Multi-frequency Acoustic Mapping of Marine Benthos

Data-driven Multibeam Classification in the Dutch North Sea

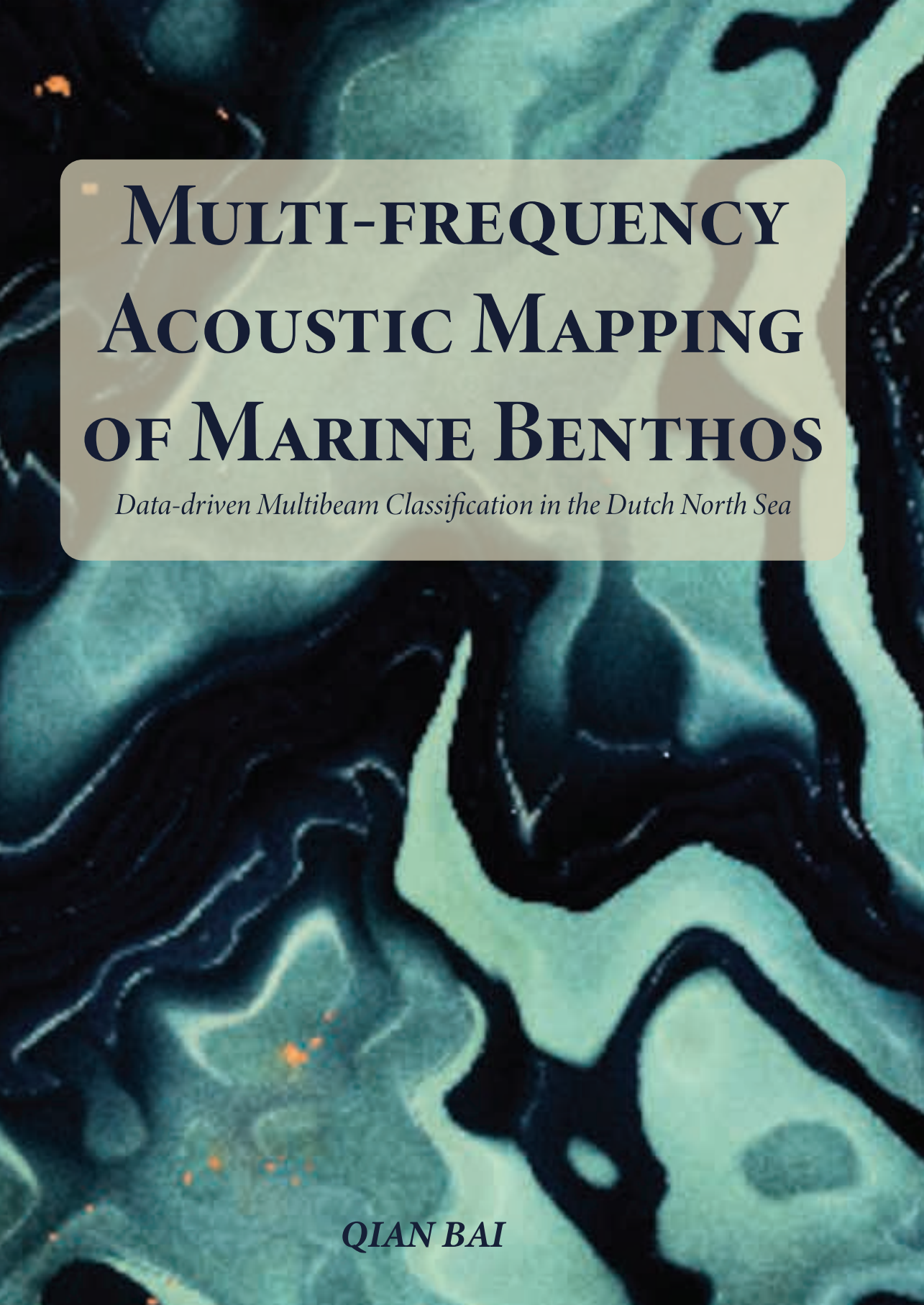
Bai, Q.

DOI
10.4233/uuid:da455477-65dc-44c5-8657-32397549189c

Publication date
2026

Document Version
Final published version

Citation (APA)
Bai, Q. (2026). *Multi-frequency Acoustic Mapping of Marine Benthos: Data-driven Multibeam Classification in the Dutch North Sea*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:da455477-65dc-44c5-8657-32397549189c



Multi-frequency Acoustic Mapping of Marine Benthos

Data-driven Multibeam Classification in the Dutch North Sea

QIAN BAI

Multi-frequency Acoustic Mapping of Marine Benthos

Data-driven Multibeam Classification in the Dutch North Sea

Multi-frequency Acoustic Mapping of Marine Benthos

Data-driven Multibeam Classification in the Dutch North Sea

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates,
to be defended publicly on
Monday, 2 February, 2026 at 17:30 o'clock

by

Qian BAI

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairman
Em. Prof. dr. D.G. Simons,	Delft University of Technology, *promotor*
Prof. dr. ir. M. Snellen,	Delft University of Technology, *promotor*
Independent members:	
Prof. dr. L.L.A. Vermeersen,	Delft University of Technology
Prof. dr. L.G. Evers,	Delft University of Technology
Prof. dr. J. Greinert,	GEOMAR Helmholtz Centre for Ocean Research Kiel
Dr. R.C. Lindenbergh,	Delft University of Technology
Dr. R. Merino-Martínez,	Delft University of Technology, *reserve member*
Other member:	
Dr. A. Amiri-Simkooei,	Delft University of Technology

Keywords: Multibeam echosounder, benthic habitat, marine benthos, seabed mapping, multi-frequency, underwater acoustics, artificial intelligence
Printed by: IPSKAMP Printing
Cover design: Anna Glass

ISBN 978-94-6518-219-3

An electronic copy of this dissertation is available at
`https://repository.tudelft.nl/`.

Contents

Summary

With increasing offshore human activities and accelerating climate change, regular seabed habitat monitoring is essential for marine conservation and sustainable coastal development. Compared to destructive bottom sampling that is labor-intensive and optical remote sensing with limited penetration in seawater, the multibeam echosounder (MBES) provides a cost-effective solution for high-resolution, large-scale seabed mapping by simultaneously acquiring bathymetry and acoustic backscatter. In recent years, multi-spectral MBES has become a state-of-the-art acoustic mapping technique, providing nearly co-located multi-frequency measurements and largely enriching seabed characterization. Despite difficulties in obtaining calibrated backscatter, data-driven methods, especially machine learning techniques, still allow for linking MBES measurements to seabed geophysical and biological properties.

Nevertheless, MBES-based benthic habitat mapping remains challenging. Limited seabed ground truth hinders model construction and evaluation. Lack of absolute calibration poses challenges when comparing or combining MBES backscatter across surveys. Backscatter angular dependency and large volume of multi-frequency measurements further complicate data processing. This thesis addresses these challenges by exploiting the multi-spectral MBES, making optimal use of limited ground truth, and improving the MBES data processing workflow.

This thesis focuses on MBES measurements in the Dutch North Sea and habitats of marine benthos, which are organisms living on or in the seabed sediment. In the first study, two surveys north of the Wadden Sea islands provided multi-spectral MBES data (90–450 kHz) for sand mason worm habitats. Two single-frequency unsupervised classification methods, Bayesian classification leveraging beam-averaged backscatter of individual incident angles, and hierarchical clustering on mean backscatter values from three angular ranges, were applied and compared. Both methods achieved a good alignment between acoustic classes and sediment median grain size, while hierarchical clustering better distinguished areas with high gravel content or abundant sand mason worms, highlighting the value of backscatter angular information. Moreover, results across frequencies were highly similar, suggesting little distinction among 90–450 kHz for the sandy seabed in both study areas.

The next study revisits the same areas after almost three years with the same sonar configuration, allowing for MBES-based seabed change detection. Compared to previous surveys conducted in summer, two new surveys were carried out in early spring, introducing a seawater temperature difference of nearly 10°C. Backscatter strengths (after backscatter correction) within a stable reference area differed between the two survey times, potentially due to the temperature impact on sonar hardware. This misalignment was utilized for a relative backscatter calibration, which enabled calculation of backscatter differences between the survey times. A Gaussian Mixture Model, constrained by

the backscatter uncertainty estimated from the reference area, then clustered seabed change types based on these differences. With MBES backscatter at 300 kHz, the change detection results indicated possible sediment movement and a potential reduction in sand mason worm density, supported by bathymetry data and seabed ground truth.

Beyond unsupervised approaches, a semi-supervised seabed classification method was developed for a blue mussel cultivation area in the Oosterschelde, the Netherlands. With nine bottom samples, only a few nearby soundings can be labeled directly. Nevertheless, relations between labeled and unlabeled soundings were exploited to create pseudo-labeled data via the label spreading algorithm on bathymetry derivatives and backscatter features at 300 kHz, enlarging the training dataset and enabling robust training of supervised classification. Mussels were successfully distinguished from the bare sediment, as validated by the documentation for mussel cultures.

Another study conducted in the cut trough shell habitat north of the Ameland and Terschelling islands shows the effectiveness of semi-supervised graph propagation for predicting backscatter values at unknown locations in a single-frequency single-angle dataset. These missing values typically arise from limited overlap between MBES survey lines and angular variations in backscatter. This completion step leveraged connectivity in bathymetric features among soundings and generated full-coverage backscatter for each incident angle and frequency (90, 300, 450 kHz). Following this, a spectral neural network conducted feature reduction on the multi-frequency multi-angle backscatter features, leading to better clustering results than conventional linear reduction and manifold learning methods. Moreover, sandy seabed, coarse sediments composed of shell fragments, and the living and empty shell beds were successfully discriminated. Although living and empty shell beds present bathymetric similarity, they show distinct multi-frequency backscatter signatures.

Together, these four studies demonstrate the effectiveness of semi-supervised and unsupervised learning for benthic habitat mapping with limited ground truth samples. Relative calibration and single-angle backscatter completion were integrated into the MBES data processing workflow, offering a way to utilize backscatter data from multiple surveys while preserving both spatial and angular density. The multi-spectral MBES also proved effective for mapping shellfish beds in the Dutch North Sea. Finally, this thesis summarizes insights from these studies and presents guidelines for future MBES surveys in marine benthos mapping, regarding the sonar system selection, relative calibration, and seabed ground truth data collection.

Samenvatting

Door toenemende menselijke activiteiten op zee en versnellende klimaatverandering is regelmatige monitoring van habitats op de zeebodem essentieel voor mariene natuurbescherming en duurzame kustontwikkeling. Vergeleken met destructieve en arbeidsintensieve bodembemonstering en optische teledetectie met beperkte penetratie in zeewater, biedt de multibeam echosounder (MBES) een kosteneffectieve oplossing voor grootschalige zeebodemkartering met hoge resolutie, door gelijktijdig bathymetrie en akoestische backscatter te meten. In de afgelopen jaren is multispectrale MBES uitgegroeid tot een state-of-the-art techniek voor akoestische kartering, die vrijwel cogelokaliseerde metingen op meerdere frequenties levert en de karakterisering van de zeebodem sterk verrijkt. Ondanks de moeilijkheden bij het verkrijgen van gekalibreerde backscatter maken datagedreven methoden, met name machinelearningtechnieken, het nog steeds mogelijk om MBES-metingen te koppelen aan geofysische en biologische eigenschappen van de zeebodem.

Toch blijft benthische habitatkartering op basis van MBES uitdagend. Beperkte ground truth van de zeebodem belemmert modelontwikkeling en evaluatie. Het ontbreken van absolute kalibratie vormt een uitdaging bij het vergelijken of combineren van MBES-backscatter tussen surveys. De hoekafhankelijkheid van backscatter en de grote hoeveelheid multifrequente metingen maken de gegevensverwerking nog complexer. Dit proefschrift pakt deze uitdagingen aan door multispectrale MBES te benutten, optimaal gebruik te maken van beperkte ground truth, en de workflow voor MBES-dataverwerking te verbeteren.

Dit proefschrift richt zich op MBES-metingen in het Nederlandse deel van de Noordzee en op habitats van marien benthos, organismen die op of in het zeebodemsediment leven. In de eerste studie leverden twee meetcampagnes ten noorden van de Waddeneilanden multispectrale MBES-data (90–450 kHz) voor habitats van de schelpkokerworm. Twee unsupervised classificatiemethoden op basis van enkelvoudige frequentie, Bayesiaanse classificatie op basis van beam-gemiddelde backscatter bij individuele invalshoeken, en hiërarchische clustering op gemiddelde backscatterwaarden uit drie hoekbereiken, werden toegepast en vergeleken. Beide methoden toonden een goede overeenstemming tussen akoestische klassen en de mediane korrelgrootte van het sediment, terwijl hiërarchische clustering beter gebieden onderscheidde met een hoog grindgehalte of een hoge dichtheid aan schelpkokerwormen, wat de waarde van hoekinformatie in backscatter benadrukt. Bovendien waren de resultaten over de verschillende frequenties heen zeer vergelijkbaar, wat erop wijst dat er voor de zandige zeebodem in beide studiegebieden weinig onderscheid is tussen 90 en 450 kHz.

De volgende studie bezoekt dezelfde gebieden na bijna drie jaar opnieuw met dezelfde sonarconfiguratie, wat MBES-gebaseerde detectie van veranderingen op de zeebodem mogelijk maakt. In vergelijking met eerdere surveys in de zomer werden twee nieuwe sur-

veys uitgevoerd in het vroege voorjaar, met een zeewatertemperatuurverschil van bijna 10°C. Backscattersterkte (na backscattercorrectie) binnen een stabiel referentiegebied verschilde tussen de twee meetmomenten, mogelijk door de invloed van temperatuur op de sonarhardware. Deze uitlijningsfout werd gebruikt voor een relatieve backscatter-kalibratie, waarmee backscatterverschillen tussen de meetmomenten konden worden berekend. Vervolgens clusterde een Gaussian Mixture Model, begrensd door de backscatteronzekerheid geschat uit het referentiegebied, typen zeebodemverandering op basis van deze verschillen. Met MBES-backscatter bij 300 kHz wezen de resultaten van de veranderingdetectie op mogelijke sedimentverplaatsing en een mogelijke afname in de dichtheid van schelpkokerwormen, ondersteund door bathymetrische gegevens en ground truth van de zeebodem.

Naast unsupervised benaderingen werd een semi-supervised classificatiemethode voor de zeebodem ontwikkeld voor een gebied met mosselkweek in de Oosterschelde. Met negen bodemsamples kunnen slechts enkele nabijgelegen soundings direct gelabeld worden. Toch werden relaties tussen gelabelde en ongelabelde soundings benut om pseudo-gelabelde data te creëren via het label spreading-algoritme op bathymetrische afgeleiden en backscatterkenmerken bij 300 kHz. Dit vergrootte de trainingsdataset en maakte robuuste training van supervised classificatie mogelijk. Mosselen werden succesvol onderscheiden van het kale sediment, zoals gevalideerd door de documentatie voor mosselculturen.

Een andere studie, uitgevoerd op een schelpenbank van halfgeknotte strandschelpen ten noorden van de eilanden Ameland en Terschelling, toont de effectiviteit van semi-supervised graafpropagatie voor het voorspellen van backscatterwaarden op onbekende locaties in een dataset met enkelvoudige frequentie en hoek. Deze ontbrekende waarden ontstaan doorgaans door beperkte overlap tussen MBES-vaarroutes en door hoekvariaties in backscatter. Deze stap voor voltooiing maakte gebruik van de connectiviteit in bathymetrische kenmerken tussen soundings en genereerde backscatter met volledige dekking voor elke invalshoek en frequentie (90, 300, 450 kHz). Daarna voerde een spectraal neuraal netwerk feature-reductie uit op de multifrequente, multihoek backscatterfeatures, wat leidde tot betere clusteringresultaten dan conventionele lineaire reductie- en manifold learning-methoden. Bovendien werden zandige zeebodem, grovere sedimenten bestaande uit schelpfragmenten, en levende en lege schelpenbedden succesvol onderscheiden. Hoewel levende en lege schelpenbedden bathymetrisch sterk op elkaar lijken, vertonen ze onderscheidbare multifrequente backscatter-signaturen.

Samen tonen deze vier studies de effectiviteit aan van semi-supervised en unsupervised learning voor benthische habitatkartering met beperkte ground truth-monsters. Relatieve kalibratie en voltooiing van backscatter onder enkele hoek werden geïntegreerd in de workflow voor MBES-dataverwerking, wat een manier biedt om backscatterdata uit meerdere surveys te benutten terwijl zowel ruimtelijke dichtheid als hoekdichtheid behouden blijven. Multispectrale MBES bleek ook effectief voor het in kaart brengen van schelpdierbanken in het Nederlandse deel van de Noordzee. Tot slot vat dit proefschrift inzichten uit deze studies samen en presenteert het richtlijnen voor toekomstige MBES-surveys voor mariene benthoskartering, met betrekking tot de keuze van het sonarsysteem, relatieve kalibratie en de verzameling van ground truth-gegevens van de zeebodem.

Nomenclature

Latin symbols

symbol	description
a	Receiver elements on a line array
A	Ensonified footprint size
BS	Backscatter strength
$\overline{BS}$	Mean backscatter strength of one Gaussian component
BS'	Backscatter strength after relative calibration
c	Sound speed
	The largest index of data labels
C	Relative backscatter calibration term
d	Distance from array elements to the acoustic wavefront
	Water depth
	Length of a point vector
d_{50}	Median grain size
D	Directivity pattern
	Degree matrix of a graph
E	Graph edge
EL	Received echo level of an MBES
f	Acoustic frequency
F	Classification rule in label spreading
G	Receiver gain
	Graph
h	Histogram of backscatter measurements
k	Number of folds in cross-validation
	Number of neighbors to construct the k-NN graph
l	Number of labeled points
L	Array length
$\mathscr{L}$	Data labels
	Loss function for a neural network
m	Number of Gaussian components
	Holdout percentage
M	Number of receiver elements on a line array
	Number of histogram bins in Bayesian classification
n	Number of backscatter measurements per histogram bin
	Number of points in a dataset
$N\,(N_\phi, N_f)$	Number of incident angles and acoustic frequencies

symbol	description
p	Point vector in a point dataset
$\mathbf{p}$	Positional encoding
P	Overall probability of one measurement in a dataset
$\mathscr{P}$	A point dataset
Q	Cost function of label spreading
R	One-way travel distance
S	Normalized similarity matrix
SH	Transducer sensitivity
SL	Source level
t	Two-way travel time
	Survey time
	Mini-batch size
T	Temperature
TL	Transmission loss
v	Degrees of freedom
	Prediction target of the semi-supervised backscatter completion
V	Graph vertex
w	Weighting factor of Gaussian components
W	Similarity matrix
x	Data point in the backscatter difference
	Easting coordinate
X	Data matrix
y	Northing coordinate
	Reduced feature vector
Y	Predicted label matrix in label spreading
z	Bathymetry

Greek symbols

symbol	description
α	Seawater absorption coefficient
δ	Spacing between receiver elements on a line array
ΔBS	Difference in backscatter strength between two surveys
ϵ ($\epsilon_{al}, \epsilon_{ac}$)	Seabed slope in the along- and across-track direction
λ	Wavelength
μ	Lambert parameter
	Mean backscatter strength of one Gaussian component
	Average backscatter strength of a reference area
	Regularization parameter in label spreading
μ_Δ	Average backscatter difference in Gaussian Mixture Model
ϕ	Incident angle relative to seabed slopes
ϕ_{fl}	Incident angle under a flat-seabed assumption
ρ	Density

symbol	description
σ	Standard deviation of one Gaussian component
	Backscatter standard deviation of a reference area
	Scaling parameter in the similarity matrix
σ_Δ	Standard deviation of backscatter difference in Gaussian Mixture Model
τ	Time delay of receiver elements
	Pulse length
θ	Steering angle
Θ	Beam angle
ω	Circular frequency
$\Omega\,(\Omega_t, \Omega_r)$	Beam width (Beam opening angle) of transmission and reception
χ^2	Chi-square goodness of fit criterion
χ^2_v	Reduced chi-square goodness of fit criterion

Acronyms

General acronyms

ARC	Angular Response Curve
AUV	Autonomous Underwater Vehicle
DEM	Digital Elevation Model
DL	Deep Learning
GB	Gradient Boosting
GMM	Gaussian Mixture Model
GNSS	Global Navigation Satellite System
GSF	Generic Sensor Format
IHO	International Hydrographic Organization
k-NN	k-Nearest Neighbors
LiDAR	Light Detection and Ranging
LOSO	Leave-One-Station-Out
MAE	Mean Absolute Error
MBES	Multibeam Echosounder
ML	Machine Learning
MLP	Multi-Layer Perceptron
MPA	Marine Protected Area
PC	Principal Component
PCA	Principal Component Analysis
RF	Random Forest
SAS	Synthetic Aperture Sonar
SBES	Single-Beam Echosounder
SBP	Sub-Bottom Profiler
SSS	Side-Scan Sonar
SVM	Support Vector Machine
USBL	Ultra-Short BaseLine
VIF	Variance Inflation Factor

Folk classes

fS	Fine Sand
mG	Muddy Gravel
mS	Medium Sand
gmS	Gravelly Muddy Sand

cS	Coarse Sand
msG	Muddy Sandy Gravel

1

Introduction

The sea becomes great by embracing all rivers.

Lin Zexu (1785—1850)

1.1. Background

As the ocean's foundation, the seabed fosters biodiversity, supplies vital natural resources, supports global communication networks, and provides opportunities for exploration and recreation. Life on and in the seabed, known as marine benthos, is crucial to the health of marine ecosystems. They are the food source for fish, birds, and marine mammals. Some of them act as ecosystem engineers, modifying the sediment physical properties with their behaviors and thus turning their habitats into biodiversity hotspots.

However, accelerating climate change and the intensification of offshore activities are placing stress on seabed habitats. Bottom trawling, dredging, sand nourishment, and other extractive practices physically disrupt benthic communities [1, 2], while the rapid expansion of offshore wind farms can alter the foraging and dispersal patterns of sensitive species [3, 4]. Regular monitoring of seabed habitats is therefore essential for marine conservation and sustainable coastal development.

In addition to these ecological concerns, ocean conservation policies create a regulatory need for detailed seabed mapping. Worldwide, Marine Protected Areas (MPAs) are a fundamental measure for managing human activities and preserving marine biodiversity [5, 6]. In the North Sea, engineering practices such as sand extraction have been prohibited within a distance of 100 meters from living shellfish beds since 2008 [7]. Implementing MPAs and meeting such regional obligations for ecologically significant areas require effective methods for classifying and delineating benthic habitats.

Traditional monitoring of the seabed relies on bottom sampling techniques such as boxcores, grabs, and dredges. Each seabed sample usually has a limited surface coverage, ranging from a few square decimeters to tens of square meters. Laboratory analysis of those samples can reveal detailed information on sediment compositions and benthic communities. Besides, underwater imagery and videos are commonly collected in seabed surveys during the past decades. Compared to extractive sampling which is potentially destructive to the benthic habitat, drop camera systems can hover above the seabed and record transects with a broader view [8]. Nevertheless, they are best suited for observing hard substrates, since the image quality is largely affected by water turbidity. The acquired seabed information is also limited to the sediment surface. In general, both approaches are labor-intensive and time-consuming, resulting in only sparsely distributed measurements. To achieve spatially continuous maps, assumptions about areas between sampling locations are needed for interpolation [9], making these techniques inadequate for high-resolution habitat mapping on a broad scale.

