

Accurate Mapping of an Extended Shell Bed Area in the North Sea with Multi-spectral Multibeam Backscatter Data

Qian Bai¹, Sebastiaan Mestdagh², Alireza Amiri-Simkooei¹, Thaiënne van Dijk³, Mirjam Snellen¹

¹ Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands
- (q.bai, a.amirisimkooei, m.snellen)@tudelft.nl

² Dept. of Ecosystems and Sediment Dynamics, Deltares, Delft, The Netherlands - sebastiaan.mestdagh@deltares.nl

³ Dept. of Subsurface Systems and Technologies, Deltares, Utrecht, The Netherlands – thaienne.vandijk@deltares.nl

Keywords: Seabed Backscatter, Multibeam Echosounder, Multi-spectral, Shellfish Bed, Seabed Habitat, North Sea.

Abstract

The multibeam echosounder (MBES) has been widely used in seabed mapping, considering its ability to collect continuous and broad-scale seabed measurements efficiently. The presence of shellfish or dead shell material can alter the geophysical properties of the sediment and thus affect the MBES backscatter intensity, making acoustic surveys with the MBES a potential non-invasive solution for regularly monitoring the benthic habitats of shellfish aggregations. Although there exists an increasing interest in mapping marine benthos with MBES measurements recently, the use of multi-spectral backscatter data is still limited. Thus, this research aims to enhance the acoustic mapping of benthic habitats using multi-spectral MBES data, with a focus on a shell bed region in the Dutch North Sea. With backscatter measurements from three frequencies, 90, 300, and 450 kHz, we achieved seabed classification in two steps. First, a semi-supervised backscatter completion was conducted to generate full-coverage backscatter data for each incident angle, mitigating the limited overlap between adjacent survey lines. We then classified the multi-angle backscatter data from each individual frequency using the Gaussian Mixture Model. Our results indicate an improved seabed classification performance compared to the classical Bayesian method. Comparisons of classification maps across frequencies also show their different abilities to distinguish the shell bed region from other coarse sediments, demonstrating the value of leveraging multi-spectral backscatter data in seabed habitat mapping.

1. Introduction

Mapping the occurrence of marine benthos is crucial for preserving seabed habitats and planning offshore human activities (Misiuk and Brown, 2024). Traditional monitoring methods, such as bottom sampling, provide precise point-based seabed measurements, but are costly, low in data density, and potentially destructive. By contrast, the multibeam echosounder (MBES) offers an efficient and non-invasive alternative for continuous and broad-scale seabed mapping. MBES emits acoustic signals in a wide swath perpendicular to the sailing direction and collects the backscattered signals (Parnum and Gavrilov, 2011). The beam steering technique facilitates distinguishing signals backscattered from different directions.

MBES backscatter intensity is affected by seabed geophysical properties, which are not only determined by sediment characteristics such as grain sizes (Hu et al., 2023) but can also be altered by marine benthos (Bai et al., 2023). Benthic species, such as shellfish aggregations, can modify the seabed hardness and roughness due to their hard shell material. It is therefore possible to link the spatial patterns of MBES backscatter data with the presence of marine benthos. Moreover, some shellfish, such as the cut trough shell (*Spisula subtruncata*), can aggregate on a large scale underneath the sediment surface and form a shell bed (de Fouw et al., 2024). Considering the various penetration depths of acoustic signals with different wavelengths, shell bed mapping can thus benefit from the multi-spectral MBES measurements. Additionally, single-frequency backscatter data can show ambiguity for seabed materials with sizes similar to or larger than the acoustic wavelength (Snellen et al., 2018). Multi-spectral analysis might help to resolve this ambiguity.

Previous research has already applied multi-spectral backscatter data to enhance the discrimination of substrate types based on the sediment grain size (Gaida et al., 2018; Runya et al., 2021). Apart from this, it has been indicated that MBES backscatter data with acoustic frequencies lower than 200 kHz might be more sensitive to the presence of sand mason worms (Feldens et al., 2018), showing the potential of multi-frequency surveys for benthic habitat mapping. Menandro et al. (2023) also found that MBES data with frequencies 170, 280, and 400 kHz helped to predict not only the presence but also the coverage proportions of rhodolith beds. However, the applications of multi-spectral MBES backscatter in shellfish bed mapping are still rare.

