, "Backscatter calibration of high-frequency multibeam echosounder using a reference single-beam system, on natural seafloor". *Marine Geophysical Research.***39**, 55-73 (2018).
5. "The secrets behind EM backscatter and backscatter calibrations", *17th Forum for the Exchange of Mutual Multibeam Experiences (FEMME).*, WORKSHOP 6 (2023).
6. T.K. Stanton, W-J. Lee, and K. Baik, "Echo statistics associated with discrete scatterers: A tutorial on physics-based methods", *J. Acoust. Soc. Am.* **144**(6), 3124-3171 (2018).
7. I. Mopin, "Accurate characterisation of seafloor acoustic response to improve seabed identification", PhD. thesis, ENSTA Bretagne, 301 (2023).
8. L. Fonseca, X. Lurton, R. Fezzani, J-M. Augustin, and L. Berger, "A statistical approach for analyzing and modeling multibeam echosounder backscatter, including the influence of high-amplitude scatterers", *J. Acoust. Soc. Am.* **149**(1), 215-228 (2021).
9. R. Fezzani, L. Berger, N. Le Bouffant, L. Fonseca, and X. Lurton, " Multispectral and multiangle measurements of acoustic seabed backscatter acquired with a tilted calibrated echosounder", *J. Acoust. Soc. Am.* **149**(6), 4503-4515 (2021).
10. M. Malik, "Sources and impacts of bottom slope uncertainty on estimation of seafloor backscatter from swath sonars", *Geosciences* **9**(4), 183. (2019).
11. J., D'Errico, "Movingslope" (https://www.mathworks.com/matlabcentral/fileexchange/16997-movingslope), accessed on 2023-12-19, (2023).
12. A. P. Lyons and D. A. Abraham "Statistical characterization of high-frequency shallow-water seafloor backscatter", *J. Acoust. Soc. Am.* **106**(3), 1307-1315 (1999).