Remote sensing technologies extend seabed observation from discrete samples to continuous measurements. Satellite-derived and airborne LiDAR bathymetry enable large-scale analysis of seabed geomorphology and the construction of ocean floor Digital Elevation Models (DEMs) comparable to their terrestrial counterparts [10]. Reflectance data from imagery and laser scanning point clouds have also proved effective for mapping sediment types, aquatic vegetation, and coral reefs [11–13]. However, the rapid attenuation of electromagnetic waves in aquatic environments limits their applications to clear, shallow coastal waters and intertidal flats [14]. Sound waves, by contrast, can travel much further in water, making acoustic remote sensing the primary tool for mapping deeper waters or turbid environments such as the Dutch North Sea.

1.1.1. Acoustic interaction with the seabed

When the distance sound travels is sufficiently long compared to its wavelength, sound can be assumed as plane waves [15]. Under this assumption, an acoustic wave travels through seawater and meets the water-seabed boundary at one uniform incident angle. The interaction between sound and seabed material can then be mainly described by three physical processes: reflection, transmission, and scattering. The contribution of each process is affected by the seabed irregularity and the geoacoustic contrast between seawater and sediment, and is further controlled by the incident angle and acoustic frequency.

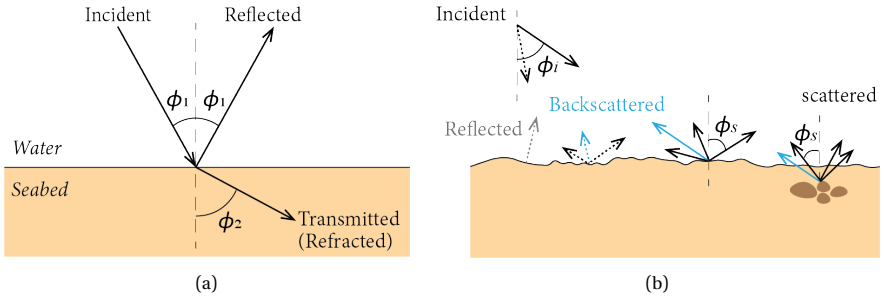


Figure 1.1: Acoustic interaction with the seabed illustrated for (a) an ideally smooth and homogeneous seabed and (b) a seabed with a rough surface and certain volume heterogeneity.

Reflection and transmission

At an ideally smooth seabed surface, apart from a small fraction of acoustic energy absorbed within the sediment, specular reflection and transmission dominate. Relative to the normal of the interface, the angle of reflection equals the incident angle ϕ_1, whereas the transmission (refraction) angle ϕ_2 (Fig. 1.1a) follows Snell's law

$$\frac{\sin\phi_2}{\sin\phi_1} = \frac{c_2}{c_1}, \tag{1.1}$$

with c_1 and c_2 the sound speed in seawater and sediment, respectively. The reflection coefficient R, describing the relation between the incident sound pressure p_i and reflected pressure p_r, is expressed as

$$R(\phi_1) = \frac{p_r}{p_i} = \frac{Z_2\cos\phi_1 - Z_1\cos\phi_2}{Z_2\cos\phi_1 + Z_1\cos\phi_2}, \tag{1.2}$$

where Z_1 and Z_2 represent the acoustic impedance of seawater and sediment, respectively. The acoustic impedance quantifies the resistance of a medium to particle motion during the propagation of sound waves and is determined by the bulk density ρ and sound speed c as $Z = \rho c$. Moreover, the transmission coefficient T is defined as

$$T(\phi_1) = \frac{p_t}{p_i} = \frac{2Z_2\cos\phi_1}{Z_2\cos\phi_1 + Z_1\cos\phi_2} = 1 + R(\phi_1), \tag{1.3}$$

with p_t the transmitted sound pressure [16].

Scattering

Natural seabeds seldom fulfill the above smooth-surface assumption because of the interface roughness and volume heterogeneity relative to the acoustic wavelength. Such seabed irregularity scatters sound away from the direction of specular reflection [17]. During a hydrographic survey, if an acoustic signal is sent from the vessel to the seabed, two types of returns can be collected: i) specular reflection at or near the normal incidence for a relatively smooth seabed surface and ii) scattering towards the direction of the vessel, with the scattered angle equal the incident angle. The latter is called backscattering, which is a fundamental part of active acoustic remote sensing.

Similar to the reflection and transmission coefficients, scattering can also be quantified based on the relation between the incident and scattered sound. With the acoustic intensity I, which measures the mean energy carried by an acoustic wave per unit area, the scattering cross-section σ_s is expressed as

$$\sigma_s(\phi_i, \phi_s) = \frac{I_s(\phi_s)}{I_i(\phi_i)}, \tag{1.4}$$

where ϕ_i and ϕ_s are the incident and scattered angle, respectively (Fig. 1.1b). I_i and I_s denote the incident and scattered acoustic intensity at a distance of 1 m from the scattering object. σ_s is a dimensionless ratio defined per unit area and unit solid angle [18]. $\phi_s = \phi_i$ then delivers the backscattering cross-section σ_b. In practice, the backscatter strength $BS = 10\log(\sigma_b(\phi_i))$ in decibel per m^2 at 1 m is commonly used.

The angular variation of $BS(\phi_i)$ is dependent on the seabed characteristics. For rough interfaces and high acoustic frequencies, this relationship can be simplified by Lambert's law. Assuming the incident acoustic wave ensonifies an area A on the seabed, the corresponding acoustic power can be expressed as $P_i = I_i A\cos\phi_i$. I_s is then written as [19]

$$I_s = \mu P_i \cos\phi_s = \mu I_i A \cos\phi_i \cos\phi_s, \tag{1.5}$$

assuming that P_i is radiated proportional to $\cos\phi_s$ according to a constant Lambert parameter μ, which is intrinsic to seabed properties [20]. σ_s then becomes

$$\sigma_s(\phi_i, \phi_s) = \mu A \cos\phi_i \cos\phi_s, \tag{1.6}$$

leading to the backscattering cross-section $\sigma_b(\phi_i) = \mu(\cos\phi_i)^2$ when A represents the unit area (1 m^2) and

$$BS(\phi_i) = 10\log(\sigma_b(\phi_i)) = 10\log\mu + 20\log(\cos\phi_i). \tag{1.7}$$

Interpreting backscattered acoustic signals is an inverse problem. By analyzing BS and its angular patterns, we try to infer seabed properties that have an impact on them. The acoustic frequency also plays a key role here. Scattering of a high-frequency acoustic wave, whose wavelength is smaller or at the same scale as the relief of seabed surface structures, can be more affected by the interface roughness. Low-frequency signals might penetrate deeper into the seabed, making the acoustic wave scattered not only by the surface but internal heterogeneities such as gas bubbles and sediment layering.

Impact of marine benthos

Marine benthos influence sound–seabed interaction mainly from two aspects. Some epibenthos, organisms living on the sediment surface, increase the seabed roughness with their hard shells [21, 22]. Their assemblages with a certain spatial scale, such as oyster and mussel reefs, can also increase the seabed surface hardness and thus the acoustic impedance. In addition, endobenthos, which are species dwelling usually in the upper 25 cm of the sediment body, alter seabed surface and bulk characteristics via activities such as tube-building and burrowing [23, 24] (Fig. 1.2). Changes in these properties have a potential impact on the reflection, transmission, and scattering of acoustic signals, making the associated benthic habitats therefore detectable with acoustic remote sensing.

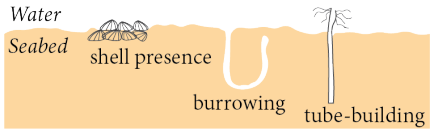


Figure 1.2: Marine benthos influence the seabed characteristics via their occurrence and activities such as burrowing and tube-building.

1.2. Acoustic benthic habitat mapping

1.2.1. Sonar systems

The idea of using sound to detect underwater objects dates back to the early 20th century. The first echo-sounding device was patented in 1913, one year after the sinking of *Titanic* [25]. During World Wars I and II, this technique was adapted to detect submarines and the term "sonar" (sound navigation and ranging) entered the naval vocabulary. Since then, a family of sonar systems has evolved for mapping the seafloor:

- **Single-beam echosounder (SBES)**. Adopted for routine hydrographic surveys since the 1920s, the SBES vertically transmits an acoustic signal in a narrow beam to the seafloor while mounted on a vessel. Operating frequencies of SBES systems range from 12 to 400 and even 700 kHz, depending on the application in deep or shallow waters [16]. Water depths can be derived based on the two-way travel time and sound speed. The returned echo amplitude and waveform also provide information on objects in the water column, the acoustic impedance of the seabed, and potentially the subsurface properties [26]. Due to the limited beam footprint, large gaps remain between survey lines. Nevertheless, the SBES is still widely used in applications such as fisheries, considering its low cost and relatively simple configurations [27].

- **Side-scan sonar (SSS)**. Developed in the 1950s and commercialized during the 1960s, the SSS comprises paired transducers that emit fan-shaped beams at both port and starboard sides [16]. The SSS is usually operated at 100–500 kHz and towed behind the vessel a few meters above the seabed, returning signals that

represent backscatter intensity as a function of time. With a small pulse length (typically 0.1 ms or less), high-resolution seabed images can be obtained, offering insights for trawl marks, shipwrecks, and other manmade and geomorphic structures. Precise localization of seabed objects in the SSS image, however, can be challenging [28].

- **Multibeam echosounder (MBES)**. The MBES was first demonstrated in the 1970s and widely adopted in hydrographic practices during the 1980s [16]. Mounted beneath the vessel, the MBES transmits sound in a wide swath perpendicular to the sailing direction and forms hundreds of narrow beams with known angles during reception [29]. The range of operating frequencies of MBES systems is similar to that of the SBES. With an MBES, co-registered bathymetry and *BS* data can be obtained for each beam. The angle information further enables the investigation of backscatter angular dependence, which is intrinsic to different seabed types [30]. Integrated with the navigation system and motion sensors onboard, the location of each beam can be accurately inferred, allowing nearly continuous coverage of the seafloor.

- **Synthetic aperture sonar (SAS)**. Demonstrated at sea in the late 1990s, SAS adapts the synthetic aperture principle from radar (radio detection and ranging), coherently combining successive pings to form a virtual extended array [31]. Although SAS is operated at a roughly similar frequency range as high-resolution SSS, it yields an along-track spatial resolution at the centimeter scale, which is range-dependent and exceeds the performance of conventional SSS. Mounted on autonomous or towed platforms, modern SAS systems nowadays deliver detailed seabed information. Nevertheless, preserving the imaging quality requires sophisticated positioning and a very stable platform, largely increasing costs in budget and time compared to other systems.

- **Sub-bottom profiler (SBP)**. Developed in the 1930s, SBPs refer to seismic systems that employ acoustic signals with a low frequency ($\leq$ 10 kHz) to sense the seabed subsurface [32]. The SBP can be mounted on a vessel, with the transmitter and receiver combined as a transducer, or towed behind the vessel, when the sound source and receiver are separated. The adopted low-frequency sound can penetrate tens of meters into unconsolidated sediments, helping to reveal crucial information on seabed stratification and buried infrastructure.

Although each sonar system presents distinct advantages, the MBES provides a balance of spatial coverage, efficiency, and seabed information richness, making it the most cost-effective tool for large-scale seabed mapping.

1.2.2. From sediment mapping to benthic habitat characterization

Sediment types are a key indicator for benthic communities. For example, sand mason worms inhabit the sandy bottom mixed with shell fragments to build tubes [33], while oysters and mussels tend to attach to hard substrates such as rocks [34]. Continuous sediment maps have also been the primary objective in acoustic remote sensing. With

hydro-acoustic data, especially *BS*, we can efficiently divide sediments into "fine", "medium", and "coarse" [35], or "soft" and "hard" [36]. Beyond such rough classification, directly linking acoustic measurements to the mean or median grain size has proved effective [37, 38]. Similar predictive modeling was also conducted for the percentages of sand, mud, and gravel [39]. Furthermore, the Folk scheme [40] classifies common sediments based on the median grain size into different types such as Sandy Clay, Fine Sand, Coarse Sand, Sandy Gravel, and Rock, which becomes standard ground truth classes in many acoustic seabed mapping studies [20, 41, 42].

Besides the grain size, maps of other geoacoustic parameters, such as acoustic impedance, spectral strength (which describes the sediment roughness), and the volume scattering parameter, can be derived from backscatter measurements based on laboratory-built physical models. Such model inversion has been applied to bottom echo envelopes measured by the SBES [43] and angular *BS* data from the MBES [30, 44]. In addition, acoustic measurements can help to achieve continuous maps for sediment compositions such as carbon and sulfide [45, 46]. High-resolution bathymetry from the MBES has also facilitated the mapping of seabed geomorphology types [47, 48].

We may link information on sediment types to the occurrence of marine benthos [49]. Nevertheless, considering the impact of benthos on seabed geoacoustic properties, a comprehensive classification of the benthic habitat characteristics is increasingly needed nowadays. A recent study [14] summarized four main types of benthic habitat maps anticipated in research and industry:

- **Abiotic properties**, such as the aforementioned sediment maps which can be surrogates for biological patterns.
- **Spatial distribution of a single benthic organism**. For example, MBES data have been successfully adopted for mapping the presence of specific sponges in Australia [50] and a type of Arctic clam [51].
- **Benthic community characteristics**, such as distributions of co-occurring organisms or biodiversity. Previous research classified the seabed with MBES data into different groups of benthic organisms, such as algae, invertebrates, or a mixture of several species [52, 53]. MBES bathymetry and backscatter were also used as predictors for species diversity and functional traits of benthic communities [54].
- **Bio-physical benthic classification**, which involves discrete categories depicting both biotic and abiotic properties. This type of mapping helps to achieve a holistic view of the study areas and is commonly considered in regional seabed research. In Canada and Australia, acoustic remote sensing techniques, especially the MBES, have been applied to simultaneously map sediment types based on grain sizes and the occurrence of benthos, such as bivalves, algae, and sea urchins [55–57]. In the North Sea, acoustic measurements also enable researchers to distinguish shells [58] or kelp [59] from the sandy bottom.

1.2.3. From single-frequency to multi-frequency analysis

Seabed irregularities that cause scattering of sound are relative to the acoustic wavelength. Take sediment grain size as an example. When the grain size acts as the controlling

factor for seabed variations and is smaller than the acoustic wavelength, we can expect a positive correlation between the grain size and seabed *BS*. However, such correlation might be ambiguous when the grain size approximates the acoustic wavelength, and can become negative after it further increases [41, 60]. It is therefore possible that fine and coarse sediments result in similar *BS* for a single acoustic frequency. In addition to the ambiguity regarding sediment grain size, the presence of marine benthos, shell remains, and gas bubbles further complicates the interpretation of single-frequency backscatter measurements.

Previous research has accounted for bathymetry and its derivatives to mitigate this challenge. Bathymetry was found to be a more important indicator than seabed *BS* for an algae habitat [53]. Depth residuals, which represent the deviation of measured bathymetry from a fitted surface patch, also helped to identify an artificial stone layer in the riverbed sediment [35]. Nevertheless, benthic habitats with simple seabed morphology exist [61]. The use of depth residuals, on the other hand, requires additional consideration for distinguishing the actual seabed roughness from bathymetric uncertainties.

Apart from investigating various acoustic features, exploring underwater acoustics at multiple frequencies has been another long-standing focus of research. Surveying the same area several times, each with a different operating frequency, is a possible solution [62, 63]. However, obtaining such a multi-frequency dataset for the same seabed is challenging in a dynamic sediment environment. To achieve multi-frequency data in a single survey, the dual-frequency SBES has been widely adopted [64]. Relatively simple configurations of the SBES provide the flexibility of installing two transducer elements in one system. Mounting multiple SSS or MBES systems tuned to distinct frequencies on the same vessel was also implemented [65, 66]. Although this method presents additional challenges for accounting for various beam patterns of different systems in post-processing, it can still improve the discrimination of seabed types [66]. Moreover, since the penetration of acoustic signals varies with wavelength, multi-frequency analysis can facilitate the understanding of sediment volume characteristics [67, 68].

Recent advances in the multi-spectral MBES system bring new potential for multi-frequency seafloor mapping. A multi-spectral MBES is capable of modifying the operating frequency on a ping-by-ping basis and allows for simultaneously collecting multi-frequency acoustic returns with a single system [69], which enables a further improvement in broad-scale seabed characterization.

1.2.4. From physics-based to data-driven modeling

In real-world benthic habitat characterization, we hope to build a model that takes acoustic measurements as input and predicts the desired mapping target. Physics-based methods realize this goal from the opposite direction, by first modeling the acoustic reflection, transmission, and scattering processes. For example, the widely-accepted APL-model [18] estimates *BS* at different incident angles for a given acoustic frequency and sediment type, which is defined by geoacoustic parameters such as median grain size, spectral strength, and the volume scattering parameter. The sediment is usually assumed as a semi-infinite, dissipative, and homogeneous fluid medium. Based on the APL-model, we can invert seabed geoacoustic parameters using the field-measured *BS*.

The development of such physical models has laid a solid foundation for understanding the sound–seabed interaction. Some studies also assessed the impact of marine benthos by examining the deviations of measured *BS* from model predictions based solely on sediment properties [24]. Nevertheless, how to account for the occurrence of marine benthos in physical models is still an ongoing research topic.

In addition, physical model inversion relies on well-calibrated *BS* data. Due to the varied sonar sensitivity in different water environments, absolute backscatter calibration is difficult to achieve for every single survey and requires a deeper understanding of the sonar mechanical properties [70, 71]. Although the APL-model has been extended to uncalibrated multibeam backscatter data through an optimization-based calibration curve estimation [72], the inversion performance was found to be affected by the narrow range of sediment types.

By contrast, data-driven modeling is more flexible regarding the uncalibrated backscatter measurements. The modeling does not explicitly involve the physical processes, but instead learns empirical relationships between acoustic data and seabed properties. Assuming that the sonar calibration terms are fixed values during a single survey, the uncalibrated *BS* is still representative of variations in seabed characteristics. Moreover, other types of data such as bathymetry can be included. Unlike physics-based approaches, data-driven models are not confined to geoacoustic parameters, but can be used for describing diverse bio-physical properties, providing a powerful alternative for benthic habitat mapping.

On the other hand, model construction and/or validation of data-driven methods require the seabed ground truth, which is usually provided by bottom samples. We can classify these methods based on how they utilize the ground truth information. For instance, *supervised* methods fit an empirical model entirely using labeled data, i.e., the input and target variables at known locations. This empirical relationship can be described by parametric statistical models, such as ordinary least-squares linear regression, generalized linear models, and generalized additive models, in which contributions of covariates are represented by the fitted coefficients. Besides, machine learning (ML) methods, such as Random Forest, and deep learning (DL) methods, such as neural networks, provide higher flexibility for fitting the input–target relationship with a larger number of parameters. They require few assumptions on data distributions, have better abilities to deal with multi-dimensional features, and are commonly applied in classification and regression tasks for seabed mapping [73–79]. Despite this, developing robust supervised ML and DL models often requires a large amount of labeled data for training, which can be challenging in hydro-acoustic seabed surveys.

Without relying on labeled data for model construction, *unsupervised* methods identify patterns in acoustic data and group them as different classes. Afterward, a small set of ground truth data is still necessary for validating the classes and assigning them to geological or biological facies. Such methods in seabed mapping range from fitting assumed probability distributions to *BS* histograms [80] to clustering acoustic data with ML algorithms such as K-means [35, 47]. Regarding DL paradigms, self-supervised techniques that pre-train feature extractors on unlabeled data have also been applied to enhance seabed classification [81]. Bridging the gap between fully supervised and unsupervised learning, *semi-supervised* methods start with limited labeled data and

exploit the structure of unlabeled features through techniques such as self-training and pseudo-labeling, to refine decision boundaries and improve the robustness of model training [82–84].

1.3. Scope of this thesis

1.3.1. Technical challenges

Sustainable coastal development and marine conservation policies nowadays demand efficient and non-destructive mapping of benthic habitats. The multi-spectral MBES, which can acquire co-registered bathymetry and acoustic backscatter at several frequencies in one survey, offers broad spatial coverage and enhanced discrimination ability to meet such requirements. Moreover, various data-driven methods provide high potential and flexibility for modeling the relations between multi-spectral MBES measurements and seabed characteristics. Despite these advantages, benthic habitat mapping with MBES data still faces several critical challenges:

- **Limited ground truth data**. Seabed ground-truthing, usually achieved by bottom samples and videos, can be costly and restricted by regulations, resulting in only sparse reference data for model development and validation.
- **Lack of absolute backscatter calibration**. Although uncalibrated BS from a single survey can be linked to seabed variations, comparing or combining MBES data from multiple surveys requires solutions for eliminating the impact of possible drifts in sonar characteristics.
- **Across-track angular and spatial variations of MBES backscatter**. With limited overlaps between survey lines, it can be challenging to achieve spatially continuous BS data and high-resolution backscatter angular information.
- **Large data volume of multi-frequency measurements**, which requires efficient processing and effective feature extraction for high-dimensional data.

1.3.2. Research objectives

Based on the current trend and technical challenges of acoustic benthic habitat mapping, the main research objectives of this thesis are four-fold:

i) Mapping the occurrence of "acoustically hard" marine benthos, including epibenthos that modify the seabed surface characteristics and endobenthos that affect the sediment volume properties, and their associated sediment types with multi-spectral MBES measurements.

ii) Developing data-driven acoustic seabed mapping methods with optimal use of limited ground truth data.

iii) Improving the MBES data processing workflow to effectively combine various multi-frequency features, such as bathymetry, seabed backscatter, and BS angular responses, from multiple surveys in seabed mapping.

iv) Providing survey guidelines for future mapping and monitoring of marine benthos, with a focus on shellfish beds in the Dutch North Sea.

1.3.3. Study locations

This thesis focuses on benthic habitats in the Dutch coastal area (Fig. 1.3). As one of the busiest marine environments worldwide, the Dutch North Sea embraces shipping, wind farm construction, sand extraction, fisheries, and other economic and engineering activities. Accurate benthic habitat mapping is a priority for planning offshore activities and managing marine ecosystems in the Dutch North Sea.

The relevant seabed surveys in this thesis were mainly conducted north of the Wadden Sea islands, covering habitats of tube worms and bivalves that hold significant conservation values (Location A and B). Besides, one of the study sites (Location C) is located southwest of the Netherlands, overlapping with a well-known mussel cultivation region. MBES datasets collected during these surveys build a solid basis for developing epibenthos and endobenthos mapping methods in this thesis.