Moreover, MBES backscatter measurements show both angular and spatial variations in the across-track direction, bringing challenges in data processing and interpretation. Angular normalization during post-processing is a common way to compensate the angular variations and generate spatially continuous backscatter mosaic products for further analysis. However, the information of backscatter angular dependence, which is intrinsic to seabed properties, will be lost during this process (Misiuk and Brown, 2022). On the other hand, typical angular range analysis treats one swath (or half-swath) as a classification unit (Fonseca and Mayer, 2007), which sacrifices the spatial density of MBES data.

In this regard, this research aims to accurately characterize marine benthos by multi-spectral MBES backscatter analysis with an improved use of the angular backscatter data, focusing on an extended shell bed area in the Dutch North Sea. We applied a semi-supervised method, called label propagation, to predict spatially dense backscatter data at a single incident angle and frequency from the sparse input, which helped to reduce noise

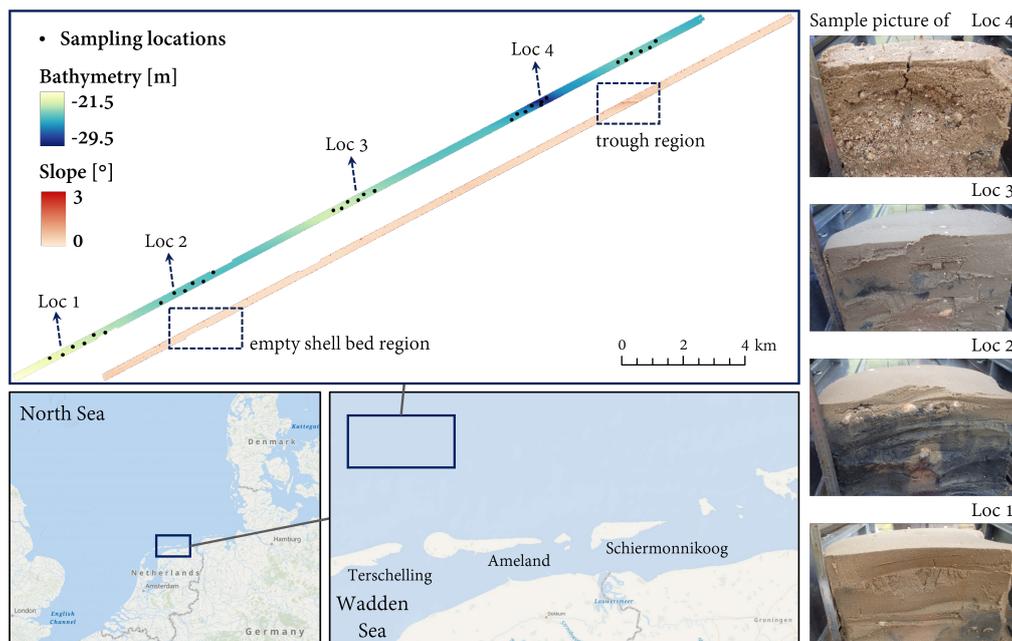


Figure 1. Location of the study area, displayed with bathymetry, slope, sampling locations, and bottom sample pictures for several seabed types.

and ensure spatial consistency in the following acoustic classification. After this backscatter completion step, we classified the achieved dense backscatter data from all incident angles for each frequency using the Gaussian Mixture Model (GMM) clustering. Our results for three acoustic frequencies, 90, 300, and 450 kHz, indicate that only classification for 90 and 450 kHz can distinguish an empty shell bed with buried dead shells from the sandy sea bottom and coarser sediments with larger sediment grain sizes. This demonstrates the benefits of using multi-spectral backscatter to achieve a distinct acoustic signature of marine benthos, which can help large-scale benthic habitat monitoring in the long term.