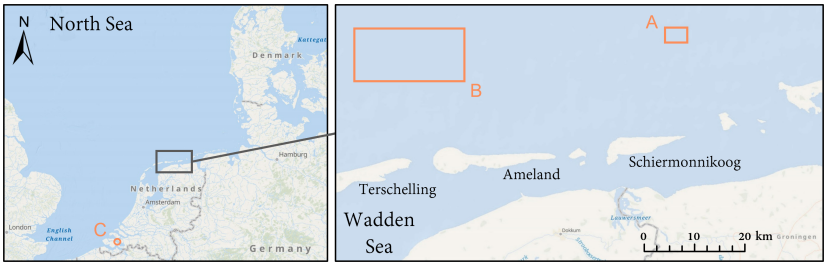


Figure 1.3: An overview of the study locations relevant to this thesis.

1.4. Thesis outline

To provide background knowledge for the following chapters, we take a closer look at the MBES data processing theory in **Chapter 2**. Beam geometry and its implication on the ensonified footprint area are briefly introduced. We also present the steps of data pre-processing, with detailed descriptions of backscatter correction and relative calibration. Afterward, a common workflow for processing MBES data with data-driven methods is described.

In **Chapter 3**, we compare two unsupervised backscatter classification methods: Bayesian classification, which takes beam-averaged BS at a single incident angle as input, and hierarchical clustering for angular mean values of BS, on two multi-spectral MBES datasets acquired in the Dutch North Sea (Location A and B in Fig. 1.3). The potential of BS angular features for distinguishing the sand mason worm habitat and coarse sediments mixed with dead shells is shown.

As the follow-up research of Chapter 3, **Chapter 4** involves MBES data collected after almost three years for the same study areas. After a relative backscatter calibration to account for the impact of water temperatures on sonar characteristics, we conduct

change detection via Gaussian Mixture Model clustering on the backscatter difference, revealing possible sediment movement and seasonal changes in sand mason worms between the two survey times.

In **Chapter 5**, we apply semi-supervised classification for mussel detection in a blue mussel farming area (Location C in Fig. 1.3). With a single-frequency MBES dataset and only nine ground truth samples, our method delivers largely improved classification maps for mussel presence/absence, compared to conventional supervised techniques.

In **Chapter 6**, we focus on characterizing the shellfish habitat (Location B in Fig. 1.3) in different seasons with multi-frequency MBES measurements. We achieve spatially continuous *BS* data per angle with semi-supervised learning and enhance the classification performance on the multi-frequency multi-angle data with an unsupervised feature reduction neural network.

Based on insights from previous chapters, **Chapter 7** provides recommendations for future seabed surveys, with a focus on ensuring the quality of MBES measurements for mapping shellfish beds. As a summary of this thesis, **Chapter 8** presents conclusions and an outlook for future research on benthic habitat mapping.

Table 1.1: Research objectives (Section 1.3.2) covered by the following chapters

Chapter \ Objective	i	ii	iii	iv
3	✓	✓		
4		✓	✓	
5	✓	✓		
6	✓	✓	✓	
7				✓

2

MBES data processing

The ocean is not a place, but a process.

Carl Safina (1955—)

This chapter introduces the theoretical background and practical knowledge of MBES data processing in seabed mapping. The principle of beamforming and the beam geometry are illustrated. The pre-processing steps necessary to prepare MBES measurements for seabed mapping are outlined, with an emphasis on backscatter correction based on the sonar equation and relative calibration needed for integrating data from multiple surveys. Following this, common steps for using MBES data in supervised and unsupervised data-driven methods are described.

2.1. Multibeam theory

As an active acoustic remote sensing technology, the MBES transmits sound and then receives the returned signals. Most MBES systems are designed under the Mill's cross principle, with the transmitter and receiver physically separated and orthogonal to each other [16]. The transmitter emits an acoustic signal spanning across a wide swath. After the transmitted pulse reaches the seabed, the receiver captures scattered or reflected acoustic signals from specific angles via *beamforming*, resulting in a set of beams and bringing the concept of *multibeam* (Fig. 2.1).

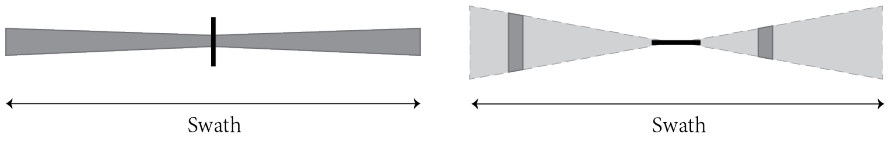


Figure 2.1: Top view of directivities of an MBES under the Mill's cross setup for (Left) the transmission array and (Right) two beams formed by the reception array.

2.1.1. The principle of beamforming

The property of concentrating acoustic energy on a particular angle in space is called *directivity* [15]. At transmission, it describes the angular distribution of the transmitted energy; on reception, it represents the acoustic response as a function of the arrival angle of the wavefront [16]. Directivity relies on constructive interference of acoustic waves at a specific angle and destructive interference in other directions. Therefore, at least two projector or receiver elements are required, forming a transmission or reception array.

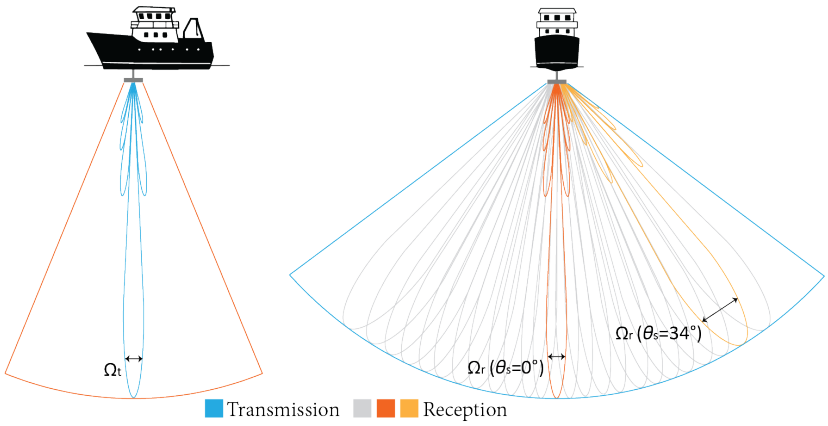


Figure 2.2: MBES directivity patterns of the transmission and reception array with the (Left) along-track and (Right) across-track view (adapted from [85]). Ω_t and Ω_r represent the beam width during transmission and reception, respectively. θ_s denotes the steering angle at reception.

For a conventional MBES, the array can be imagined as a line of identical elements. Projector elements on the transmission array provide an angular distribution that is

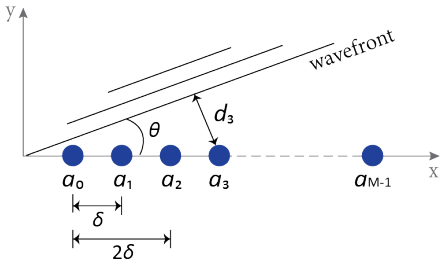


Figure 2.3: A plane acoustic wave impinging on a line array with M identical receiver elements. δ is the spacing between elements, while θ indicates the direction of the incoming wave.

wide in the across-track direction and has a narrow opening angle in the along-track direction (Fig. 2.2). During reception, the directivity pattern is related to the direction (θ) of the incoming plane acoustic wave (Fig. 2.3). Given M receiver elements and a fixed spacing δ between them, the acoustic response of the array is a summation of the response of all elements, which can be expressed as

$$g(\theta) = \frac{\sin(MK)}{\sin K}, \text{ with } K = \frac{\pi\delta}{\lambda}\sin\theta, \tag{2.1}$$

where λ indicates the acoustic wavelength. The directivity pattern, defined as the squared magnitude of the normalized $g(\theta)$, is

$$D(\theta) = \left|\frac{g(\theta)}{M}\right|^2 = \left|\frac{\sin(MK)}{M\sin K}\right|^2. \tag{2.2}$$

$D(\theta)$ peaks at $K = 0$ (i.e., $\theta = 0$), which is the center of the main lobe. The main lobe is characterized by a -3 dB width on each side of its maximum, which is also indicated as the *beam width*. For a broad array with length $L \approx M\delta$, the beam width can be calculated as $\Omega = \lambda/L$.

With $\theta = 0$, the incoming acoustic wave is perpendicular to the array, indicating that sound is best located when the wavefront reaches each receiver element at the same time. In other words, we can rotate the reception array to capture sound from the desired angle. Because physically steering the array is impractical, beamforming is applied to electronically achieve such filtering. By delaying the time of acoustic signals arriving at the various receiving elements, the signals received at these elements are aligned in time.

For instance, electronically steering the reception array to the direction of θ_s requires a time delay at the n-th element a_n relative to the reference element a_0 (Fig. 2.3) of

$$\tau_n = \frac{d_n}{c} = \frac{n\delta}{c}\sin\theta_s, \tag{2.3}$$

where c is the sound speed. τ_n is equivalent to a phase shift

$$\varphi_n = \omega\tau_n = \frac{2\pi f}{c}n\delta\sin\theta_s = \frac{2\pi}{\lambda}n\delta\sin\theta_s, \tag{2.4}$$

where $\omega = 2\pi f$ is the circular frequency.

As a result, the directivity pattern becomes

$$D(\theta) = \left| \frac{\sin[M(K-K_s)]}{M\sin(K-K_s)} \right|^2, \text{ with } K_s = \frac{\pi\delta}{\lambda}\sin\theta_s, \tag{2.5}$$

peaking at $\theta = \theta_s$. On the other hand, such filtering achieved by beamforming will broaden the main lobe for $\theta_s \neq 0$ (Fig. 2.2) according to the length of the virtually steered array and result in the beam width

$$\Omega' = \frac{\lambda}{L\cos\theta_s}. \tag{2.6}$$

2.1.2. Beam geometry

Array characteristics and beamforming determine the transmission and reception beam width of an MBES, which govern spatial resolution and the instantaneously ensonified footprint area A [m^2] on the seabed. A is also closely related to the water depth and local slopes of the seabed.

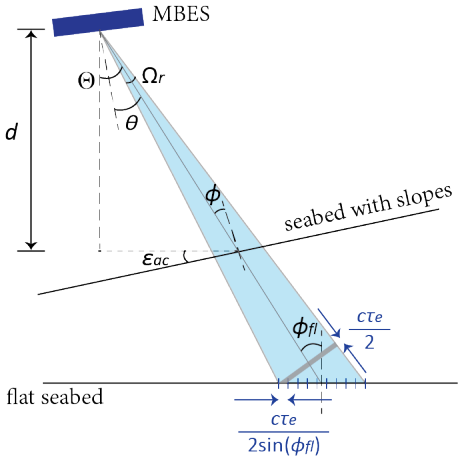


Figure 2.4: Across-track view of an oblique beam (adapted from [86]), with Θ the beam angle, θ the steering angle, ϕ_{fl} the incident angle under the flat-seabed assumption, and ϕ the actual incident angle relative to seabed slopes. d denotes the water depth measured by the MBES. Ω_r indicates the across-track beam width regarding the steering angle, while τ_e is the effective pulse length.

Let's first consider a flat seabed and an incident angle ϕ_{fl} (Fig. 2.4). Given the one-way travel distance R [m] of an acoustic signal, A in the beam-limited regime equals [86]

$$A_b = R^2\Omega_t\Omega_r, \tag{2.7}$$

where Ω_t and Ω_r are the transmission and reception beam width in radians that can be estimated by Eq. 2.6. In the pulse-limited regime, A is determined by the resolution of the time signal (Fig. 2.4) and becomes

$$A_p = \Omega_t R \frac{c\tau_e}{2\sin(\phi_{fl})}, \tag{2.8}$$

where τ_e is the effective pulse length in seconds.

Now assuming a seabed with the across-track slope ϵ_{ac} and along-track slope ϵ_{al} in radians, which can be estimated by a 2D finite difference method using bathymetry data [87], A_b and A_p are then modified to

$$A_b = \frac{R^2\Omega_t\Omega_r}{\cos(\epsilon_{ac})\cos(\epsilon_{al})}, \tag{2.9}$$

and

$$A_p = \Omega_t R \frac{c\tau_e}{2\sin(\phi_{fl}-\epsilon_{ac})\cos(\epsilon_{al})}, \tag{2.10}$$

respectively. In practice, we determine A as the smaller value between A_b and A_p. For the nearly flat seabed in the Dutch North Sea, we can calculate A_b according to Eq. 2.7 to avoid including uncertainties of the slope estimation.

Moreover, the actual incident angle relative to the seabed is calculated as

$$\phi = \arccos\left[\frac{\cos(\phi_{fl}) + \epsilon_{ac}\sin(\phi_{fl})}{\sqrt{1+\epsilon_{al}^2+\epsilon_{ac}^2}}\right]. \tag{2.11}$$

Bathymetry measurements

An MBES estimates the water depth d relative to it via measuring the two-way travel time t for a given beam angle Θ (Fig. 2.4). Θ can differ from the steering angle θ due to the mounting angle of the MBES and the platform motion [88]. d is calculated as

$$d = \frac{ct}{2}\cos\Theta, \tag{2.12}$$

with a constant sound speed c in the water column. In reality, c can vary with depth considering the water temperature. Therefore, an accurate d is usually retrieved by ray-tracing, which approximates the water column with multiple vertical layers of different sound speeds to reconstruct the travel path of an acoustic signal.

2.2. Pre-processing for seabed mapping

For seabed mapping applications in this thesis, MBES data pre-processing aims to achieve cleaned bathymetry data and the beam-averaged backscatter strength (BS).

- **Bathymetric cleaning** removes depth anomalies possibly caused by false bottom detection, by considering the neighborhood statistics and uncertainties from, for example, the echosounder itself, motion sensor, and sound speed profile [89, 90]. Various filtering algorithms to optimize this process have been implemented in common commercial hydrographic software [91].

- **Backscatter pre-processing** corrects each sounding for the acoustic wave propagation in the water column, the ensonified footprint on the seabed, and the sonar characteristics during data collection. This procedure can be divided into two steps: i) backscatter correction and ii) relative calibration.

2.2.1. Backscatter correction

Apart from geophysical and biological seabed properties, the echo level received by an MBES is affected by many factors, including acoustic frequency, incident angle, water column properties, seabed geomorphology, and certain sonar settings. To accurately make use of MBES backscatter measurements for seabed characterization, it is essential to account for these factors.

Sonar equation

Given an acoustic frequency f and incident angle ϕ, BS [dB per m^2 at 1 m] exclusively representing the sea bottom material properties can be achieved according to the sonar equation [16]:

$$BS(\phi, f) = EL - SL - G - SH - D_T(\theta_T) - D_R(\theta_R) + TL - 10\log A, \tag{2.13}$$

where EL indicates the received echo level [dB] and SL is the source level [dB re 1 μPa at 1 m]. $D_T(\theta_T)$ and $D_R(\theta_R)$ [dB] are the directivity patterns at the transmission steering angle θ_T and the reception steering angle θ_R. G and SH refer to the receiver gain [dB] and transducer sensitivity [dB re 1 V/μPa], respectively. A indicates the ensonified footprint area [m^2] (see calculations in Section 2.1.2). In addition, TL accounts for the two-way transmission loss through the water column via

$$TL = \underbrace{40\log R}_{\text{geometrical spreading}} + \underbrace{2\alpha R}_{\text{seawater absorption}}, \tag{2.14}$$

with R the one-way travel distance of the acoustic signal from the sonar to the seabed and α the absorption coefficient [dB/m].

Absolute and relative backscatter correction

Accounting for all factors described in Eq. 2.13 results in the absolute backscatter correction, which is challenging in practice due to the limited knowledge of SH and the directivity patterns. Nevertheless, these unknown parameters can be considered constant during a single survey. By properly handling the other factors, we can still achieve the uncalibrated or relative BS representing the spatial variation in seabed properties.

Backscatter mosaic

Although the MBES offers a nearly continuous seabed coverage, the measured BS is incoherent in space due to its angular dependency and a limited overlap between adjacent swaths (Fig. 2.5). A common approach to achieve a spatially continuous backscatter image, namely the backscatter mosaic, is to remove the BS angular variation. With the relatively corrected BS, a mean angular response curve (ARC) can be calculated with

an along-track sliding window. The BS level of each incident angle is then normalized based on the average BS at reference incident angles, which are usually between 30° and 60° [58, 92, 93].

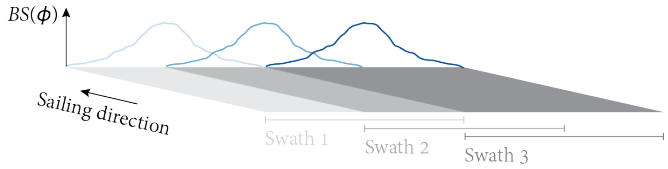


Figure 2.5: Illustration of BS measurements achieved by three adjacent swaths with a 50% overlap. ϕ indicates the actual incident angle relative to the seabed.

2.2.2. Relative calibration

Although the relative backscatter correction is sufficient for analyzing data collected in a single survey, comparing or combining backscatter data across multiple surveys requires additional consideration for potential changes in the sonar characteristics. For an MBES, drift in the transducer sensitivity and directivity can happen over time. Recent research also revealed a correlation between MBES backscatter and seawater temperature, which is possibly attributed to the variation in acoustic properties of the transducer material regarding different temperatures [71].

An ideal solution is to conduct absolute calibration periodically for an MBES system, usually through tank measurements of the transducer properties or by utilizing an artificial reference object with a known BS-frequency dependency [94]. However, strict requirements for infrastructure and equipment pose challenges to making absolute backscatter calibration a universal strategy. The potential temperature-dependency of BS further complicates the calibration procedure.

Acquiring the absolute BS is still essential for understanding the seabed geoacoustic properties and developing physical models. Nevertheless, alternative solutions can be considered to ensure that backscatter data collected in different surveys are mutually comparable, especially for the use in data-driven MBES mapping.

This strategy can be a relative backscatter calibration based on a stable seabed, or reference area. Consider a single MBES system employed for measuring the reference area at time t_1 and t_2. For each incident angle ϕ and acoustic frequency f, the relative BS of the reference area at both survey times, $BS(\phi, f, t_1)$ and $BS(\phi, f, t_2)$, can be utilized to generate a relative calibration term

$$C(\phi, f) = BS(\phi, f, t_2) - BS(\phi, f, t_1). \tag{2.15}$$

Given an unchanged reference area, the relative calibration is then conducted by compensating $C(\phi, f)$ in backscatter data from one of the surveys. This method can also be suitable for a combined use of measurements from different MBES systems.

Reference area selection

In principle, the reference area should remain sufficiently stable over time in terms of bathymetry and geomorphology, as validated by successive surveys [70, 94]. The reference area selection follows several criteria. First, the area should be flat and stable in bathymetry. The bathymetric change of the area between t_1 and t_2 should remain within the uncertainty of MBES depth measurements. The flatness can be examined by the seabed slope estimations, and the occurrence of simple unidirectional patterns is acceptable [70]. Second, the area should exhibit homogeneity in backscatter data. In addition to acoustic measurements, the stability of the reference area should be validated by regular seabed ground-truthing. The change in sediment grain size, benthos composition, and presence of dead shell material should be assessed, considering their potential influence on the backscatter measurements.

2.3. Workflow for data-driven MBES mapping

Depending on the choice of supervised or unsupervised methods, the general workflow for mapping seabed properties using MBES measurements consists of five or six stages:

- **Data pre-processing** described in Section 2.2.

- **Feature extraction**, which is usually conducted on bathymetry and backscatter data to derive additional discriminative descriptors of the seabed.

  Based on the bathymetry map, derivatives such as the depth residual, slope, and curvature at one or multiple spatial scales can be calculated. For mapping benthic habitats that present distinct seabed geometry (e.g., coral reefs), geometric features extracted from the 3-dimensional (3D) bathymetric point cloud are also useful indicators. Regarding backscatter data, local textural features can be derived from the grayscale backscatter mosaic [95], whereas mean level, slope, and other statistics calculated on angular response curves (ARCs) help to leverage the discrimination ability of BS angular variations [30]. For multi-frequency measurements, descriptors for inter-frequency contrast, such as the depth difference, can also enrich the acoustic feature set.

  Apart from manual feature extraction, neural networks are a common tool nowadays to extract informative patterns from the high-dimensional feature space in both supervised and unsupervised modes. With all extracted features, normalization or standardization is usually necessary to avoid vast scale differences among features.

- **Feature reduction**. Although feature extraction is crucial for analyzing patterns in MBES measurements, it can increase the data dimensionality and induce redundant information among features. Therefore, a feature reduction step is needed, which can be achieved by: i) removing features that are highly correlated with other features or have little mutual information with the target variable (in the supervised case); ii) projecting initial features into a lower-dimensional space with various techniques, including linear methods, such as the largely accepted

principal component analysis (PCA) that preserves most data variance, and non-linear manifold learning methods, which account for complex data structures but can be more computationally expensive.

- **Data split**, which is required in terms of supervised approaches to divide the labeled acoustic data into training, validation, and test datasets.
- **Model construction** on the reduced features. Unsupervised methods estimate parametric probabilistic distributions or cluster the features into different groups based on ML or DL algorithms. Regarding supervised methods, training data are adopted for fitting the model parameters, while the validation set is used for hyper-parameter tuning, helping to avoid overfitting to the training data. When the labeled data have a small size, using an individual validation set can be replaced by cross-validation. For example, the training data can be split into k folds with the same size, with each fold used once as the validation data. After tuning the hyper-parameters based on the averaged performance of k runs, the final model can be fitted by the whole training dataset.
- **Model evaluation**. In the case of supervised methods, model evaluation is conducted on the untouched test set. Unsupervised mapping results should also be validated with certain ground truth data or by expert interpretation.

Moreover, semi-supervised variants can include additional steps such as pseudo-labeling to leverage information from the large pool of unlabeled data, which is not mentioned in this general workflow.

3

Indications of sand mason worm habitats from multi-frequency multibeam backscatter data

To identify the marine benthos occurrence using the multibeam echosounder (MBES), relationships between MBES backscatter data and seabed compositions need to be investigated. Following the procedure of achieving the relative backscatter strength (BS) introduced in the previous chapter, we explore the potential of mapping marine benthos through BS classification on two multi-spectral MBES datasets (90, 300, and 450 kHz) collected near the Wadden Sea islands in the North Sea. Two unsupervised classification methods that employ distinct strategies to balance the spatial and angular BS variations are compared. One is the Bayesian classification, which exploits BS data from individual incident angles. The other method leverages the BS angular response curve (ARC) at the cost of spatial data density, by classifying mean BS values from three different angular ranges with hierarchical clustering. Per-frequency classification results using both methods show a good correspondence with the sediment median grain size. Nevertheless, based on a principal component analysis (PCA) of all available bottom sample properties, hierarchical clustering results reveal a better distinction between contributions from the gravel content and sand mason worm density. Comparisons among multiple frequencies, on the other hand, show little difference regarding the relationship with bottom samples. Although the BS difference between frequencies might positively correlate with certain sample properties, multi-spectral analysis with the employed frequency range in this study brings limited additional information compared to the single-frequency classification.