2. Material and Methods

2.1 Study Area and Acoustic Dataset

We obtained the acoustic dataset in a study area located north of the Wadden Sea islands in the Dutch North Sea (see Figure 1). In recent years, shellfish *S. subtruncata* are known to have formed dense beds in this area. We surveyed the study area in February 2024 with the multi-spectral multibeam system R2Sonic 2026 (R2Sonic, Austin, TX, USA). R2Sonic 2026 is a single-head system and was pole-mounted on the research vessel. Backscatter data for three frequencies, 90, 300, and 450 kHz, were collected, with a beam opening angle of 2.3°, 0.7°, and 0.5°, respectively. A swath coverage of 130° was adopted. To include sufficient seabed variations in the MBES measurements, we surveyed a long transect of about 20 km, which showed a variation in water depths from 21.5 to 29.5 m. As observed from the pictures of the boxcore samples collected at 31 sampling locations, the transect covered the sandy seabed sediment (Loc 1 in Figure 1), a shell bed region (Loc 2), the sandy sediment with a few shell fragments on the surface (Loc 3), and a trough with coarse sediments composed of dead shell accumulations (Loc 4). Sample pictures for the six sampling locations in the northeast of our study area are similar to Loc

3 and are thus not presented here. The trough region showed the largest water depth and seabed slope. Due to the seasonal life cycle of *S. subtruncata*, with strong growth in spring and summer and high mortality in winter, the shell bed region contained only a few living shells. Nevertheless, a layer of empty shells was found just underneath the sediment surface of seabed samples taken in this region.

Analysis of the boxcore samples further indicates the variation of dead shell content in our study area (see Figure 2). The shell bed region contained the highest weight percentage of dead complete shells (%empty shell), while the trough region contained the most broken shell fragments. Gradual changes in %empty shell can also be noticed next to sampling locations at the shell bed region.

2.2 Semi-supervised Backscatter Completion

To eliminate the influence of seawater absorption, sonar settings (such as the source level and receiver gain), and the beam footprint size on raw backscatter measurements, we first conducted backscatter correction according to the sonar equation (Lurton, 2002). This backscatter pre-processing step delivered the beam-averaged backscatter strength (BS) [dB per m^2 at 1 m] for each incident angle and acoustic frequency (Figure 3).

The limited overlap between adjacent survey lines due to restrictions in survey time and budget can bring across-track spatial gaps in backscatter data from a single incident angle. To preserve both the angular and spatial density of backscatter data in further analysis, such as classification, we conducted a backscatter completion step by applying the semi-supervised graph-based method, called label propagation (Zhu and Ghahramani, 2002). Label propagation was proposed to propagate label information from a small amount of labeled data to unlabeled data points in a dataset. The algorithm achieves this propagation by treating all data points as nodes in a graph and leveraging the connectivity among them, which is determined by the pairwise similarity between data points.

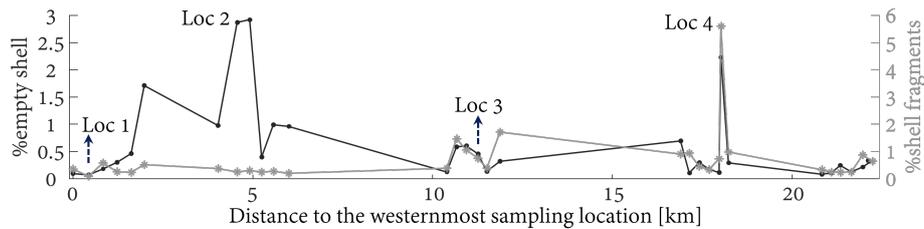


Figure 2. Dead shell content of all boxcore samples. Empty shells and shell fragments were analyzed separately.

Given a single-frequency MBES dataset, we constructed a dense graph with (x, y, z, s) of each point, in which x and y represent the easting and northing coordinates, respectively. z is bathymetry and s indicates seabed slope. In contrast to *BS*, bathymetry data can be regarded as continuous measurements across incident angles. By defining the similarity between data points based on the Euclidean distance calculated on (x, y, z, s) , the graph involves the local smoothness in seabed topography. Among these dense nodes, only a part of them has $BS(\theta)$, with θ a single incident angle. Afterward, instead of using the label probability matrix as the prediction target, like in the original label propagation algorithm, we propagated $BS(\theta)$ from nodes measured with θ to nodes with other incident angles. This backscatter completion was repeated for every incident angle, providing full-coverage multi-spectral multi-angle backscatter maps for each frequency.