This chapter has been published in Frontiers in Earth Science 11 (2023), p. 1140649. [96]

3.1. Introduction

Knowledge regarding the occurrence of marine benthos is essential for the assessment and conservation of seafloor habitats. Traditional biodiversity monitoring relies on bottom sampling, providing accurate point measurements of the seafloor. However, such methods achieve only sparsely distributed information and are time-consuming, especially in the analysis phase. In recent years, acoustic remote sensing technologies such as the single-beam echosounder (SBES), side-scan sonar (SSS), and multi-beam echosounder (MBES), have been extensively used for large-scale surveying, producing high-resolution maps [97]. In particular, the MBES has drawn widespread attention in seafloor characterization due to its advantage of simultaneously acquiring bathymetry and backscatter data. By emitting the acoustic signal and beamforming during reception, an MBES collects data over a swath, covering a wide range of incident angles. The backscatter strength (BS), which is derived from the received echo intensity, is influenced by the frequency, incident angle, and sediment properties such as interface roughness and volume heterogeneity [98]. Given an acoustic frequency, the angular dependency of BS is also an intrinsic property of the seafloor [16, 99].

The composition of seafloor substrates can indicate the associated biological communities. On the other hand, the occurrence and bioturbation activity of marine benthos also affect the sediment properties, which has an impact on the backscattering process of acoustic signals as a result [100]. Many regional studies have pointed out the potential correlation between backscatter data, sediment properties, and benthic macrofauna abundance [64, 101]. Using least squares curve fitting and data from a forward-looking sonar, previous research modeled BS at a fixed incident angle as a function of several sediment properties such as gravel percentage [20]. It was also found that mud content and mean grain size can be the most important sediment properties affecting BS through modeling with a random forest decision tree [102]. Moreover, Hutin *et al.* [21] showed the potential of identifying scallop beds from BS using statistical discriminant analysis, but with the combination of SBES data and epi-macrobenthos photographs. Besides the empirical relationship discovered for different regions, a lot of effort has been put in developing quantitative sediment–backscatter models based on well-calibrated MBES backscatter data [92]. For example, the widely accepted APL-model [18] estimates BS at various incident angles for a given median grain size and frequency. Having the angular response of calibrated backscatter data, the classification of sediment types can be obtained through model inversion [30, 103]. However, the APL-model is only applicable to a limited range of frequencies up to 100 kHz and absolute backscatter calibration is not always easy to achieve.

Investigating differences in sediment properties among acoustically defined groups is another important approach in seafloor characterization, in which similar features, such as BS and its derivatives, are assigned to one of several classes. For this purpose, both supervised and unsupervised classification methods are commonly employed in seabed mapping [97]. Stephens and Diesing [77] compared six supervised classification techniques on features from bathymetry and backscatter mosaics, stressing the importance of selecting suitable features for mapping seabed substrates. Porskamp *et al.* [74] applied a random forest model on MBES data and showed the benefits of using multi-scale features from backscatter mosaics in benthic habitat mapping. Deep learning

models also showed the potential for sediment classification given the reliable labeled data [81, 104]. Compared to using backscatter mosaics, in which the *BS* angular dependence needs to be removed, previous research demonstrated the importance of angular response features for predicting benthic biota classes [53]. Alevizos and Greinert [105] also combined angular range analysis and multi-dimensional image processing by generating backscatter mosaics with different reference angles. Bayesian classification [80], by contrast, avoids angular normalization by classifying backscatter data at individual angles in an unsupervised manner, which was also adopted and demonstrated effective in multi-frequency backscatter classification [106]. Recent development of the multi-spectral MBES brings the opportunity to better distinguish seafloor substrates with nearly co-located multi-frequency data from a single survey [67, 69], either by merging single-frequency classification results [106] or by employing features from multiple frequencies as input for classification [62, 107]. At the same time, a deeper knowledge of the signal penetration and bathymetric uncertainty at different frequencies is required to ensure reliable interpretations [90].

Given the increasing demand for efficient methods of identifying marine benthos to support offshore human activities while ensuring sustainable use of the sea, the objective of this study is to investigate the potential of mapping the existence of marine benthos using multi-spectral backscatter data. To this end, two backscatter-based unsupervised classification methods, i.e., Bayesian classification and hierarchical clustering of the mean *BS* from different angular ranges, are applied to two datasets acquired by a multi-spectral MBES system in the North Sea. By comparing both methods, we investigated the possibility of achieving additional indications of benthos occurrence by involving the backscatter angular information. To accomplish this, we compared the seabed compositions indicated by bottom samples with backscatter classification results for different frequencies. In addition to the single-frequency analysis, an attempt to use multi-spectral features for acoustic classification was also made to assess the benefits of combining multiple frequencies for seabed characterization regarding our study areas.

3.2. Materials and methods

3.2.1. Study areas and datasets

In July and August 2021, we conducted two surveys north of the western Wadden Sea islands, hereafter referred to as “off Ameland and Terschelling” (AT), and “Borkumse Stenen” (BoS), respectively (Fig. 3.1). Both areas showed similar water depths, ranging from 22.1 to 29.5 m. There was a trough with gradual slopes in Area AT, whereas Area BoS was shaped around a steep drop in bathymetry, with gradual bathymetric changes in the shallower western region and a more heterogeneous region in the east. During the surveys, a multi-spectral MBES system R2Sonic 2026 was used to collect MBES data at four frequencies, 90, 200, 300, and 450 kHz, with beam opening angles of 2.3°, 1.1°, 0.7°, and 0.5°, respectively. A swath coverage of 130° was adopted. The system settings of the MBES were kept the same during data acquisition in July and August 2021. In this study, we first processed the MBES data in the QPS software Qimera for bathymetric cleaning and converted them into the generic sensor format (GSF). Afterward, data were separated into different frequencies. For each acoustic frequency, we derived the

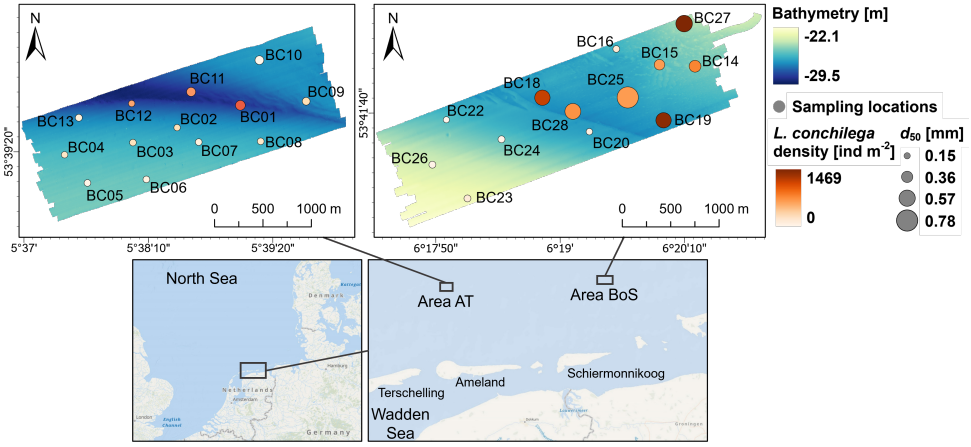


Figure 3.1: Study areas in the North Sea, shown with bathymetry, sampling locations, median grain size (d_{50}), and the densities of sand mason worm (*Lanice conchilega*) in units of individuals per m^2. Area AT is located north of the islands of Ameland and Terschelling, while Area BoS (Borkumse Stenen) is in the Borkum Reef Ground.

beam-averaged BS in dB per m^2 at 1 m which can represent the seafloor properties from the received echo level via backscatter correction (see Section 2.2.1).

In both Area AT and BoS, 13 sampling locations were selected for the sediment and macrofauna analysis. At each location, we took 3 bottom samples using a 0.078 m^2 boxcore. Two replicates were taken for macrofauna analysis and a third for sediment. The macrofauna replicates were sieved on board through a 1 mm mesh to extract the fauna from the sediment. Fauna were stored in a 4% formalin solution before analysis in the lab. All animals were counted, but only "acoustically hard" animals (such as molluscs or the sand mason worm *Lanice conchilega*) were identified down to species level, other groups only to class or order level. Sediment was analyzed by separating dry weight sieve fractions with mesh sizes of 63, 125, 250, 500, 1000, 2000, and 4000 μm. These fractions were used to determine properties such as median grain size (d_{50}) and weight percentages of gravel, sand, and mud. In addition, volume percentages of shell fragments, living bivalves (the most abundant hard-shell animal), and stones (> 4 mm) were determined.

3.2.2. Bayesian classification

Considering that Area BoS was surveyed only one month after Area AT and the sonar settings were kept unchanged, backscatter data from both areas were combined for classification in this study. By doing so, we can relate the classification results to a larger number of samples and link the sediment properties of the two surveyed areas.

The Bayesian classification method used in this study was developed by Simons and Snellen [80]. In Bayesian classification, the beam-averaged BS is regarded as a random variable, which varies only with the seafloor properties for a given frequency and incident angle. As a result of the averaging over many independent scatter pixels within a beam,

the obtained BS per beam is assumed to follow a Gaussian distribution according to the central limit theorem. Here, a scatter pixel indicates the signal footprint, which is the ensonified extent of an acoustic pulse on the seafloor (see Fig. 2.4). If there are m sediment types in the surveyed area, the histogram of BS at one incident angle can be modeled as a summation of m Gaussian distributions as follows:

$$h(BS_j|\mathbf{x}) = \sum_{k=1}^{m} w_k \exp\left[-\frac{(BS_j - \overline{BS}_k)^2}{2\sigma_k^2}\right], \tag{3.1}$$

where BS_j represents the backscatter strength of the jth bin in the histogram and $\mathbf{x}$ refers to all $3m$ unknown parameters $(\overline{BS}_1, ..., \overline{BS}_m, \sigma_1, ..., \sigma_m, w_1, ..., w_m)$. $\overline{BS}_k$ and σ_k are the mean and standard deviation of BS for the kth sediment type. In addition, w_k is the weighting factor. The unknown parameters can be determined by solving the non-linear least squares problem

$$\min \sum_{j=1}^{M} \left[n_j - h(BS_j|\mathbf{x})\right]^2, \tag{3.2}$$

with n_j the number of observations in the jth bin and M the total number of bins in the histogram. In the end, the non-overlapping acceptance region of each sediment type (or Bayes class) is determined by BS values at the intersections of the m fitted Gaussian distributions according to the Bayesian decision rule for multiple (m) hypotheses, given that the *a priori* probabilities of all m sediment types are equal (Fig. 3.2).

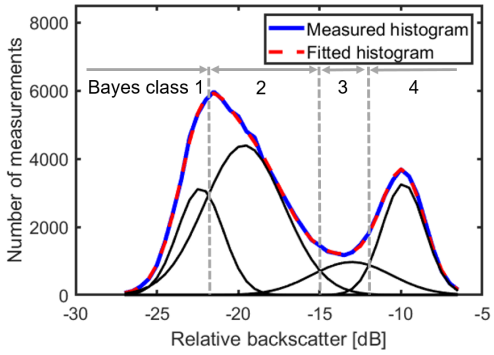


Figure 3.2: An example of acceptance regions for each Bayes class, with the number of sediment types $m = 4$. The measured and fitted histograms are also illustrated.

To achieve an optimal selection of m, i.e., the number of Bayes classes, we solved the above least squares problem for an increasing number of assumed sediment types, followed by a χ^2 goodness-of-fit criterion. The χ^2 statistic is calculated as

$$\chi^2 = \sum_{j=1}^{M} \frac{\left[n_j - h(BS_j|\mathbf{x})\right]^2}{n_j}, \tag{3.3}$$

assuming Poisson statistics for the random variable n_j. This χ^2-distributed statistic has $\nu = M - 3m$ degrees of freedom. The goodness-of-fit criterion is further defined as the reduced χ^2-statistic ($\chi^2_\nu = \chi^2/\nu$). If ν is large enough, χ^2_ν can also be well approximated by a Gaussian distribution with mean 1 and standard deviation $\sqrt{2/\nu}$. Therefore, in our implementation, χ^2_ν was calculated for each m for a series of incident angles. The iteration of searching for an optimal m stopped when the value of χ^2_ν was close to $1 \pm \sqrt{2/\nu}$ and not improving anymore.

Beams from larger incident angles contain more scatter pixels, resulting in a Gaussian distribution with a lower standard deviation and thus a higher discrimination ability for sediment types. Considering the robustness of the classification, we first selected several reference incident angles between 40° and 60° based on their performance in the χ^2 goodness-of-fit and consistency in the fitted Gaussian distributions. Then the average percentage distribution of the acquired acceptance regions at these reference angles was applied to other incident angles. To ensure a similar number of scatter pixels for beams with different incident angles and water depths, we also averaged the *BS* over a few adjacent pings and beams [108]. The number of pings was 4 and 3 for the depth range 20–25 m and 25–30 m. The tolerance angle for averaging over beams was 5.0°, 3.2°, and 1.4° for incident angles 20°–25°, 25°–55°, and 55°–65°, respectively. Beams with an incident angle smaller than 20° were excluded in Bayesian classification, since they have too few scatter pixels and the central limit theorem cannot be satisfied.

3.2.3. Hierarchical clustering of mean *BS* from different angular ranges

After backscatter correction, the *BS* angular variation represents an intrinsic property of the seafloor. Although the random fluctuation of *BS* from several individual incident angles is accounted for in Bayesian classification, the angular information is neglected. Therefore, we applied a second classification method, in which we leveraged *BS* data from different angular ranges. We divided incident angles from 0° to 65° into three angular ranges, i.e., near-range (from 0° to 25°), far-range (from 25° to 55°), and outer-range (from 55° to 65°), which is similar to methods used in [30] and the angular range analysis in the Fledermaus Geocoder Toolbox (FMGT) by QPS. Moreover, the half-swath backscatter angular response on the port or starboard side was regarded as a unit for feature extraction and classification. *BS* values from each angular range were averaged, resulting in three features (near-mean, far-mean, and outer-mean) for a half-swath angular response. Also, an along-track averaging over 10 pings was adopted to reduce noise in the backscatter measurements. For each frequency, we constructed a covariance matrix from the three mean backscatter features, on which a principal component analysis (PCA) was then conducted. We classified data of each frequency separately by inputting the first principal component (PC), which can represent more than 90% of the total variance, to hierarchical clustering [109] with the Euclidean distance similarity metric and complete linkage algorithm.

The minimum number of clusters was based on the results of analyzing the Bayesian classification, since Bayes classes include the statistical fluctuations of *BS* measurements. However, as features from a wider angular range were included, it is also possible that the optimal number of clusters/classes is higher in hierarchical clustering. Therefore, for each frequency, starting from the number of Bayes classes, we successively

added another cluster during hierarchical clustering to see if a more discriminative classification can be achieved.

3.2.4. Investigation of classification results using sample properties

The relationship between seabed properties and acoustic classes, i.e., Bayes classes or hierarchical clusters, was investigated for 26 bottom samples taken in both Area AT and BoS, maximizing the coverage of different sediment types. Nine sample properties (d_{50}, *L. conchilega* density, total density of molluscs, weight percentages of gravel, sand, and mud, volume percentages of stones, dead shells, and living bivalves) were considered in the first place. However, we excluded properties showing very high collinearity with others to achieve a better representation of the sample space. We quantified the collinearity using the Variance Inflation Factor (VIF). VIF measures how much the variance of one regression coefficient in a multiple regression model is inflated due to multi-collinearity among predictor variables, and thus can be computed for each predictor. If the VIF of one predictor is 1, there exists no correlation between this predictor and the remaining predictors. VIF larger than 10 might indicate a serious collinearity problem, while VIF between 5 and 10 can be cause for concern [110]. Based on this, we selected the sample properties iteratively. The one with the highest VIF among 9 properties was first removed. Afterward, the VIF values were calculated again for the remaining 8 properties. The iteration stopped when the VIFs of all included properties were below 5.

On these selected properties for 26 samples in total, a PCA was performed to extract the primary information contained in the sample space. With the first and second PC as two axes, the samples and property vectors can be further displayed together by projecting them onto the span of those two PCs. In this *biplot*, similar samples will be aggregated together and correlations between properties can be visualized. Specifically, the high correlation is indicated by a small angle between two property vectors. This will be further explained in Section 3.3.3. In combination with the acoustic class corresponding to each sampling location, we are then able to inspect if the backscatter classification reveals differences in certain sample properties.

3.3. Results

3.3.1. Bayesian classification

Since classification results for the 200 kHz data are very similar to the other two higher frequencies, here we only present the results for 90, 300, and 450 kHz. According to the χ^2_ν values and the associated uncertainty, 6, 4, and 4 Bayes classes were selected for 90, 300, and 450 kHz, respectively (Fig. 3.3). Furthermore, several reference incident angles, i.e., 53° and 54° for 90 kHz, 43° and 47° for 300 kHz, and 43°, 45° and 51° for 450 kHz, were chosen for the classification of all incident angles between 20° and 65°. The classification process was also repeated for two more classes in each frequency, demonstrating that no additional information is brought to the classification maps by adding more Gaussians.

Compared to the backscatter mosaics (Fig. 3.4), the ascending Bayes class number corresponds to an increase in BS (Fig. 3.5). For all frequencies, the highest class occurs

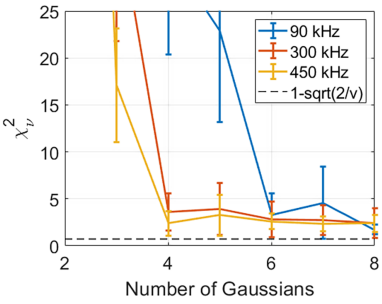


Figure 3.3: χ^2_ν values averaged over incident angles from 40° and 60° as a function of the number of Gaussians for the frequencies 90, 300, and 450 kHz.

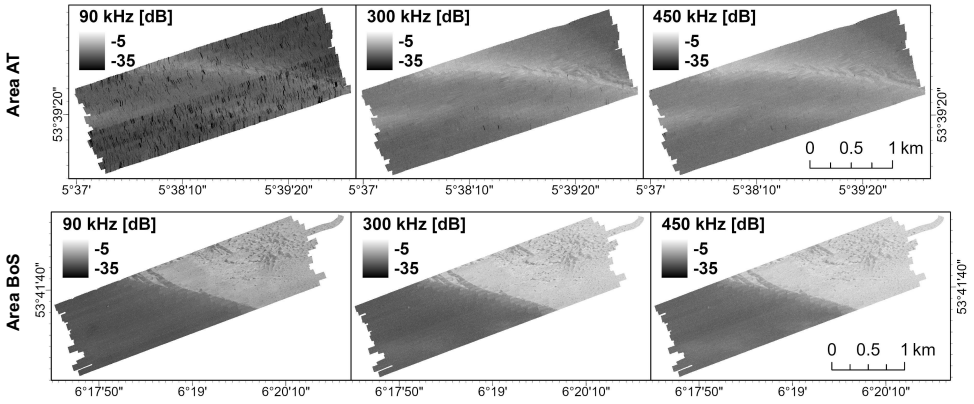


Figure 3.4: Backscatter mosaics (90, 300, and 450 kHz) of both study areas, with the angular variation normalized by BS between 40° and 60°.

in locations with deeper bathymetry, i.e., the trough of Area AT and the eastern seabed of Area BoS. In AT, regions near the sampling location BC04 also fall into a relatively high Bayes class, indicating a pattern that is significant in backscatter data but not found in the bathymetry. Compared to 300 and 450 kHz, Bayesian classification of 90 kHz makes additional distinctions for low BS values, adding extra classes. As for Area AT, although the coverage of classes 1, 2, and 3 in 90 kHz generally align with class 1 and 2 in higher frequencies, they have a noisy spatial distribution, likely due to artifacts in the backscatter data caused by rough weather conditions during the measurement. In Area BoS, a small patch of class 2 of 90 kHz near the sampling location BC26 is not observed in other frequencies. Moreover, differences among frequencies exist in the second highest Bayes class. In contrast with higher frequencies, class 5 in 90 kHz is less identified near the trough in Area AT, but occurs slightly more on the eastern side of Area BoS. However, class 3 in 300 and 450 kHz has a broader coverage on the western seabed near the bathymetric drop in the middle of Area BoS.

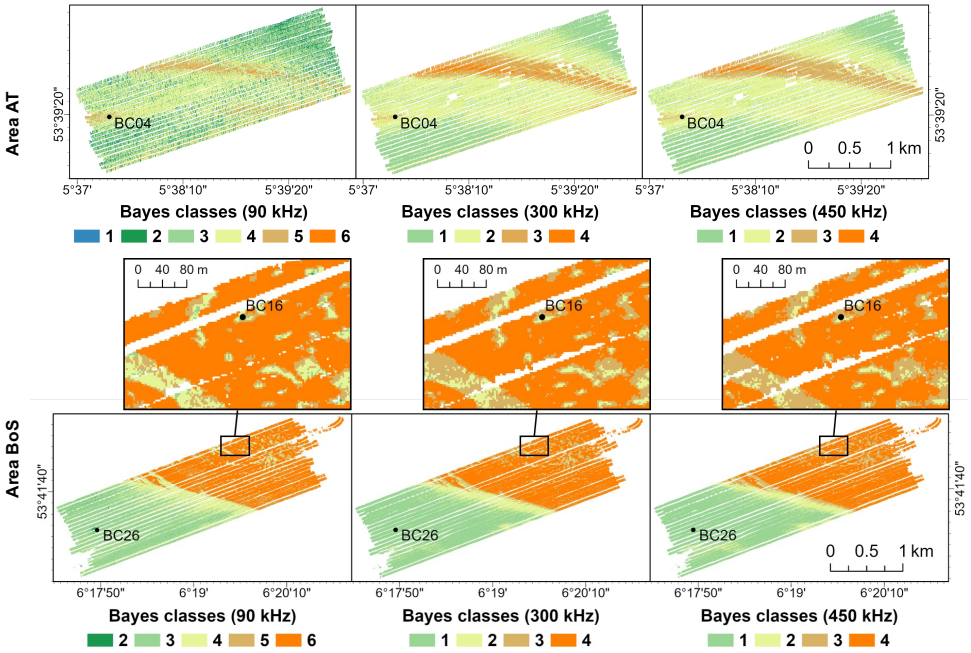


Figure 3.5: Bayesian classification maps for 90, 300, and 450 kHz.

3.3.2. Hierarchical clustering of mean *BS* from different angular ranges

Guided by the optimal number of classes in Bayesian classification, which was assumed to be the minimum number of classes, 7, 5, and 6 clusters were selected in hierarchical clustering for 90, 300, and 450 kHz, respectively. The classification maps of hierarchical clustering (Fig. 3.6) show a good correspondence with Bayes classes. Especially for 300 kHz, clusters from 1 to 4 indicate similar regions as Bayes classes. Furthermore, cluster 5 reveals a new spatial pattern in Area BoS compared to the Bayesian classification results, corresponding to regions with the highest mean *BS* near the nadir and in the far-range. Such additional information is also found in the hierarchical clusters of 90 and 450 kHz. Moreover, the highest clusters, i.e., cluster 7, 5, and 6 in 90, 300, and 450 kHz, respectively, are not observed in Area AT, which is likely due to a wider range of d_{50} in Area BoS (see Section 3.3.3). Cluster 3 in 90 kHz and cluster 1 in 450 kHz are not present in Bayesian classification either. In Area BoS, both clusters correspond to the low *BS* values in the western region, but might be affected by the striped artifacts in the nadir backscatter data. The clustering results of Area AT for 90 kHz are also impacted by rough weather backscatter artifacts, showing several noisy classes. Furthermore, since feature extraction for hierarchical clustering was based on data from half of the surveyed swath, the spatial resolution of the generated classification maps is limited, and across-track seabed variations cannot be accounted for as in Bayesian classification. Considering regions around the sampling location BC16 in Area BoS, heterogeneous backscatter features in small scales (10–20 m) are discriminated by Bayesian classification, but are

not resolved by hierarchical clustering (see zoomed-in plots in Fig. 3.5 and 3.6).

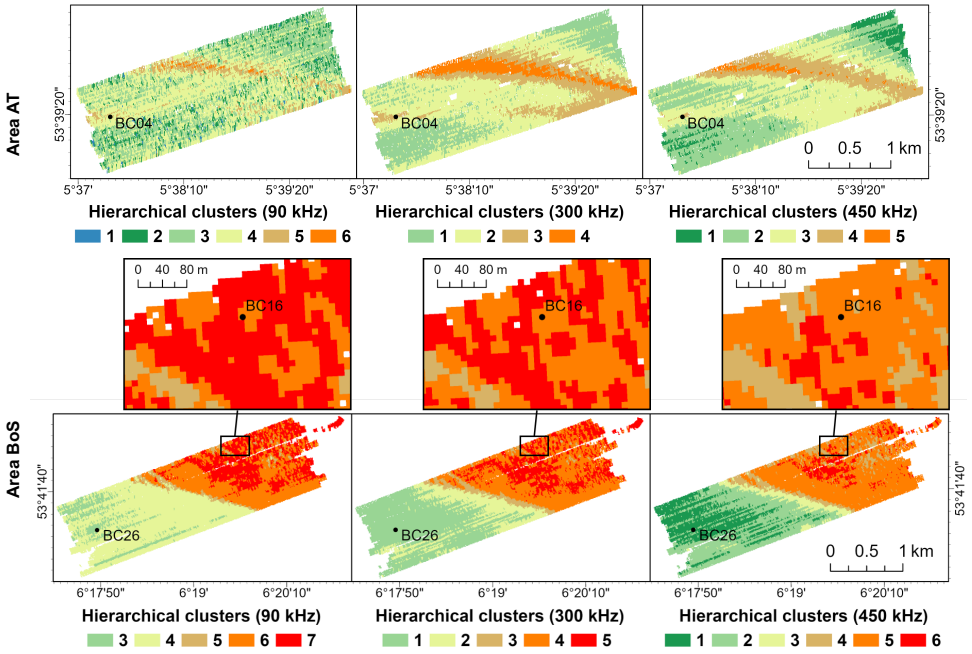


Figure 3.6: Classification maps of hierarchical clustering of the mean *BS* from three angular ranges for 90, 300, and 450 kHz

3.3.3. Investigation of classification results using sample properties

The sample analysis results show that d_{50} in Area AT ranged from 0.16 to 0.30 mm, indicating a relatively homogeneous composition from fine sand to medium sand. By contrast, Area BoS had a wider range of sediment types from fine to gravelly sand. For both classification methods, the ascending acoustic classes generally correspond to an increase in d_{50} (Fig. 3.7). Note that near the sampling location BC16, the sediment showed a small median grain size, but is classified as the highest acoustic class within 50 m for every frequency, possibly due to a high spatial variation of sediment types in that region. According to the Bayesian classification results, BC16 might be located on a small-scale region with softer sediment, while coarser sediment existed in its surroundings (see Fig. 3.5).

Furthermore, following the process of property selection using the VIF values, *L. conchilega* density, and the percentages of gravel, sand, mud, dead shells, and living bivalves were finally selected for PCA in the sample space. In PCA, the first two PCs account for nearly 70% of the bottom sample data variability. Moreover, most sample properties, including the percentages of gravel and dead shells and the density of *L. conchilega*, contribute to the first PC (Fig. 3.8). Furthermore, the *L. conchilega* density correlates positively with the percentages of gravel and dead shells, as indicated by the

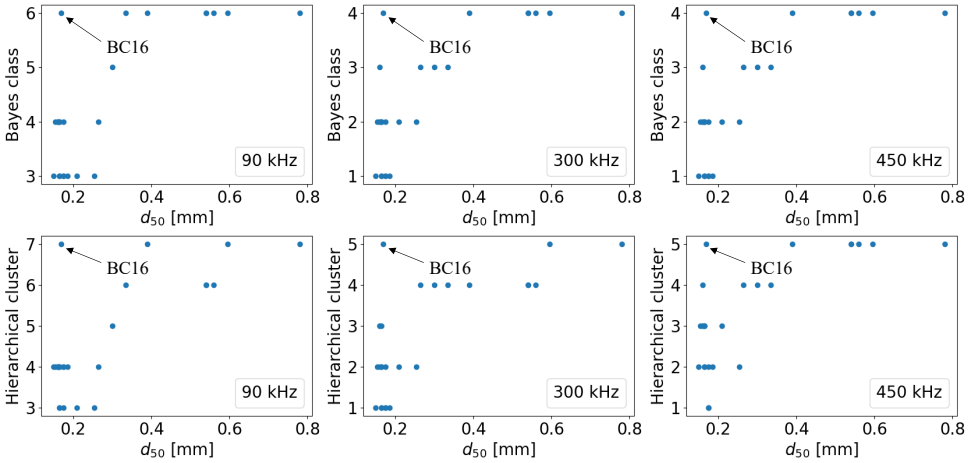


Figure 3.7: Scatter plots of the acoustic classes versus d_{50}. Each point represents a sampling location, whose classification is determined by the most frequent class within a radius of 50 m around it.

small angles between their property vectors on the biplot. By contrast, mud and sand percentages are almost uncorrelated considering that their property vectors are nearly perpendicular to each other. BC22 and BC27 show the least similarity to other samples due to their high percentages of mud and gravel, respectively.

In both the Bayesian classification and hierarchical clustering, the two lowest classes shown in the biplot correspond well to the close aggregation of samples with finer sediment and smaller *L. conchilega* density. However, only the results of 90 kHz show a clear distinction between the two lowest acoustic classes. Hierarchical cluster 3 in 90 kHz, including locations near BC26 in Area BoS, corresponds to a higher percentage of sand. Sample BC01 from Area AT and BC15 from Area BoS showed a similar d_{50} (around 0.3 mm) and are close in the biplot. Both sampling locations are categorized as the same class, except for 90 kHz. By contrast, although BC04 (from Area AT) and BC15 (from Area BoS) are even closer in the biplot, BC04 is classified as a lower acoustic class than BC15 for each method and frequency. These two sampling locations differ mostly in their d_{50}, a property not included in the PCA. Due to the high collinearity with other properties, d_{50} was not considered in the PCA, making its contribution unseen in the biplot.

Differences between Bayesian classification and hierarchical clustering are mostly visible in the biplots for 90 and 300 kHz. For both frequencies, locations in the bathymetrically heterogeneous region in the northeast of Area BoS (BC25 and BC27) are categorized as a higher acoustic class in hierarchical clustering compared to Bayesian classification, consistent with the divergence between the percentage of gravel and both *L. conchilega* density and the volume percentage of bivalves. Generally, hierarchical clusters indicate different contributions to *BS* from the gravel content and benthos better than Bayes classes. In the hierarchical cluster map of 450 kHz, a similar spatial pattern is observed for the highest acoustic class as in 90 and 300 kHz. However, due to a more variable distribution of the hierarchical clustering results near BC25 and BC27, both sampling locations do not fall into the highest hierarchical cluster and the differ-

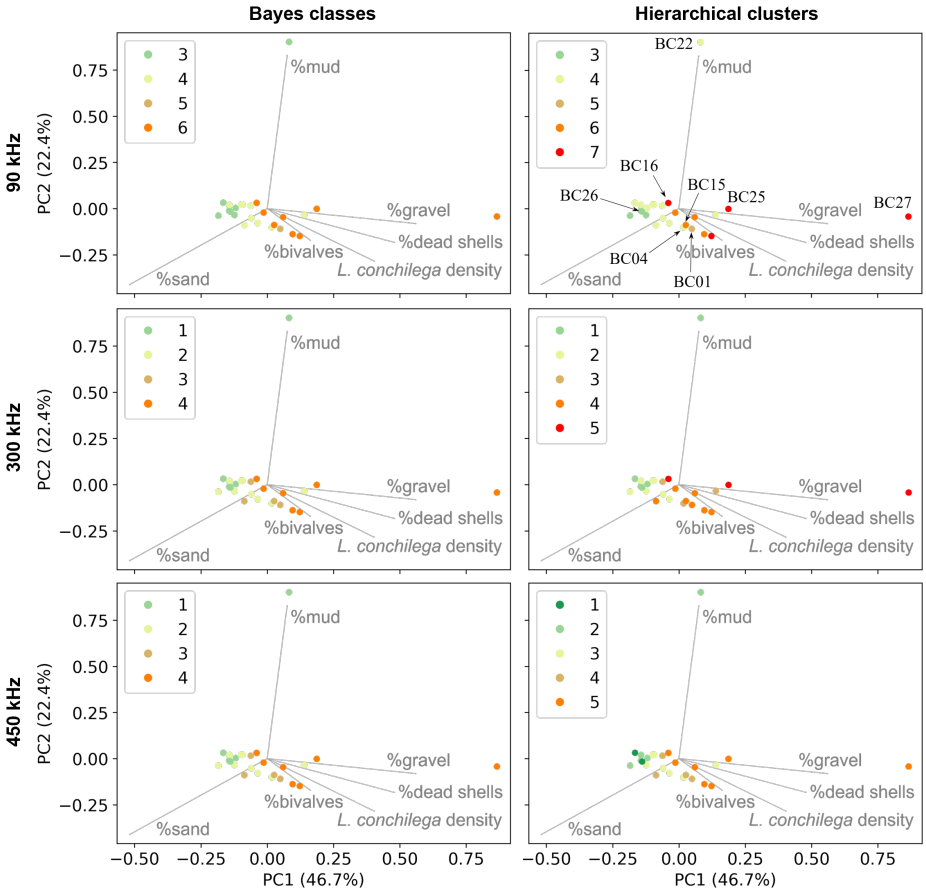


Figure 3.8: PCA biplots with Bayes classes or hierarchical clusters annotated for every sampling location, and the selected properties indicated as vectors. The most frequent acoustic class around each sampling location is indicated by the color of the points.

ence between the two classification methods in the biplot as mentioned above is absent for 450 kHz.

3.3.4. Use of multi-spectral features for acoustic classification

Besides the analysis of single-frequency classification results, it is observed that the backscatter difference between two frequencies can be an indicator of the variation in bottom sample properties, such as d_{50} and *L. conchilega* density. Considering the similar *BS* between 300 and 450 kHz, the backscatter difference between 90 and 300 kHz was taken as an example (Fig. 3.9). *BS* at 300 kHz is lower than at 90 kHz in the near-range, but higher in the outer-range, with the difference up to 4 dB. Far-mean in both frequencies show less difference. Compared to near-mean, the backscatter difference in far-mean and outer-mean show particularly a positive correlation with d_{50}

and *L. conchilega* density, indicating the potential of combining multiple frequencies for sediment classification.

Therefore, we also conducted hierarchical clustering on two multi-spectral feature sets, respectively. The first feature set contains mean *BS* from three angular ranges in both 90 and 300 kHz, while the second set includes only the backscatter difference between 90 and 300 kHz in far- and outer-range. The classification in both cases shows generally the same results as hierarchical clustering for single frequencies, indicating very limited additional information brought by using multi-spectral features in acoustic classification regarding the MBES data in this study.

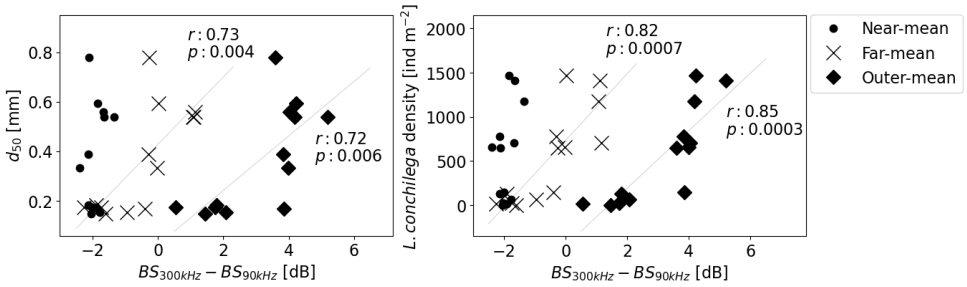


Figure 3.9: Scatter plots of d_{50} and *L. conchilega* density as a function of the difference between 300 kHz and 90 kHz regarding the mean *BS* from three angular ranges for sampling locations in Area BoS. For far-mean and outer-mean, the correlation between the backscatter difference and sample properties is described by the Pearson correlation coefficient r and the associated p-value. Backscatter differences from Area AT are not included considering the artifacts at 90 kHz.

3.4. Discussion and conclusion

We applied two backscatter-based classification methods, i.e., Bayesian classification and hierarchical clustering of mean *BS* from three angular ranges, to two multi-spectral MBES datasets collected near the western Wadden Sea islands in the North Sea. As an approach extensively used by previous research on various multi-beam backscatter data [80, 86, 106, 111], Bayesian classification accounts for *BS* statistical fluctuations at a certain incident angle and also proves effective in describing changes of the sediment composition in this study. According to a comprehensive analysis of the sample space through PCA, the resulting Bayes classes generally align with the distribution of sample properties. Especially along the first PC in the biplots, the ascending order of Bayes classes can be linked to gravel content and shell fragments, which is consistent with the findings based on another relatively homogeneous sandy environment in the Ameland inlet close to Area AT [58]. However, such a relationship between acoustic data and sediment type might change for a more muddy or gravelly seabed [102]. Besides, the tube-building activity of *L. conchilega* is known to influence the geoacoustic properties of the sediment such as rigidity and porosity [23], allowing an assessment of their occurrence from the backscatter data [112]. In our study areas, regions with a significant density of *L. conchilega* are also determined as the highest Bayes class, potentially indicating the impact of this species on *BS* measurements.

Regarding the second classification method, the spatial distribution of hierarchical clusters shows a good correspondence with Bayes classes in general. Compared to Bayesian classification, in which only several beam angles around 50° are used as the reference, additional information, e.g., the highest hierarchical cluster, is obtained by considering mean *BS* from different angular ranges in a hierarchical clustering algorithm. With this additional acoustic class shown in the biplots, the contributions of *L. conchilega* density and volume percentage of bivalves to *BS* are better distinguished from the gravel content and dead shells. Information from the whole angular response curve has also been demonstrated to be critical in determining the difference in macrofauna abundance by previous research [38]. Hasan *et al.* [53] stressed the importance of using the mean *BS* from 30° to 50° when classifying biota including various types of algae and invertebrates. The low number of sampling stations in regions with coarser sediment and higher macrofauna density in our measurements, on the other hand, might limit the ability to predict biological communities from backscatter data [100]. Moreover, sediment properties of a local region revealed by bottom samples sometimes do not align with the spatial patterns of acoustic classes in hierarchical clustering results. To account for features from an angular response curve, the finer-scale across-track discrimination is lost in our hierarchical cluster maps due to a limited spatial resolution of the MBES half-swath width.

Among frequencies, differences in the relationship between acoustic classes and the sediment composition are not clearly observed in this study, although small differences can be noticed in the classification maps. Previous research showed a better correlation between *BS* and mean grain size at 30 kHz than 300 kHz regarding a seafloor containing sediment from sand to gravel [113]. However, for our study areas with mostly sand, this contrast is not observed for 90 and 300 kHz. By further using the backscatter difference at 90 and 300 kHz as features, a combination of multiple frequencies in acoustic classification do not reveal remarkable differences compared to the single-frequency classification results in this study either. For future studies of such a sandy environment, frequencies lower than 90 kHz can be considered to better interpret the coarse sediment distribution and the subsurface properties. Moreover, employing multi-spectral features is still not trivial as it requires determining whether or not the same sediment layer is detected by acoustic signals from multiple frequencies. Investigation for the penetration depth also requires a better understanding of the bathymetry difference between multiple frequencies and the associated bathymetry uncertainties [114]. Considering the underlying link between the seafloor morphology and biological communities [49, 115, 116] , this can also help to increase the potential of using bathymetry data and its derivatives for characterizing the benthos distribution and improve the interpretation of acoustic classification results.

In summary, although limitations in using the multi-spectral MBES data exist in this study, the approach of linking backscatter classification results to sample properties reveals the potential for mapping the distribution of marine benthos using acoustic measures, especially when taking the *BS* angular variation into account.

4

Detecting benthic habitat changes with multibeam backscatter using a constrained Gaussian Mixture Model

Where the stream ends, I sit and watch the clouds rise.

Wang Wei (699—761)

Following the North Sea measurements described in the previous chapter, we conducted another survey in the same study areas after nearly three years to inspect potential changes in seabed habitats. However, developing robust change detection models with multibeam echosounder (MBES) data is challenging due to the high costs and limited spatial coverage of seabed ground truth data. Lack of absolute backscatter calibration also hinders the comparison between repeated MBES backscatter measurements. To mitigate these issues, we propose an unsupervised method in this chapter to detect seabed changes, by fitting a Gaussian Mixture Model (GMM) to the backscatter difference between two datasets. To eliminate the impact of possible drifts in echosounder characteristics on the backscatter difference, a relative calibration can be conducted based on a stable reference area. We then model the unchanged class as a zero-mean Gaussian distribution, with its variance constrained by the backscatter uncertainty estimated from the reference area. The change detection results reveal seasonal and temporal variations in benthic communities, such as sand mason worms, and are consistent with the sediment movement in one of the study areas. Moreover, by processing each incident angle individually, it is shown that MBES backscatter from the angular range 40°–60° might have the largest ability for seabed change detection.

This chapter has been published in Applied Acoustics 246 (2026), p. 111222. [117]

4.1. Introduction

Regular seabed monitoring is important for preserving and managing marine habitats [14, 97]. Traditional ecological monitoring techniques, such as box coring and trawling, provide precise but spatially sparse seabed information. These techniques also disturb the benthic habitats, which restricts the locations and frequency of sampling [8]. By contrast, acoustic techniques, especially the multibeam echosounder (MBES), offer a non-destructive and efficient approach for seabed characterization by simultaneously collecting bathymetry and backscatter measurements.

Previous research has successfully employed MBES backscatter data to characterize sea bottom properties, such as the sediment grain size [118–120], the occurrence of marine benthos [121, 122], and manmade structures [35, 62]. Considering the ability of MBES for continuous and broad-scale seabed mapping across various applications, repeated MBES surveys provide a promising solution for inspecting seabed changes [41].

Generally speaking, change detection from multi-temporal remote sensing datasets has been a prominent research topic for decades [123]. Given the ground truth data, supervised methods including machine learning and deep learning approaches can be applied for change detection. One common supervised approach is the post-classification comparison, where two classified images are compared to generate a change matrix [124]. Similarly, Montereale-Gavazzi *et al.* [125] classified MBES data from six surveys using the Random Forest algorithm. In their study, comparisons of the classification maps revealed temporal changes of three seabed sediment types over a ten-year period. However, collecting ground truth data for environmental monitoring can be time-consuming and labor-intensive. In seabed mapping applications, the aquatic environment poses additional challenges for accessing the seabed and achieving ground truth samples with accurate localization [58]. In this regard, unsupervised change detection methods, which minimize the dependence on ground truth data, offer a valuable alternative for monitoring dynamic seafloor environments.

Unsupervised classification methods have also been employed to detect seabed changes regarding the post-classification comparison. For example, Gaida *et al.* [58] identified changes in underwater coastal nourishment materials from eight surveys by comparing the unsupervised Bayesian classification [80] results of the MBES backscatter data. To avoid misinterpreting the possible drifts in sonar system characteristics as the seabed changes, they classified the backscatter data from each survey separately. However, some overall changes in the study area might be overlooked when comparing patterns of the separately classified maps without ground truth data from each survey. Feature differencing, on the other hand, offers a more straightforward approach for comparing datasets acquired at different times. Analyzing the difference between two co-registered datasets is a widely adopted method for change detection [126]. Seabed monitoring using the bathymetric differences from repeated surveys has also been investigated [127]. Seabed change detection using the MBES backscatter differences, however, remains an area that requires further research.

Analyzing the backscatter differences among distinct surveys presents difficulties mainly due to the absence of absolute backscatter calibration. Some sonar parameters such as directivity patterns can be proprietary and hence unavailable to users. Moreover, the aging of MBES components or changes in the seawater environment can also affect

the hardware sensitivity. In addition, the angular variation of the backscatter intensity introduces across-track inconsistencies of the backscatter data, which further complicates the process of feature differencing.

Building on the aforementioned gaps and challenges, this research aims to identify seabed habitat changes in an unsupervised manner by analyzing the backscatter differences between two MBES datasets. To make the two backscatter datasets comparable, we can conduct a relative calibration using a stable reference area, which has been indicated as an effective approach for making the temporal backscatter data intrinsically relative [70, 128]. Following this, the seabed changes are analyzed for each incident angle separately by clustering the backscatter differences using a Gaussian Mixture Model (GMM). Gaussian distribution is a widely used assumption for change detection [129]. Regarding the MBES backscatter data, statistical properties of the beam-averaged backscatter strength (BS) also allow the use of Gaussian distributions. To ensure the separation between unchanged and changed seabed areas, constraints are introduced to the GMM: the unchanged cluster is modeled as a zero-mean Gaussian, where its variance is initialized by the uncertainty estimated from backscatter data within the reference area.

We applied the proposed change detection method to repeated MBES datasets acquired in two study areas in the Dutch North Sea. Our detection results align with the observed geomorphological changes in seabed sediments and suggest a potential reduction in living benthos such as *Lanice conchilega* (sand mason worms) between two measurement periods, showing that the proposed method can be a cost-effective solution for monitoring the benthic habitat. The constrained GMM method also reveals the difference in change detection performance regarding different incident angles, facilitating the design of future surveys for acoustic seabed mapping.

4.2. Methods

The workflow of our seabed change detection method can be divided into three steps: data pre-processing (Section 2.2), backscatter difference preparation, and change detection using the constrained GMM (see Fig. 4.1). Details of the second and third steps are presented as follows.

4.2.1. Backscatter difference preparation

The difference in BS between two MBES datasets collected at time t_1 and t_2 for the same study area is the input to our seabed change detection method. Considering the angular dependence of BS, we analyze BS for each incident angle ϕ separately. To avoid including the change in sonar characteristics in the backscatter difference, $\Delta BS(\phi)$, we first conduct a relative calibration on $BS(\phi, t_2)$ using a reference area. The general requirements of the reference area are introduced in Section 2.2.2. Since such an independent reference area was inaccessible during our surveys, we selected a homogeneous and stable region within our study area (see Section 4.4.2).