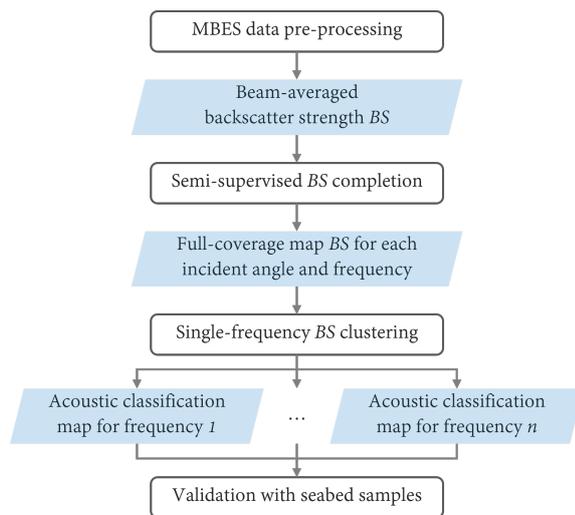


Figure 3. Our workflow for acoustic classification of multi-spectral MBES backscatter data.

2.3 Single-frequency Acoustic Classification

With backscatter data from various incident angles, we conducted acoustic classification for each frequency to investigate their differences in discriminating seabed types (Figure 3). To reduce redundant information in *BS*, we conducted a principal component analysis (PCA) on *BS* from all angles and extracted the first several principal components (PCs) that can explain $\geq 90\%$ of the data variance. The PCs were then classified based on a Gaussian Mixture Model (GMM). GMM is a widely used clustering algorithm, which assumes that the input features can be represented by a combination of k multivariate Gaussian distributions. By finding the mean, variance, and mixing weight parameter of each distribution through the Expect-

ation–Maximization optimization, the input features can be divided into k classes. GMM is flexible in capturing classes with different variances and allows to model overlapping classes. For our seabed mapping application, this unsupervised classification method also helps to exploit the data structure given a limited number of seabed ground truth samples.

We compared the acoustic classification maps, achieved by the GMM classification on completed dense *BS* data (GMM-Dense), with another well-developed unsupervised seabed classification method, Bayesian classification (Simons and Snellen, 2009). The Bayesian method assumes that *BS* for a specific seabed type follows a Gaussian distribution based on the central limit theorem, given a fixed incident angle and frequency. A summation of k Gaussians can then be fitted to the histogram of the single-angle single-frequency *BS* within the study area. The method involves a χ^2 -statistical test to evaluate the fitting performance, which also helps to search for the optimal k . However, since Bayesian classification works angle by angle in principle, a strategy of combining results from different angles is needed. As described by Gaida et al. (2018), several reference incident angles between 40° and 65° can be selected based on the χ^2 -test results and used to define the class percentage rule, which will be extended to other angles larger than 20° . *BS* from the nadir beams are excluded due to the possible violation of the central limit theorem. We followed this approach to implement Bayesian classification in this research. Moreover, starting from the optimal k indicated by Bayesian classification, we incrementally increased the number of classes in GMM-Dense classification until no additional spatial patterns emerged.

We validated our GMM-Dense classification based on the box-core sample analysis results. A qualitative comparison was conducted between the seabed types indicated by the sample pictures (Loc 1–4 in Figure 1) and the acoustic classes at these sampling locations regarding each frequency. In addition, we investigated if the acoustic classification can account for the variations in dead shell properties (%empty shell and %shell fragments) within our study area.