Afterward, for each incident angle ϕ, we compute the average BS within the reference area from t_1 and t_2, denoted as $\mu(\phi, t_1)$ and $\mu(\phi, t_2)$. The relative calibration term C is then determined as

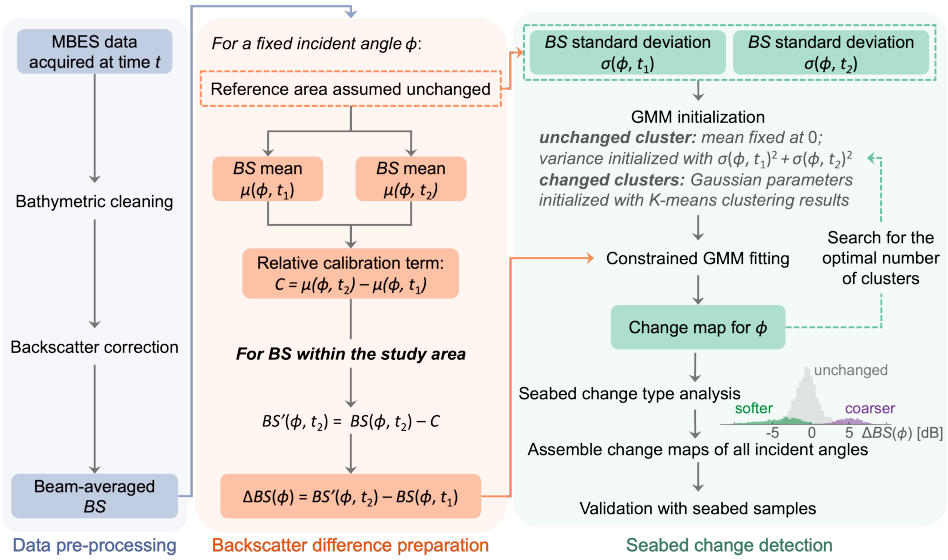


Figure 4.1: Workflow for seabed change detection using MBES backscatter measurements from two surveys.

$$C(\phi) = \mu(\phi, t_2) - \mu(\phi, t_1). \tag{4.1}$$

If $BS(\phi, t_2)$ represents any BS value in the study area for ϕ at t_2, the relative calibration can be achieved by

$$BS'(\phi, t_2) = BS(\phi, t_2) - C(\phi), \tag{4.2}$$

which accounts for the MBES systematic difference between t_1 and t_2. Following this, we achieve the corrected backscatter difference as input to the change detection algorithm through

$$\Delta BS(\phi) = BS'(\phi, t_2) - BS(\phi, t_1). \tag{4.3}$$

4.2.2. Seabed change detection: constrained GMM

We propose to detect seabed changes by clustering $\Delta BS(\phi)$ using the GMM, since the nature of the beam-averaged backscatter strength allows us to model $\Delta BS(\phi)$ as Gaussian distributions. Given the ping-by-ping emission of an MBES, $BS(\phi)$ can be treated as a random variable affected by the seabed geoacoustic properties. As introduced in the previous chapter, when a sufficient number of scatter pixels are present in a beam, $BS(\phi)$ from a single seabed type follows a Gaussian distribution according to the central limit theorem.

Therefore, given a fixed incident angle ϕ, the overall probability density function (PDF) of BS at time t_1 can be represented by a Gaussian mixture as

$$f_1(x) = \sum_{i=1}^{m_1} \frac{w_{1,i}}{\sqrt{2\pi\sigma_{1,i}^2}} \exp\left[-\frac{(x-\mu_{1,i})^2}{2\sigma_{1,i}^2}\right], \tag{4.4}$$

where x is one observation from $BS(\phi, t_1)$, m_1 is the total number of seabed types present at time t_1, and $w_{1,i}$, $\mu_{1,i}$, and $\sigma_{1,i}$ are the weight coefficient, mean, and standard deviation of the i-th seabed type, respectively.

Similarly, for the same study area and ϕ, the PDF of BS at time t_2 can be written as

$$f_2(x) = \sum_{j=1}^{m_2} \frac{w_{2,j}}{\sqrt{2\pi\sigma_{2,j}^2}} \exp\left[-\frac{(x-\mu_{2,j})^2}{2\sigma_{2,j}^2}\right]. \tag{4.5}$$

For seabed change detection, assuming that MBES measurements at t_1 and t_2 are independent, the PDF of their difference $\Delta BS(\phi) = BS(\phi, t_2) - BS(\phi, t_1)$ can then be derived as

$$f_\Delta(x) = \sum_{k=1}^{m_1 m_2} \frac{\tilde{w}_{\Delta,k}}{\sqrt{2\pi\tilde{\sigma}_{\Delta,k}^2}} \exp\left[-\frac{(x-\tilde{\mu}_{\Delta,k})^2}{2\tilde{\sigma}_{\Delta,k}^2}\right]. \tag{4.6}$$

where $\tilde{w}_{\Delta,k} = w_{1,i} w_{2,j}$, $\tilde{\mu}_{\Delta,k} = \mu_{2,j} - \mu_{1,i}$, and $\tilde{\sigma}_{\Delta,k}^2 = \sigma_{1,i}^2 + \sigma_{2,j}^2$ for each pair (i, j). x becomes one value in $\Delta BS(\phi)$.

In theory, the difference between $f_2(x)$ and $f_1(x)$ yields $m_1 m_2$ components corresponding to all possible transitions between seabed types. In practice, only a subset of these components is physically meaningful since many transitions do not occur within the study area. Moreover, at each location, only a single type of bottom change can occur, which limits the number of relevant components. Thus, we approximate Eq. 4.6 by n components:

$$f_\Delta(x) = \sum_{k=1}^{n} \frac{w_{\Delta,k}}{\sqrt{2\pi\sigma_{\Delta,k}^2}} \exp\left[-\frac{(x-\mu_{\Delta,k})^2}{2\sigma_{\Delta,k}^2}\right], \tag{4.7}$$

in which the first component represents the unchanged seabed and components 2 to n represent different types of seabed changes. The weight coefficient $w_{\Delta,k}$, $k = 1, ..., n$, satisfies $\sum_{k=1}^{n} w_{\Delta,k} = 1$. $\mu_{\Delta,k}$ and $\sigma_{\Delta,k}$ are the mean and standard deviation of the k-th seabed change type, respectively. $w_{\Delta,k}$, $\mu_{\Delta,k}$, and $\sigma_{\Delta,k}$ are the unknown parameters to be searched through optimization.

Regarding the unchanged cluster, several constraints can be added to its Gaussian distribution. After appropriate backscatter correction and relative calibration, the mean of the unchanged cluster, μ_1, can be assumed to be 0. Theoretically, its variance $\sigma_{\Delta,1}^2$ is calculated as

$$\sigma_{\Delta,1}^2 = \sigma_u^2(t_1) + \sigma_u^2(t_2), \tag{4.8}$$

with $\sigma_u^2(t_1)$ and $\sigma_u^2(t_2)$ the uncertainty of $BS(\phi)$ measurements from the two distinct surveys, respectively. Since the random fluctuation of BS is affected by the seabed types,

it is difficult to determine $\sigma_u^2(t_1)$ and $\sigma_u^2(t_2)$ without prior knowledge of the study area. Nevertheless, we can facilitate the search for $\sigma_{\Delta,1}$ using the pre-selected reference area.

Within the reference area, where the seabed is homogeneous, $BS(\phi)$ can be seen as repeated measurements of the same seabed type. Since the reference area remains stable over time and experiences no change measurable with an MBES between t_1 and t_2, we can estimate the variance of backscatter difference for this specific seabed type as

$$\sigma'^2_{\Delta,1} = \sigma_r^2(t_1) + \sigma_r^2(t_2), \tag{4.9}$$

with σ_r^2 the variance of $BS(\phi)$ from the reference area. We then use $\sigma'_{\Delta,1}$ to initialize $\sigma_{\Delta,1}$ during optimization.

Regarding the remaining unknown parameters, the weights w_k of different Gaussian components are initialized equally. In addition, $\mu_{\Delta,k}$ and $\sigma_{\Delta,k}$ of the changed clusters are initialized using the cluster mean and standard deviation from the K-means clustering [130] results. The K-means cluster whose mean is closest to zero is excluded for this. K-means clustering minimizes within-cluster data variances and might help separate GMM clusters. Following this, the unknown parameters in Eq. 4.7 are estimated by the Expectation-Maximization (EM) optimization algorithm [131].

4.2.3. Evaluation

We evaluate the change detection performance using the Silhouette coefficient [132], which ranges from -1 (poor fit) to 1 (well clustered). The Silhouette coefficient is used to compare the clustering results across different incident angles. It also helps to determine the optimal number of clusters/change types (n in Eq. 4.7).

We also validate the detected seabed changes by comparing with the ground truth bottom samples. We examine how predicted changes correspond to variations in sample properties such as sediment grain size and benthic community presence, thereby linking our detections to specific seabed conditions. For locations sampled during both survey times, we also compute the accuracy and F1-score for the binary "changed" and "unchanged" classification. The accuracy describes the proportion of correctly predicted samples, while the F1-score is calculated as

$$\text{F1} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{4.10}$$

where Precision measures the proportion of predicted "changed" locations that truly experienced changes, and recall is the proportion of actual "changed" samples correctly identified as "changed". Both accuracy and F1-score range from 0 (worst) to 1 (best).

4.3. Study Areas and Datasets

4.3.1. Study areas

To compare with MBES data collected in 2021 for the study areas "off Ameland and Terschelling" (AT) and "Borkumse Stenen" (BoS) (see Chapter 3), we conducted hydrographic surveys in the same region in 2024 (Fig. 4.2). In this subtidal region, intensive economic activities take place, including fishing, offshore wind energy generation, and

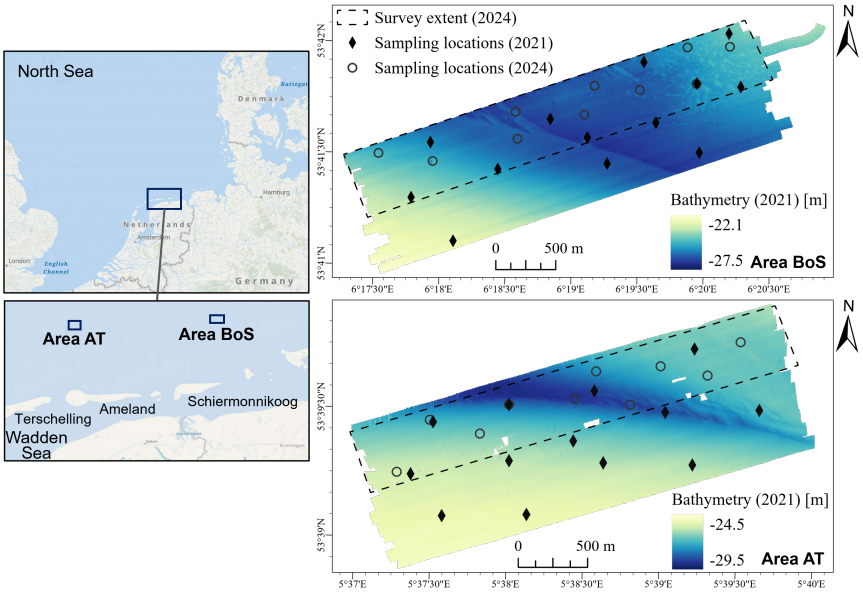


Figure 4.2: Study areas. (Left) Location of the study areas in the North Sea. (Right) Bathymetry measured in 2021, displayed with the seabed sampling locations in 2021 and 2024.

installation of cables and pipelines on the seabed [133]. Regular monitoring in this region is therefore essential for achieving a sustainable use of the seabed environment. Due to limitations in time and budget, the 2024 survey covered only half of the initial survey extent in each study area. In addition, a seasonal difference in the seawater temperature (≈10°C) can be noticed regarding surveys in the two years (see Table 4.1).

Table 4.1: Survey time and seawater temperature of the study areas

Study area	Survey time	Seawater temperature[1] [°C]
BoS	August 9–13, 2021	17.2±0.08
AT	July 12–16, 2021	16.7±0.06
BoS	March 13–14, 2024	7.3±0.14
AT	March 12–13, 2024	7.5±0.14

4.3.2. Acoustic datasets

During all surveys, we collected the acoustic data using the same pole-mounted MBES system, R2Sonic 2026, under the same sonar settings. Data with an acoustic frequency of 300 kHz are presented in this study. The beam opening angle was set to 0.7°, with a nominal pulse length of 150 μs. Under the equiangular beam spacing mode, each

[1]Mean±standard deviation for all water depths

ping contained 256 beams, covering a swath of 130°. In addition, adjacent track lines had approximately 50% overlap. To enable the comparison of $BS(\phi)$ between different surveys, we adopted the same sailing direction for the track lines in 2021 and 2024. During the measurements, we used the hydrographic data acquisition software Qinsy to store the MBES bathymetry and backscatter data as database (.db) files. Conductivity, temperature, and depth (CTD) measurements were collected regularly to assist the accurate backscatter correction for sound absorption through the water column.

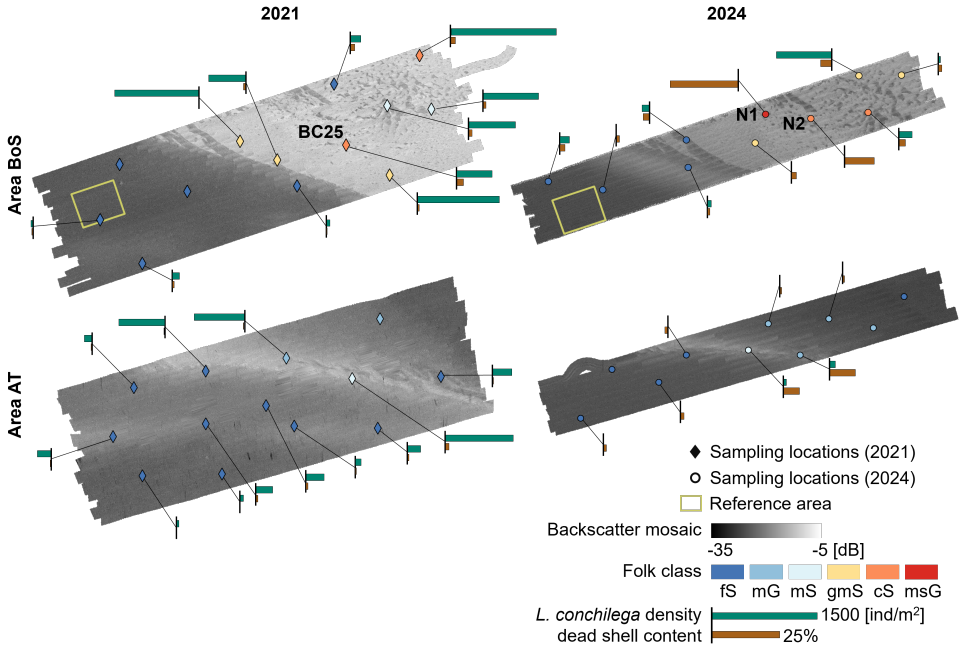


Figure 4.3: Folk class, *L. conchilega* (sand mason worm) density, and dead shell content for all boxcore samples collected in 2021 and 2024, displayed with the backscatter mosaic of each area. The backscatter mosaics were achieved using the software FMGT and only used for visual assessment. The 6 Folk classes correspond to an increase in the median grain size (fS: fine sand; mG: muddy gravel; mS: medium sand; gmS: gravelly muddy sand; cS: coarse sand; msG: muddy sandy gravel). *L. conchilega* density less than 20 [ind/m^2] or dead shell content smaller than 0.2% are not shown.

4.3.3. Seabed properties revealed by bottom samples

In 2021, 13 boxcore sampling locations were selected in Area BoS and AT for ground truthing. We used a cylindrical boxcore sampler with a surface area of 0.078 m^2. Pictures of the boxcore samples were taken onboard to provide a simple visual inspection. Macrofauna and sediment analysis were conducted later in the laboratory. In 2024, we selected 10 locations in BoS and AT for seabed sampling. Apart from boxcore samples, we took underwater videos above the seabed at each sampling location in 2024. Each video footage covered a transect of about 50 m, offering a qualitative assessment of the seabed conditions.

Sample analysis results from 2021 revealed a homogeneous seabed consisting of fine sand on the western side of Area BoS (Fig. 4.3). By contrast, the eastern seabed varied in sediment types from fine to coarse sand, with an abundant presence of *L. conchilega*, which are common in the subtidal sandy sediments of the North Sea [134]. *L. conchilega* are known for building tubes with sand grains and shell fragments. These tubes protrude outside the seabed surface, altering the seabed geoacoustic properties to some extent [112]. Moreover, at location BC25, many clay chunks were found underneath the seabed surface. Area AT, on the other hand, had a narrower variation in sediment types (fine to medium sand). *L. conchilega* were also found in AT, especially in the coarser sediment of the trough.

In 2024, the spatial distribution of different sediment types in both study areas remained largely consistent with 2021. However, a new sediment type (msG) was identified in Area BoS from one boxcore sample taken in a region unsampled in 2021 (see N1 in Fig. 4.3). This region contained a high dead shell content in 2024. Clay chunks were also found at location N2, which is close to the location BC25 sampled in 2021. For other sampling locations in 2024 on the eastern seabed of BoS, the *L. conchilega* density was smaller compared to 2021. This is possibly due to the influence of seasonality on mortality and recruitment of the worms, considering that the life cycle of *L. conchilega* follows a seasonal pattern. Adult worms that survive the storms in autumn and winter spawn in spring, and juveniles usually settle on the seabed in late spring and summer [135, 136]. In the central trough of Area AT, a lower density of *L. conchilega* and a higher percentage of dead shells were also found in spring 2024 than in summer 2021.

4.4. Results and Discussion

4.4.1. Relations between BS and seabed properties

To infer the type of seabed changes using $\Delta BS(\phi)$, relations between $BS(\phi)$ and seabed properties, such as the sediment grain size and benthos abundance, need to be examined first. In this way, we can link the increase or decrease in $BS(\phi)$ with certain seabed habitat changes. Fig. 4.4 presents the correlation analysis between seabed characteristics revealed by bottom samples and averaged BS from three angular ranges for our study areas. BS_1 from 0–25°, BS_2 from 25°–55°, and BS_3 from 55°–65° were achieved by averaging over 50 pings around each sampling location.

In general, BS_1, BS_2, and BS_3 are highly correlated, particularly in Area BoS surveyed in 2021. In 2021, *L. conchilega* in BoS were more abundant in coarse sediments (high Folk classes). The dead shell content also shows a moderate correlation with the Folk class. Moreover, the averaged BS across all three angular ranges correlate well with the Folk class. BS_2 and BS_3 exhibit moderate correlations with the dead shell content, and may be strongly affected by the *L. conchilega* density.

By contrast, BS_1, BS_2, and BS_3 show limited correlations with the *L. conchilega* density and dead shell content in the 2024 BoS survey. The relations between the Folk class and *L. conchilega* density also change a lot. Much fewer *L. conchilega* were found in most samples taken in 2024 than that in 2021. In addition, bottom samples from locations N1 and N2 (Fig. 4.3) contained sediments with larger grain sizes and higher dead shell content than those collected in 2021. The differences in sampling locations

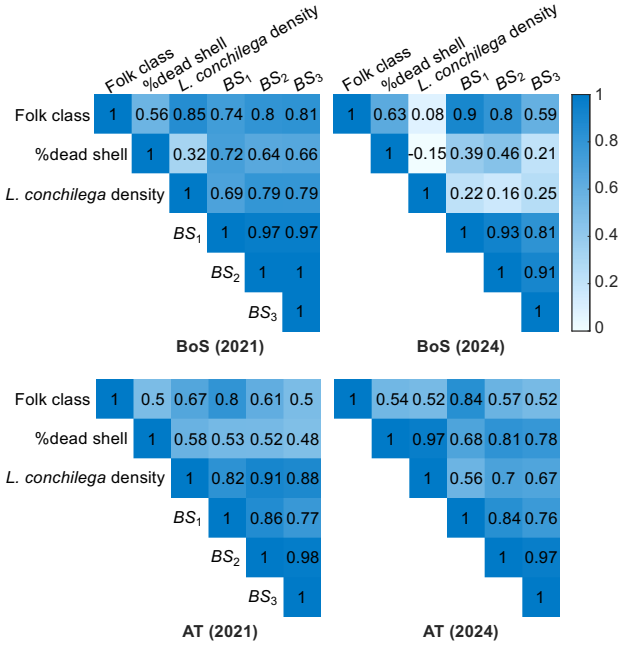


Figure 4.4: Correlation matrix for each study area between the boxcore sample properties and mean BS from three angular ranges, with BS_1 from incident angles 0–25°, BS_2 from 25°–55°, and BS_3 from 55°–65°. Folk classes in the study areas (fS–msG) are numbered from 1 to 6 when calculating the Pearson's correlation coefficient.

might also contribute to the change in the correlation between the Folk class and *L. conchilega* occurrence. The two seabed properties might share a non-linear relationship, considering the preference of *L. conchilega* for sediments containing fine to medium sand [112]. Moreover, coarse materials from N1 and N2, such as dead shells, can exceed the acoustic wavelength and may no longer increase BS. Despite these variations, BS_1 and BS_2 remained highly correlated with the Folk class.

In Area AT, there are apparent differences among BS from the three angular ranges regarding their relations with the seabed properties. For both surveys in 2021 and 2024, BS_1 exhibits the strongest correlation with the Folk class. For 2021 data, BS_1, BS_2, and BS_3 show similar moderate correlations with the dead shell content and high correlations with the *L. conchilega* density. In 2024, bottom samples contained significantly fewer *L. conchilega* and generally higher dead shell content than in 2021. Compared to the results for 2021, correlations between BS and the *L. conchilega* density are weaker for AT surveyed in 2024, while BS_2 and BS_3 are highly correlated with the dead shell content.

4.4.2. Relative calibration results

The selected reference area (Fig. 4.3) covers approximately 300 × 360 m of seabed. Within this area, the average absolute bathymetric difference between the two surveys in 2021 and 2024 is 0.08 m. To assess whether this difference is consistent with expected measure-

ment variability, we used the *A priori* Multibeam Uncertainty Simulation Tool (AMUST), which was developed to quantify the vertical uncertainty of MBES measurements [90, 137]. Based on the echosounder specifications given in Section 4.3.2, the depth uncertainty is estimated to be in the range of 0.074–0.088 m at the 95% confidence level regarding water depths in Area BoS. Although part of the observed 0.08 m difference may be attributed to systematic errors, which are not accounted for in this assessment, the close agreement with the expected random uncertainty suggests that the seabed within the reference area has remained effectively stable between the two surveys. This is also confirmed by the average seabed slope, which remains consistent at 0.7° for both years. Backscatter mosaics further support this stability by a visual assessment of the seabed homogeneity within the reference area. Furthermore, boxcore samples next to the reference area demonstrate stability in the sediment type (fS) and very little presence of benthos, although there is a slight increase in dead shell content (≈0.8%) from 2021 to 2024.

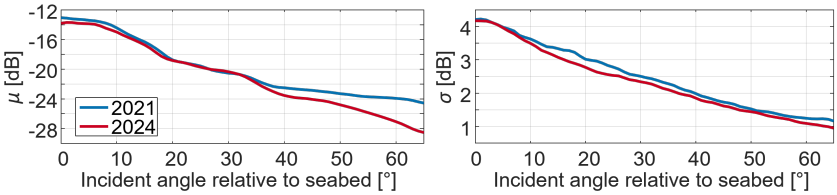


Figure 4.5: (Left) Average half-swath ARCs and (Right) standard deviations of *BS* from different incident angles within the reference area in 2021 and 2024. Incident angles with the same absolute value from the port and starboard side are considered identical, as the R2Sonic 2026 operates as a single-head system.

The averaged backscatter angular response curves (ARCs) from the reference area, $\mu(2021)$ and $\mu(2024)$, show consistency between the 2021 and 2024 surveys for incident angles 20°–30° (see Fig. 4.5). However, there exists an increasing difference between $\mu(2021)$ and $\mu(2024)$ for incident angles larger than 30°, reaching approximately 4 dB for angles above 60°, which is consistent with the finding from a recent study on the instrumental temperature-dependence of backscatter measurements [71]. Since the difference in sound absorption due to changes in seawater temperatures has been accounted for during the backscatter correction (Eq. 2.13), this effect might be attributed to the hardware sensitivity of the MBES system. In addition, we investigated the potential variation in the calculated two-way sound attenuation due to seawater absorption, by computing the difference between $2\alpha R$ in Eq. 2.14 using the minimum and maximum sound absorption coefficient (α) acquired for a water depth of 25 m. α was estimated based on *in-situ* CTD measurements during the survey period. Such difference does not exceed 0.3 dB from 30° to 65° (see Fig. 4.6), which is similar to values revealed by previous studies [58, 71] and is not significant compared to the 4 dB difference observed from $\mu(2021)$ and $\mu(2024)$. Variations in this maximum difference of estimated seawater absorption across surveys can be attributed to seasonal differences in daily temperature fluctuations as well as different numbers of CTD measurements due to varying survey durations.

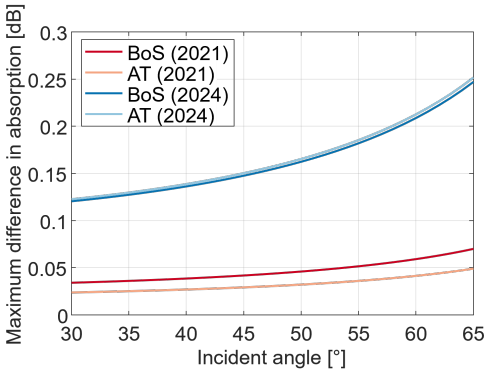


Figure 4.6: Difference between seawater absorption in dB calculated using the minimum and maximum sound absorption coefficient acquired for a water depth of 25 m, displayed for incident angles 30°–65°.

In this research, $\mu(2024)$ - $\mu(2021)$ was used as the relative calibration term for both Area BoS and AT, considering that the seawater temperature difference between the two study areas was small (Table 4.1). On the other hand, the standard deviations of BS within the reference area from the 2021 and 2024 surveys are generally consistent, with a decreasing trend as the incident angle increases (Fig. 4.5). This shows the stability regarding the uncertainty of MBES backscatter measurements from different surveys.

After applying the relative calibration, ΔBS within the reference area presents an example for backscatter difference of the unchanged seabed type. With the incident angle 55°, it is observed that the histogram of the resulting ΔBS can be approximated by a zero-mean Gaussian distribution, whose variance is calculated as $\sigma_r(2021)^2 + \sigma_r(2024)^2$ at 55°(Fig. 4.7). This is consistent with the assumption for constrained GMM.

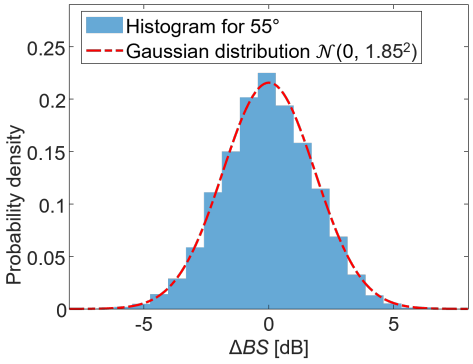


Figure 4.7: Histogram of $\Delta BS(55°)$ within the reference area after relative calibration, displayed with the zero-mean Gaussian distribution $\mathcal{N}(0, \sigma_r(55°, 2021)^2 + \sigma_r(55°, 2024)^2) = \mathcal{N}(0, 1.85^2)$.

4.4.3. Change detection: the optimal number of clusters

The Silhouette coefficient quantitatively evaluates the performance of unsupervised change detection and search for the optimal number of clusters (m in Eq. 4.7). We compare the Silhouette coefficients across five clustering methods: K-means, GMM with random initialization (GMM-r), constrained GMM with random initialization (CGMM-r), GMM with Gaussian parameters initialized by K-means clustering (GMM-k), and constrained GMM with K-means initialization (CGMM-k).

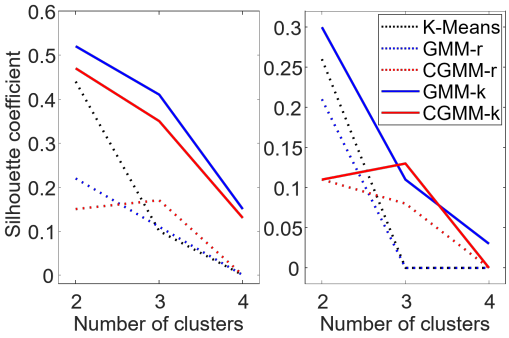


Figure 4.8: Comparison of the Silhouette coefficient from five clustering methods for Area (Left) BoS and (Right) AT, with the number of clusters ranging from 2 to 4. The presented Silhouette coefficient is an average value computed from the clustering results for incident angles 40°–60°. The five clustering methods are: K-means, GMM clustering with Gaussian parameters initialized randomly (GMM-r), GMM clustering with Gaussian parameters initialized by the K-means clustering results (GMM-k), constrained GMM clustering with random initialization (CGMM-r), and constrained GMM clustering with K-means initialization (CGMM-k).

To compare the performance across different numbers of clusters (Fig. 4.8), we computed the average Silhouette coefficient based on clustering results for incident angles between 40° and 60°. This angular range was selected because ΔBS, for this angular range, exhibits the largest capability to distinguish different seabed change types (see results in Section 4.4.4). For Area BoS, GMM-k and CGMM-k outperform the other methods, although their 4-cluster Silhouette coefficient decreases significantly compared to the results for two and three clusters. With only two clusters, K-means produces a Silhouette coefficient comparable to GMM-k and CGMM-k, but its performance deteriorates significantly for three and four clusters. In general, GMM-based methods can further improve the K-means clustering results by using them to initialize the Gaussian parameters. From GMM-r to GMM-k, the significant improvement in the Silhouette coefficient also highlights the effectiveness of K-means initialization in GMM optimization. For Area BoS, GMM-r and CGMM-r fail to achieve satisfactory clustering results. However, the superior performance of 3-cluster CGMM-r compared to K-means and GMM-r suggests the potential benefits of incorporating the proposed constraints.

Clustering for Area AT generally yields lower Silhouette coefficient values compared to results for BoS. There might be smaller differences among seabed change types in AT. With two clusters, GMM-k performs the best, though its improvement over K-means is not significant. CGMM-k shows the best performance for three clusters. GMM-k

and CGMM-r also show similar performance as CGMM-k regarding three clusters. By contrast, K-means and GMM-r fail to effectively distinguish three clusters. When using four clusters, the Silhouette coefficient for all methods approaches zero, indicating poor clustering performance.

We might choose 2 as the optimal number of clusters based on the Silhouette coefficient. However, some methods show little difference between the quantitative results of two and three clusters. The 3-cluster results might provide more information regarding a change type that differs slightly from the others. Thus, to ensure a comprehensive investigation, we also visually assess the cluster maps. For BoS, the 3-cluster results reveal an additional cluster with a significantly different spatial pattern compared to the 2-cluster results. As for AT, the 'unchanged' detections from CGMM methods are scarcely present in the 2-cluster results, with most of the study area falling into a single 'changed' cluster. When using three clusters in CGMM-r and CGMM-k, the spatial distribution of the 'unchanged' cluster remains consistent with the 2-cluster results. However, an additional 'changed' cluster with a different spatial pattern is revealed. Notably, the two 'changed' clusters also coincide with the 2-cluster results from the other methods, which might indicate that most areas in AT experienced changes. Based on these observations, we choose to use three clusters for both study areas.

4.4.4. Change detection: comparison among incident angles

With three clusters, all five methods achieve the best clustering performance for incident angles ranging from 40° to 60° (see Fig. 4.9). For Area BoS, all methods result in a Silhouette coefficient close to zero for nadir beams up to 20°. Between 20° and 40°, K-means, GMM-r, and CGMM-r have the Silhouette coefficient still near zero, while CGMM-k and GMM-k can improve the clustering results largely compared to the nadir beams. By contrast, for Area AT, most methods do not show clear separation for incident angles below 40°.

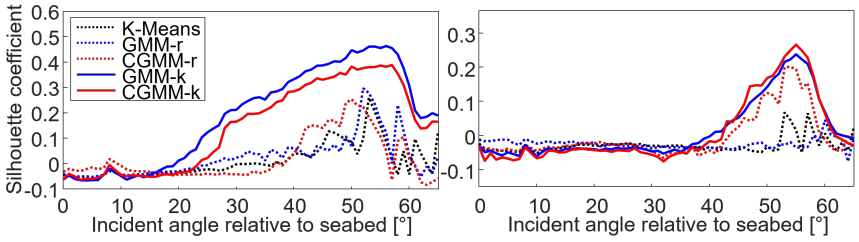


Figure 4.9: Silhouette coefficient of 3 clusters for all incident angles for Area (Left) BoS and (Right) AT. Comparisons among five clustering methods are presented.

Multiple factors contribute to the differences in clustering performances between nadir and oblique beams. Since *BS* from the nadir beams have a larger uncertainty (Fig. 4.5), distinguishing different seabed types becomes more challenging. Previous research [18] also suggested that a backscatter uncertainty of 2 dB is required to discriminate sediment types from fine to coarse sand, which is only met by the *BS* standard

deviation within the 40°–60° range. In addition, seabed roughness and volume properties mainly affect BS at oblique angles [16]. Seabed changes related to these properties may not induce nadir ΔBS values exceeding the uncertainty of the backscatter measurements. Regarding GMM methods, BS in the nadir can violate the assumption of Gaussian distributions due to the insufficient number of scatter pixels within one beam [80], which can subsequently deteriorate the clustering performance.

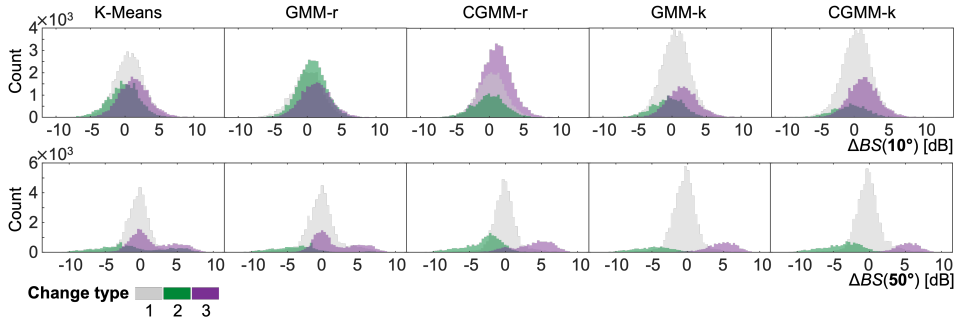


Figure 4.10: Histogram of ΔBS for 10° and 50° regarding Area BoS, colored by the clustering results. Change type 1 can be interpreted as "unchanged" for CGMM methods.

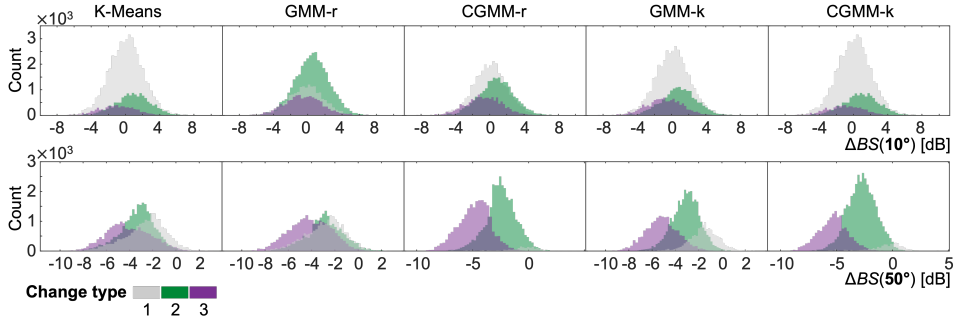


Figure 4.11: Histogram of ΔBS for 10° and 50° regarding Area AT, colored by the clustering results. Change type 1 can be interpreted as "unchanged" for CGMM methods.

Histograms of ΔBS at 10° and 50° provide a more detailed comparison of the clustering results for nadir and oblique angles (see Fig. 4.10 and 4.11). For both study areas, $\Delta BS(10^\circ)$ varies around 0 dB, with substantial overlap among the three clusters. By contrast, $\Delta BS(50^\circ)$ of Area BoS exhibits a much larger dispersion. Apart from one cluster centered at 0, cluster 2 with the negative backscatter difference is identified. Moreover, cluster 3 with $\Delta BS(50^\circ)$ around 5 dB is clearly separated by GMM-k and CGMM-k (Fig. 4.10). In contrast with BoS, $\Delta BS(50^\circ)$ of AT shows mostly negative values (Fig. 4.11). When using K-means or GMM-r, clusters 1 and 2, which are closer to 0, overlap significantly. GMM-k improves cluster separation but does not center cluster 1 around 0 dB. The proposed CGMM methods, on the other hand, ensure that cluster 1

accurately represents the “unchanged” seabed. Cluster 1 from CGMM-r and CGMM-k contains a very small number of data points, indicating that most of Area AT can be divided into the other two “changed” clusters. For both study areas, CGMM-k can simultaneously enable the identification of the “unchanged” seabed and maximize the separation among different clusters.

4.4.5. Seabed change type analysis

The 3-cluster change detection results from the CGMM methods distinguish the “unchanged” seabed and two “changed” clusters. Based on the relationship between BS and sample properties for our study areas (Fig. 4.4), we interpret the “changed” cluster as “coarser” if its mean ΔBS is larger than 0 and as “softer” if the cluster mean is negative. In addition, we validate the change detection results by comparing them with seabed change revealed by the ground truth bottom samples collected in 2021 and 2024.

Area BoS

All clustering methods reveal similar spatial patterns in the change maps for Area BoS (Fig. 4.12). Due to the unsatisfactory clustering performance of the nadir beams, only results for incident angles between 25° and 60° are presented. The change maps from K-means, GMM-r, and CGMM-r show a greater mixture of the three clusters compared to the other two methods, which aligns with the Silhouette coefficient results (see Fig. 4.8 and 4.9).

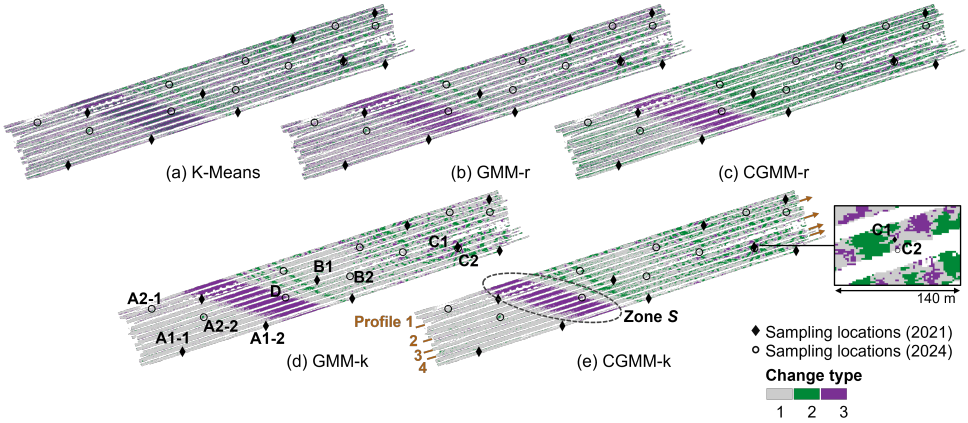


Figure 4.12: 3-cluster change maps for Area BoS. Results for incident angles between 25° and 60° are presented. For the CGMM methods, the change type 1, 2, and 3 can be regarded as “unchanged”, “softer”, and “coarser”, respectively.

For CGMM-r and CGMM-k, clusters 1, 2, and 3 can be interpreted as “unchanged”, “softer”, and “coarser”, respectively. Most of the western seabed of BoS is classified as “unchanged”. Besides, a stripe pattern (Zone *S* circled in Fig. 4.12e) suggests areas where the seabed may have become “coarser” in 2024 compared to 2021. On the eastern seabed, an alternating pattern of the “softer” and “coarser” detections is observed.

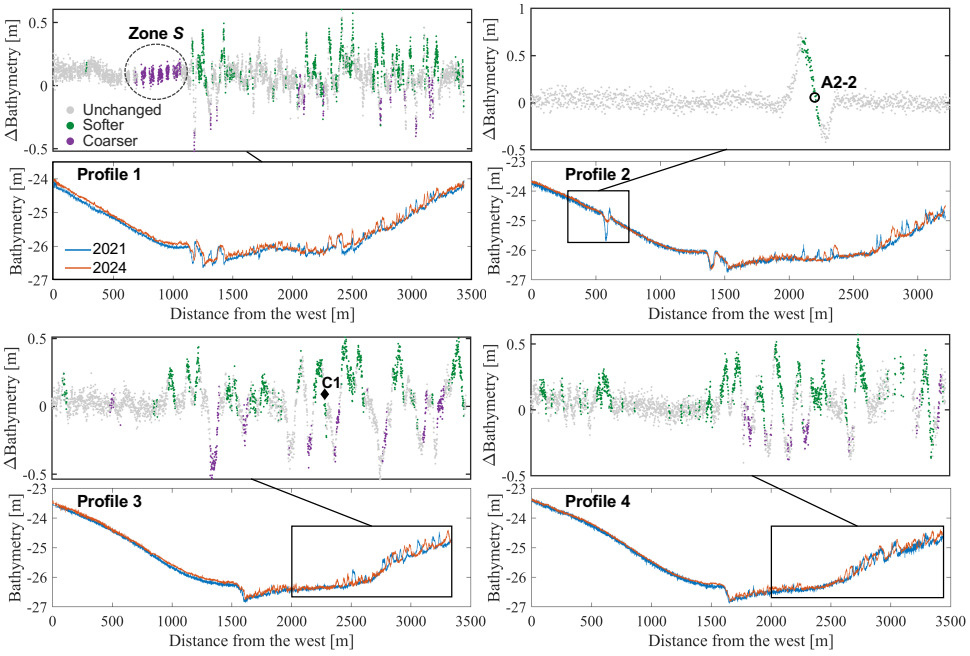


Figure 4.13: Comparison of bathymetry profiles between 2021 and 2024, colored by the CGMM-k change detection results. ΔBathymetry is calculated by subtracting the bathymetry in 2021 from that measured in 2024.

This alternating pattern on the eastern seabed may be attributed to sediment movement over time, as indicated by the bathymetry data along four profile lines in Area BoS (Fig. 4.13). In 2021 and 2024, BoS exhibited similar geomorphological features, such as the bathymetric drop in the middle and the heterogeneous reef-like structures with a height of about 40 cm in the east. However, the relocation of those reef-like structures after almost three years can be noticed. Comparing the bathymetric difference with the CGMM-k change detection results reveals that the "softer" and "coarser" seabed east of the bathymetric drop corresponds to the increase and decrease in bathymetry, respectively. The consistent amplitude of these changes further supports the possibility of horizontal sediment movement (see Profile 1, 3, 4 in Fig. 4.13). In addition, bathymetric differences along Profile 2 also show "softer" locations on the western seabed that might be due to sediment transport (near the sampling location A2-2). By contrast, the "coarser" seabed in Zone *S* in the west does not experience significant changes in bathymetry.

Seabed samples collected in 2021 and 2024 also help to validate the change detection results. The Folk class, *L. conchilega* density, and dead shell content of three sample groups are compared (see Table 4.2). Each group includes samples taken from different years but at nearby locations. Group A (A1-1, A1-2, A2-1, and A2-2 shown in Fig. 4.12d)

[2]Locations of samples from Group A, B, and C are shown in Fig. 4.12d. For Group A, average values of samples from the same year are presented.

Table 4.2: Comparison between change detection results and changes in seabed sample properties (2021 → 2024) for Area BoS

Group	A[2]	B	C
Folk class	fS → fS	gmS → gmS	mS → cS
L. conchilega density [ind/m^2]	17.1 → 64.1	**1175.2 → 0**	653.8 → 179.5
dead shell content [%]	0.26 → 1.51	0.18 → 1.72	1.30 → 2.23
Change type (CGMM-k)	Unchanged	**Softer**	–

contain sediments with the same Folk class, which aligns with the "unchanged" seabed type indicated by the CGMM-k results. However, the average *L. conchilega* density of A2-1 and A2-2 is higher than in A1-1 and A1-2. Nevertheless, very few *L. conchilega* are found for Group A in general. The dead shell content shows a slight increase in 2024 compared to 2021. Similarly, for Group B, the Folk class remains unchanged, with a moderate increase in dead shell content. However, the *L. conchilega* density drops significantly from 2021 to 2024, which is consistent with the detected seabed change. The "softer" detections from CGMM-k in the neighborhood of sampling locations B1 and B2 may result from the lower presence of *L. conchilega* in spring 2024, though more ground truthing is required to confirm this.

C1 and C2 are located close to each other, but in a heterogeneous region, which might explain why they lead to different seabed change types (see the zoomed-in plot in Fig. 4.12e). Folk classes of C1 and C2 might show the difference in the sediment from the reef-like structures (mS) and troughs between them (cS). In addition, C1 has a much higher *L. conchilega* density than C2, possibly due to the different sediment types or seasonal variations in the abundance of *L. conchilega*. Moreover, Location D on the western seabed of BoS is classified as "coarser" by CGMM-k, likely due to the presence of empty shells, as observed in boxcore pictures and video transects for sampling locations from Group A and Location D. The seabed videos also show a brownish color of the sediment at Location D, possibly indicating organic matter left by the living shellfish. However, no 2021 samples are available in Zone *S* to validate the detected seabed change.

Area AT

As observed from the CGMM-r and CGMM-k change maps, the "unchanged" detections are sparse in the northeast of Area AT (Fig. 4.14). Some across-track inconsistency can also be noticed in the "unchanged" detections, possibly due to the worse clustering performance of 20°–40° compared to 40°–60°. Nevertheless, the spatial pattern of seabed change types in Area AT is more homogeneous than that in BoS. Based on the cluster mean values of ΔBS, change types 2 and 3 can be interpreted as "slightly softer" and "softer", respectively. The "softer" seabed is mainly located near the trough in the middle of the study area, with some presence in the west, while "slightly softer" detection results align with the seabed slopes next to the trough.