3. Results

3.1 Acoustic Classification

After the semi-supervised backscatter completion, we achieved dense *BS* maps from -65° to 65° . We discarded *BS* for incident angles between $\pm 20^\circ$, considering the artifacts in the completed maps. *BS* data in the nadir can present a larger uncertainty than outer beams due to a smaller beam footprint size and thus deteriorate the completion performance. This procedure also ensured a fair comparison with Bayesian classification. Through PCA, we selected the first PC for 90 and 300 kHz, and the first

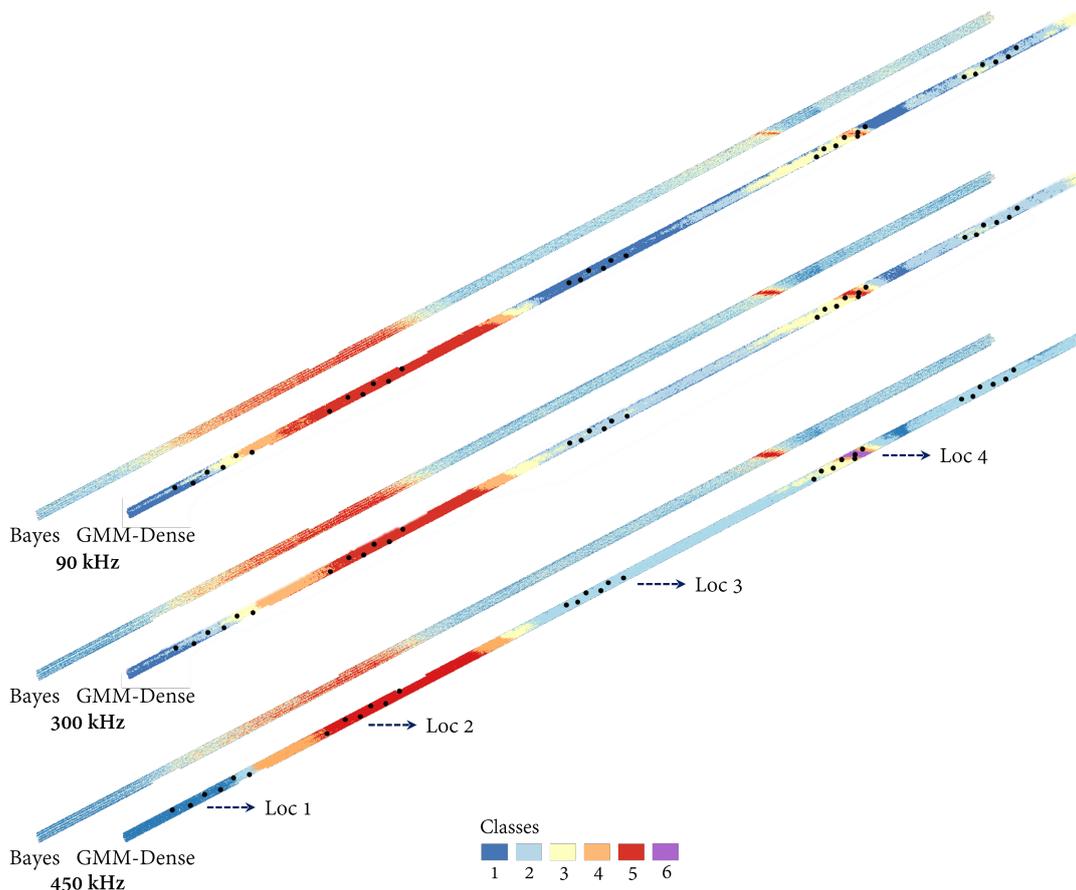


Figure 4. Classification maps achieved by Bayesian classification (Bayes) and GMM classification on our completed dense *BS* (GMM-Dense) for 90, 300, and 450 kHz.

two PCs for 450 kHz, preserving over 90% of data variance in the reduced feature space for each frequency.

Classification maps achieved by the GMM-Dense classification show better spatial consistency than the Bayesian method (see Figure 4). Due to the generalization of the classification rule from outer beams to other angles, there exists across-track inconsistency in the Bayesian classification (Bayes) maps to a certain extent. For example, a mixed pattern of class 1 and 2 can be noticed east of the trough region on the Bayes map for all three frequencies, while the GMM-Dense results reveal a clear spatial pattern for class 1 in this region. Besides, excluding nadir beams in classification results in gaps in the Bayes maps, given a limited overlap between survey lines. By contrast, spatially continuous classification can be acquired by the GMM-Dense method based on the backscatter completion step.

In addition, by leveraging the angular variations from the multi-angle *BS* data, GMM-Dense classification also reveals an additional seabed type for 450 kHz in the trough region compared to the Bayesian method. Seabed near Loc 2 and Loc 4 can therefore be distinguished by the GMM-Dense classification. For 90 kHz, although the number of GMM-Dense classes is the same as Bayesian classification, a clearer separation between class 4 and 5 in the trough is achieved.

3.2 Interpretation with Seabed Samples

We summarized the *BS* angular response curve (ARC) near sampling locations Loc 1–4 for all three frequencies (see Fig-

ure 5). ARCs for Loc 1 and 3 are similar regarding 90 and 300 kHz, while they present about 1 dB difference for 450 kHz considering $\theta \geq 40^\circ$. Boxcore pictures indicate that Loc 3 contained slightly more dead shell fragments on the sediment surface than Loc 1, which might cause stronger scattering for the highest acoustic frequency. GMM-Dense results also show that only classification for 450 kHz distinguishes the two seabed types.

Regarding Loc 2 and 4, their ARCs show much higher *BS* than the other sampling locations for all three frequencies. Situated in the deepest trough, Loc 4 contained an accumulation of dead shell material, representing the coarsest sediment in this study area. This brings the highest *BS* in the nadir for all frequencies. As observed from the sample pictures, although Loc 2 was covered by sand, there was a layer of empty shells just underneath the seabed surface. Considering a better seabed penetration ability of lower frequencies, acoustic signals from 90 kHz are more likely to sense this buried shell layer. However, *BS* from all three frequencies show higher values for the empty shell bed region than the surrounding seabed, indicating limited differences in their penetration depths. This can be explained by the shallow location of the empty shell layer. Nevertheless, differences between responses to Loc 2 and 4 can still be noticed for all frequencies.

In the middle angular range (20° – 40° for both port and starboard sides), *BS* of Loc 4 presents the largest value for 450 kHz. However, it is comparable to Loc 2 for 300 kHz, and lower than

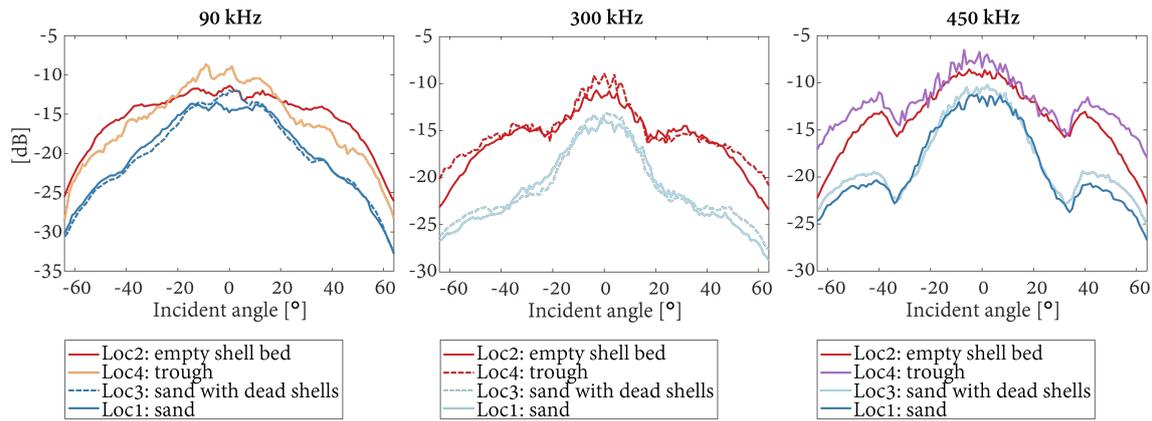


Figure 5. *BS* angular response curves of different seabed types for (Left) 90, (Middle) 300, and (Right) 450 kHz, with the line color consistent with the single-frequency GMM-Dense classification results in Figure 4.

BS from Loc 2 for 90 kHz. Moreover, *BS* with $\theta \geq 40^\circ$ of the trough is higher than the empty shell bed for 300 kHz, and is significantly higher for 450 kHz. The two seabed types, however, are less different in this angular range regarding 90 kHz. In general, ARCs of the trough and the empty shell layer show distinct shapes for 90 kHz, possibly due to the different sizes of the two types of coarse material, making them distinguishable in the GMM-Dense classes. By contrast, the trough material can have larger surface roughness, inducing much stronger scattering than the shell bed region in outer beams for 450 kHz. With a smaller wavelength, acoustic signals from 450 kHz can be more sensitive to the change in surface roughness. *BS* from the middle frequency, 300 kHz, on the other hand, presents limited difference between the two seabed types.

ability in separating all four seabed types indicated by the sample pictures, while *BS* for 90 kHz also presents intrinsic differences between two different coarse seabed types.

With acoustic classification results near all ground truth sampling locations, relationships between the GMM-Dense classes and median values of the dead shell content for each class can also be investigated (see Figure 6). The GMM-Dense classification presents a generally positive correlation with %empty shell, especially considering classes 3–5 for 90 kHz and classes 4–6 for 450 kHz. Moreover, classes 1–4 for 450 kHz are consistent with an increase of %shell fragments. Classification for 300 kHz, on the other hand, shows limited correlation with %shell fragments.

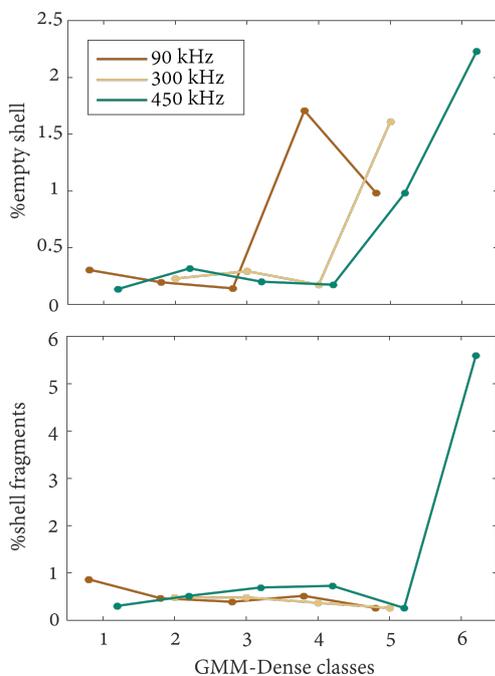


Figure 6. Relationship between GMM-Dense classes and two dead shell properties, including (Top) %empty shell and (Bottom) %shell fragments.

In summary, *BS* for 450 kHz with $\theta \geq 40^\circ$ shows the largest

4. Conclusion

For long-term benthic habitat monitoring, it is essential to employ a cost-effective approach with limited disturbance to the seabed. With the focus on an extended shell bed area in the Dutch North Sea, this research investigated an acoustic seabed classification workflow using a multi-spectral multibeam echosounder (MBES). Although seabed backscatter data from the MBES can correlate with sediment properties, they typically exhibit across-track spatial discontinuity due to the backscatter angular variation and limited overlap between survey lines. To preserve both the angular and spatial data density in acoustic classification, we applied semi-supervised label propagation to achieve a full-coverage backscatter map for each incident angle.

Leveraging the spatially dense multi-angle backscatter data, we achieved single-frequency classification maps with the Gaussian Mixture Model clustering algorithm. Comparisons among acoustic frequencies reveal their differences in distinguishing seabed types, especially the empty shell layer just underneath the sandy sediment surface and the coarse seabed composed of an accumulation of dead shells. Backscatter data of outer beams for 450 kHz were found to have the largest ability to separate the empty shell bed from other sediments, compared to 90 and 300 kHz. Although very few living shells were found in our study area due to our survey taking place in winter, our results demonstrate the value of using multi-spectral backscatter measurements for mapping the empty shell bed properties. Cataloguing these frequency-dependent acoustic signatures can also be beneficial for future seabed habitat monitoring and proper planning of offshore activities.

Acknowledgement

The authors would like to thank the Ministry of Infrastructure and Water Management of the Netherlands (Rijkswaterstaat) for their support in survey planning and measurements onboard. Special thanks to Waardenburg Ecology and Eurofins for their help in sample collection and analyses. The company QPS is thanked for licenses of their software for hydrographic data collection and processing. R2Sonic and Van Oord are also acknowledged for providing sonar equipment. Additionally, Boskalis is thanked for their technical support. Finally, the authors express their gratitude to Wageningen Marine Research for their valuable advice.

References

- Bai, Q., Mestdagh, S., Snellen, M., Simons, D. G., 2023. Indications of marine benthos occurrence from multi-spectral multi-beam backscatter data: a case study in the North Sea. *Frontiers in Earth Science*, 11, 1140649.
- de Fouw, J., van Horssen, P. W., Craeymeersch, J., Leopold, M. F., Perdon, J., Troost, K., Tulp, I., van Zwol, J., Philippart, C. J., 2024. Spatio-temporal analysis of potential factors explaining fluctuations in population size of *Spisula subtruncata* in the Dutch North Sea. *Frontiers in Marine Science*, 11, 1476223.
- Feldens, P., Schulze, I., Papenmeier, S., Schönke, M., Schneider von Deimling, J., 2018. Improved interpretation of marine sedimentary environments using multi-frequency multi-beam backscatter data. *Geosciences*, 8(6), 214.
- Fonseca, L., Mayer, L., 2007. Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. *Marine Geophysical Researches*, 28, 119–126.
- Gaida, T. C., Tengku Ali, T. A., Snellen, M., Amiri-Simkooei, A., Van Dijk, T. A., Simons, D. G., 2018. A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data. *Geosciences*, 8(12), 455.
- Hu, H., Feng, C., Cui, X., Zhang, K., Bu, X., Yang, F., 2023. A sample enhancement method based on simple linear iterative clustering superpixel segmentation applied to multibeam seabed classification. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1–15.
- Lurton, X., 2002. *An introduction to underwater acoustics: principles and applications*. Springer Science & Business Media.
- Menandro, P. S., Misiuk, B., Brown, C. J., Bastos, A. C., 2023. Multispectral multibeam backscatter response of heterogeneous rhodolith beds. *Scientific Reports*, 13(1), 20220.
- Misiuk, B., Brown, C. J., 2022. Multiple imputation of multibeam angular response data for high resolution full coverage seabed mapping. *Marine Geophysical Research*, 43(1), 7.
- Misiuk, B., Brown, C. J., 2024. Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor. *Estuarine, Coastal and Shelf Science*, 296, 108599.
- Parnum, I. M., Gavrilov, A. N., 2011. High-frequency multi-beam echo-sounder measurements of seafloor backscatter in shallow water: Part 1–Data acquisition and processing. *Underwater Technology*, 30(1), 3–12.
- Runya, R. M., McGonigle, C., Quinn, R., Howe, J., Collier, J., Fox, C., Dooley, J., O’loughlin, R., Calvert, J., Scott, L. et al., 2021. Examining the links between multi-frequency multi-beam backscatter data and sediment grain size. *Remote Sensing*, 13(8), 1539.
- Simons, D. G., Snellen, M., 2009. A Bayesian approach to seafloor classification using multi-beam echo-sounder backscatter data. *Applied Acoustics*, 70(10), 1258–1268.
- Snellen, M., Gaida, T. C., Koop, L., Alevizos, E., Simons, D. G., 2018. Performance of multibeam echosounder backscatter-based classification for monitoring sediment distributions using multitemporal large-scale ocean data sets. *IEEE journal of oceanic engineering*, 44(1), 142–155.
- Zhu, X., Ghahramani, Z., 2002. Learning from labeled and unlabeled data with label propagation. Technical report, Carnegie Mellon University.