Due to the homogeneity in each detected change type of AT, validation using bottom samples can be straightforward both qualitatively and quantitatively. Across all four ground truth sample groups, the Folk class remains unchanged, while the dead shell content shows a slight increase from 2021 to 2024 (see Table 4.3). By contrast, all groups

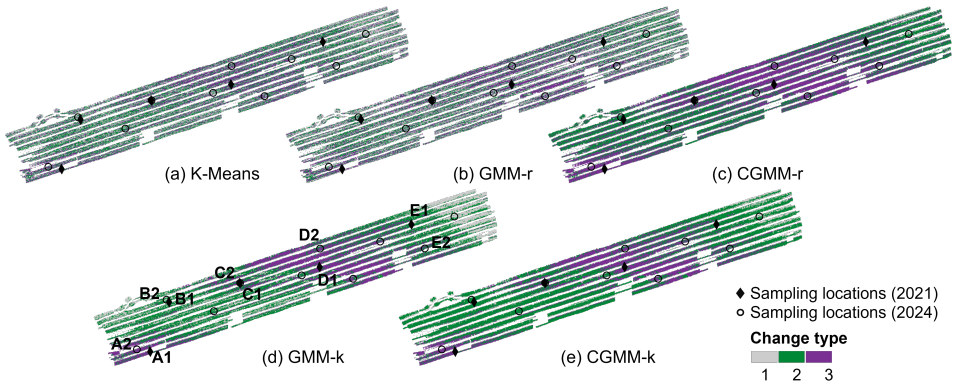


Figure 4.14: 3-cluster change maps for Area AT. Results for incident angles between 25° and 60° are presented. For the CGMM methods, the change type 1, 2, and 3 can be regarded as "unchanged", "slightly softer", and "softer", respectively.

exhibit a decrease in *L. conchilega* density over the same period. The seabed near B1 and B2 is classified as "slightly softer" by CGMM-k, corresponding to a decrease in the *L. conchilega* density of about 100 [ind/m^2]. In Groups C and D, the *L. conchilega* density drops by more than 600 [ind/m^2], which may explain their classification as "softer". In a similar manner, locations A1 and A2 are also identified as "softer", with a decrease in the density reaching 200 [ind/m^2]. $\Delta BS(50^\circ)$ for locations near Group A is about -5 dB, which is close to the boundary value between the change types "slightly softer" and "softer" given by CGMM-k (see Fig. 4.11). On the other hand, Group E is classified as "slightly softer", but no significant change is observed from the sample properties. Since defining "softer" and "slightly softer" of the ground truth samples is challenging, the accuracy and F1-score for the binary "changed" and "unchanged" classification were calculated, demonstrating the effectiveness of the CGMM-k detection results with the metric values of 0.80 and 0.89, respectively.

Table 4.3: Comparison between change detection results and changes in seabed sample properties (2021 → 2024) for Area AT

Group	A[3]	B	C	D	E
Folk class	fS → fS	fS → fS	fS → fS	mG → mG	mG → mG
L. conchilega density [ind/m^2]	**185.9 → 0**	**102.6 → 0**	**641.0 → 0**	**698.7 → 0**	12.8 → 12.8
dead shell content [%]	0.40 → 0.92	0.24 → 0.26	0.15 → 0.99	0.24 → 0.31	0.01 → 0.12
Ground truth change type	**Changed**	**Changed**	**Changed**	**Changed**	Unchanged
Change type (CGMM-k)	**Softer**	**Slightly softer**	**Softer**	**Softer**	Slightly softer
				Accuracy: 0.80	F1-score: 0.89

Across both study areas, BoS and AT, the sediment types (Folk classes) remained consistent from 2021 to 2024. The slight increase in the dead shell content is not reflected

[3]Locations of samples from Group A, B, C, and D are indicated in Fig. 4.14d.

in the change detection results. Seabed changes revealed by CGMM-k might be more related to the change in living benthos such as *L. conchilega*, which increases seabed surface roughness through their tube-building activities [138].

4.4.6. Limitations and future work

ΔBS between distinct surveys is affected by multiple seabed properties in a combined way. In this research, the inconsistency in sampling locations across years, especially in Area BoS, hinders the quantitative analysis of the relations between ΔBS and various bottom properties. This also brings difficulties in validating some of our detected seabed changes. With constraints in time and budget during our surveys, priority was given to broader spatial coverage with a limited number of bottom samples. A more comprehensive study of how different seabed properties contribute to ΔBS would require long-term monitoring with consistent sampling locations. Another limitation lies in the lack of an independent and long-term monitored reference area. For seabed monitoring surveys in the Dutch North Sea, well-established reference areas, such as the Kwinte area located in the Belgian North Sea [70], can be considered in the future.

In addition, this research focuses on case studies with relatively limited spatial extents in the Dutch North Sea. To investigate the influence of a wider range of seabed types and increased data volumes on BS uncertainty and change detection performance across incident angles, repeated MBES measurements in diverse seabed environments and on larger spatial scales are needed. Seabed types and ocean environments can also affect the optimal temporal interval of acoustic monitoring surveys. For the *L. conchilega* habitats investigated in this research, seasonal observations within a year might be beneficial. In sandy sediments affected by a high current speed, angular characteristics of BS measurements might change in hours due to variations in sand ripples [139], allowing for more frequent monitoring. Moreover, based on relative calibration, future research can investigate other machine learning based benthic habitat mapping methods to enhance the use of various types of acoustic features in seabed change detection.

4.5. Conclusion

Repeated MBES backscatter surveys can help to monitor changes in the benthic habitat, but bring the challenge of limited seabed ground truth and complexity in backscatter data processing. In this regard, we apply GMM clustering to the backscatter difference per incident angle from repeated surveys, achieving unsupervised seabed change detection. We ensure the identification of the “unchanged” and “changed” seabed, by fitting the “unchanged” cluster as a zero-mean Gaussian and constraining its variance using the estimated backscatter uncertainty from a stable reference area. The method is also flexible in describing different types of seabed changes, through fitting multiple Gaussians.

With Gaussian parameters of “changed” clusters initialized by the K-means clustering results, the proposed constrained GMM method shows promising change detection results on two study areas in the Dutch North Sea. The proposed method detects the seabed change possibly caused by a reduction in the *L. conchilega* abundance, which

might be due to the seasonal change in the seawater environment, such as the temperature. For one of the study areas, the change detection results reveal alternating patterns between the seabed becoming “coarser” and “softer”, which aligns with the sediment movement after almost three years, as indicated by the bathymetry data. By processing each incident angle individually, we show that backscatter data from 40° to 60° might have the largest ability to detect seabed changes, which is consistent with the backscatter uncertainty estimation. Regarding designing surveys for monitoring purposes, it is then helpful to enlarge the overlap between adjacent track lines.

Limitations of this research lie in the lack of an independent reference area and different seabed sampling locations in repeated surveys. Nevertheless, our study still demonstrates the value of using MBES backscatter data for benthic habitat monitoring, especially regarding changes in the benthic communities, which can be extended to long-term acoustic datasets in the future.

5

Mapping cultivated blue mussels using semi-supervised machine learning on MBES measurements

The richness I achieve comes from nature, the source of my inspiration.

Claude Monet (1840—1926)

The previous two chapters have shown the advantage of using backscatter data acquired by the multibeam echosounder (MBES) for seabed characterization. As a step further, this chapter presents a combined use of bathymetry and backscatter measurements, with a focus on leveraging a small set of ground truth data. With seabed samples and MBES data, machine learning (ML) techniques are commonly employed to model their empirical relationships and generate classification maps of an extended seabed. However, ground truth samples with a limited number and sparse spatial distributions may impede the development of robust ML models. To address this challenge, we propose to apply a semi-supervised ML method in seabed classification using MBES data. With acoustic measurements from a blue mussel (Mytilus edulis) cultivation area in the Oosterschelde, the Netherlands, and only nine boxcore samples, a pseudo-labeling strategy is introduced. The resulting pseudo-labels will enlarge the training data size, facilitate the training of three comprehensive ML algorithms (Gradient Boosting, Random Forest, and Support Vector Machine), and help to classify the study site into mussel and non-mussel areas. For this mussel culture detection task, we find the geomorphological and backscatter-related features to play complementary roles. Moreover, our classification results are demonstrated effective through expert knowledge of this cultivation area and may bring insights for future research on natural mussel habitats.

This chapter has been published in Journal of Environmental Management 369 (2024), p. 122250. [140]

5.1. Introduction

Preserving benthic habitats, such as oyster reefs and shellfish aggregations, is crucial for ensuring sustainable utilization of the ocean environment [97, 100]. These habitats can be threatened by overexploitation or other forms of bottom disturbance, such as sand extraction. To alleviate the pressure faced by marine habitats and maintain a healthy ocean ecosystem, regulations of human activities in fragile areas are necessary, which further requires effective methods for mapping the seabed habitats, especially the occurrence of marine benthos [52].

Acoustic remote sensing techniques offer efficient ways of seabed biodiversity monitoring [38, 43]. The multibeam echosounder (MBES) has been widely used over the past decades to survey extended parts of the seafloor. Although bathymetry has traditionally been the primary product of MBES, the intensity of the backscattered signal is also a crucial measurement since it can provide indications of the reflective and scatter properties of seabed materials. Geophysical properties of the seabed sediment such as grain size, roughness, and porosity, can affect the level of backscatter strength (BS) at a certain incident angle as well as the backscatter angular variation characteristics [16, 98]. Some marine benthos such as oysters and mussels can affect the sediment geoacoustic properties due to the presence of their shells [21]. Their aggregations might also modify the seafloor geomorphology, which can be linked to bathymetry. These characteristics enable the discrimination of "acoustically hard" marine benthos from the bare sediment using bathymetry and backscatter data [22].

While the MBES provides a non-destructive and efficient way for extensive seabed habitat mapping, ground truthing is still essential for establishing a link between the acoustic data and sedimentological and macrofauna properties. Limited ground truth, which is commonly encountered in seabed applications due to obligations such as minimizing the seabed disturbance and restrictions in time and budget, can pose challenges to establishing the relationships between MBES data and seabed properties in a supervised fashion. Despite this, the abundant acoustic measurements bring up the possibility of leveraging the distribution of unlabeled data to assist training. Combining both labeled and unlabeled data in training is the core idea of semi-supervised learning [82, 141], which is an important category in data-driven methods. Different techniques have been developed in semi-supervised learning, including but not restricted to graph-based learning, self-training, co-training, multi-view learning, low-density separation, and generative models [142]. Semi-supervised machine learning techniques have been successfully adopted in many fields, including classifying MBES data. For instance, Minelli *et al.* [82] used semi-supervised machine learning to enhance the detection of fish schools from the MBES water column data. However, the applications in seabed characterization, especially for marine benthos monitoring, are still scarce.

In this study, we highlight the advantage of semi-supervised methods in mitigating the problem of limited ground truth in seabed mussel habitat mapping. We applied a graph-based learning method, denoted label spreading [143], to MBES bathymetry and backscatter data to facilitate the mussel detection. Compared to other semi-supervised techniques, such as self-training that starts with training a full-supervised model on the labeled data, label spreading can be less prone to overfitting and achieve higher classification accuracy [144]. To further ensure the performance of label spreading, we

also assessed the Silhouette coefficient of pseudo-labels to help the hyperparameter selection during the graph construction.

We considered a mussel cultivation area in the Oosterschelde tidal basin of the Netherlands. The blue mussels (*Mytilus edulis*), one of the most important aquaculture products of the Netherlands, are cultivated here. Using MBES data and only nine seabed bottom samples, we classified the seabed into mussel and non-mussel regions. With label spreading, pseudo-labels were generated using a few ground truth samples. These pseudo-labels were further employed to train three machine learning algorithms: Gradient Boosting (GB), Random Forest (RF), and Support Vector Machine (SVM). We then applied the trained models to classify the entire study area. The effectiveness of the classification results was qualitatively validated by comparing them with the ecological knowledge of the study area. Given the artificial cultivation situation of this area and well-documented mussel farming locations, the validation process is straightforward. This enhances the confidence in the accuracy of the classification outcomes and will facilitate the conservation of natural mussel habitats in the future.

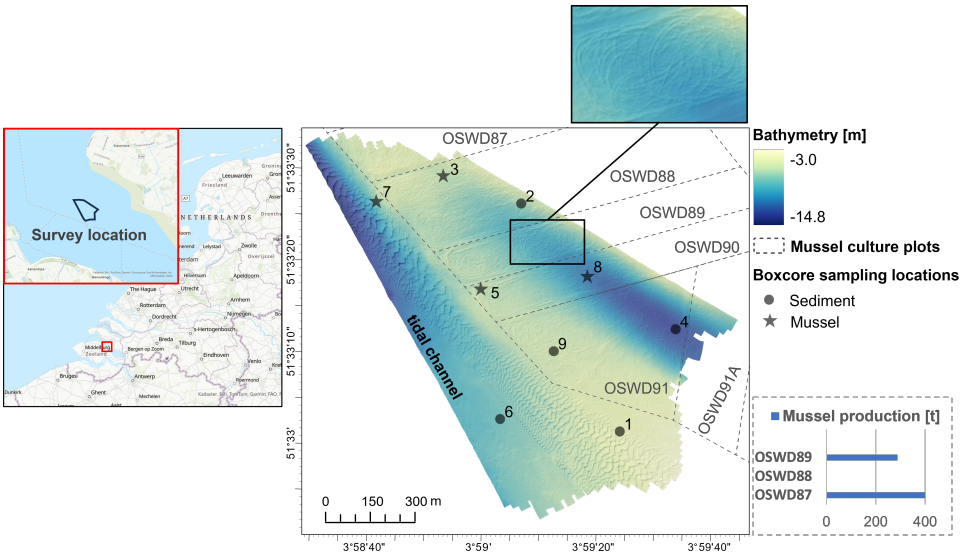


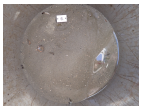
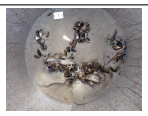
Figure 5.1: Study area. (Left) Survey location in the Oosterschelde tidal basin. (Right) Bathymetry map acquired by the MBES, with a grid size of 0.25 × 0.25 m. East of a tidal channel, mussel occurrence of nine boxcore samples and mussel culture plots covered by the study area are displayed. The shapefiles of the mussel culture plots were provided by the Netherlands Enterprise Agency (Rijksdienst voor Ondernemend Nederland, RVO). The zoomed-in figure shows the circular pattern in bathymetry caused by dredging during mussel farming. Available information regarding mussel productions of three culture plots (OSWD87, 88, and 89) was acquired from the Dutch Mussel Auction.

5.2. Study area and dataset

The study area is located in the Oosterschelde, a tidal basin in the southwest of the Netherlands (Fig. 5.1). The Oosterschelde has an important economic value due to shellfish culture and has been intensively researched during the past decades [34]. The water depth of the study area ranged from 3.0 to 14.8 m. Sand ripples extended from the northwest to the south of the surveyed area. The area is a known site for mussel bottom cultures, in which mussels are typically seeded in circular patterns on the seabed sediment [145]. Dredging activities during this type of mussel farming can form circular crests and troughs that are also visible in bathymetry (see the zoomed-in plot in Fig. 5.1).

We acquired the MBES dataset in the study area on May 2, 2023, using a dual-head system Kongsberg EM2040c (Kongsberg Gruppen, Kongsberg, Norway), which was operated at a frequency of 300 kHz. The swath coverage was 140°, with the nominal beam opening angle of 1.0° in both along- and across-track directions. We used a nominal pulse length of 37 μs during the survey. The collected MBES data, including bathymetry [m] and beam-averaged backscatter strengths in dB, were saved in the Kongsberg logging format *.kmall*.

Table 5.1: Detailed properties of 9 boxcore samples

Station	d_{50} [μm]	%mud	Mussel occurrence	Class	Boxcore picture
1 2 4 6 9	215.4 204.9 179.9 258.7 200.2	4.0 7.0 5.8 0.0 6.6	No mussels	Sediment	
3 7	170.8 264.8	34.2 0.0	A few living mussels	Mussel	
8	178.2	12.3	30% covered by mussels	Mussel	
5	114.0	42.4	70% covered by mussels	Mussel	

For ground truthing, sediment samples at 9 stations (Fig. 5.1) were collected using a cylindrical boxcore sampler with a diameter of 30 cm. Grain size distributions were determined afterward through laboratory analysis. Median grain size (d_{50}) of the samples ranged from 114.0 to 264.8 μm (Table 5.1). In addition, according to the particle size distribution classification [146], no gravel content was found in any of the stations, while mud content ranged from 0 to 42.4%, indicating a small change in sediment types from muddy sand to sand. Mussel occurrence at each sampling station was determined onboard by visual inspections. Living mussels were found in four boxcore samples, and two of them (stations 5 and 8) showed a high coverage (Table 5.1). The boxcore samples,

therefore, were classified into two classes, “sediment” (sediment without mussels) and “mussel” (sediment with mussel coverage).

Additionally, mussel culture plots documented by the Netherlands Enterprise Agency show the extent of mussel cultivation. Among them, mussel productions within three culture plots are available. It was reported that plot OSWD87 (Fig. 5.1) was covered by mussel seeds and medium-sized mussels of 400,000 kg at the time of our survey, while OSWD88 (Fig. 5.1) was empty, which means that there were no mussels or very few mussels that were too small to be harvested. Moreover, 288,000 kg of large mussels were collected from the OSWD89 plot (Dutch Mussel Auction, pers. comm.). These mussel culture plots are located east of the sand ripples in the surveyed area, while the sand ripples are on the eastern slope of a tidal channel according to the bathymetric information of the Oosterschelde documented by the Current Dutch Elevation dataset (Actueel Hoogtebestand Nederland, AHN) in 2013.

5.3. Methodology

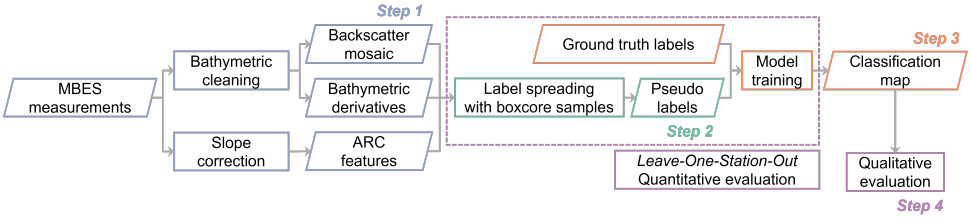


Figure 5.2: Workflow of the semi-supervised mussel classification using MBES data.

Our workflow for detecting the mussel occurrence from MBES data comprises three main steps (see steps 1–3 in Fig. 5.2). Initially, we extract and select secondary features from the MBES measurements (Section 5.3.1). These features, combined with the ground truth samples, are then employed to generate pseudo-labels (Section 5.3.2). Following this, we train machine learning models using the expanded labeled dataset and conduct classification across the entire study area (Section 5.3.3). In Section 5.3.4, we present the validation procedure of our semi-supervised classification results, including comparisons with the full-supervised method.

5.3.1. Step 1: MBES feature extraction

Extraction of effective features from the raw MBES measurements is the prerequisite for training reliable machine learning algorithms applicable for seabed habitat monitoring. Before feature extraction, it is important to clean the data and apply all required corrections. We cleaned the bathymetric data for bottom misdetections through a spline filter and manual edits in the QPS software Qimera. The bathymetric data were then rasterized as GeoTIFF with a grid size of 0.25 × 0.25 m for the entire area (Fig. 5.1).

Since mussels inhabit the seabed surface, their presence will modify the seafloor geomorphology. Surface slope and curvature are two geomorphological features that

have been widely used to describe landforms in various fields [147]. While the slope represents the steepness of each surface patch relative to the horizontal plane, curvature describes the amount of bending of a local surface along a certain direction on the surface [148]. This direction can be, for instance, along the slope line or the contours. We computed the slopes and curvatures from bathymetry in Esri ArcGIS Pro using the *Surface Parameter* tool. In this study, we calculated the Casorati curvature [149], which is a combinatorial metric, with high positive values showing areas of sharp bending in multiple directions. Moreover, the local surface patch size was determined by the variability in the local terrain.

The *BS* is also a powerful discriminator in seafloor classification and benthic habitat monitoring [150]. Although absolute backscatter calibration was not available, uncalibrated backscatter strengths collected within our survey can still be comparable. We account for both the backscatter strength values and their angular dependence by generating the backscatter mosaic and angular response curve (ARC). The cleaned soundings in Qimera were exported as generic sensor format (GSF) files, which were further used to generate the backscatter mosaic in the software FMGT with a grid size of 1 × 1 m. Since the backscatter strength is a function of the incident angle, the mosaicking process involves eliminating this variation through the angle-varying gain correction [62], which requires setting a subset of incident angles as reference. Then *BS* values at other angles are normalized to this reference.

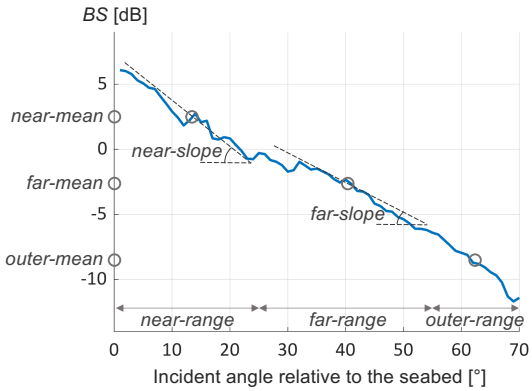


Figure 5.3: Illustration of the extracted features from a half-swath ARC.

The angular variation itself, however, is also intrinsic to seabed properties [98]. To account for this variation in our analysis, we extracted the incident angle and *BS* of each beam from the *.kmall* files using the CoFFee library in MATLAB. Before building the ARCs, a slope correction was conducted to achieve the actual incident angle ϕ [°] relative to the local seabed (see Eq. 2.11). To reduce noise, the ARCs were also averaged over 10 consecutive pings, which covered a distance of about 2 m. We divided a half-swath ARC (either port or starboard side) into three angular ranges, 0–25° (near-range), 25°–55° (far-range), and 55°–70° (outer-range) [30]. The mean values of the three angular ranges were then calculated. Considering the ARC shape, we also calculated the slopes (with respect to incident angle) of both ranges through a linear least-squares fit (Fig. 5.3).

All extracted features, including two bathymetric derivatives (slope and curvature), the backscatter mosaic, and five ARC features (near-mean, far-mean, outer-mean, near slope, and far-slope) were rasterized as 1 × 1 m grids. When the spatial resolution of a feature was worse than 1 m (such as the ARC features), bilinear interpolation was conducted. To avoid including redundant information in learning, we also analyzed the correlation among features and eliminated features having Pearson's coefficient higher than 0.8 with others (see Section 5.4.1).

5.3.2. Step 2: Pseudo-labeling of MBES features using label spreading

Label spreading is a graph-based semi-supervised learning method proposed by Zhou *et al.* [143]. It assumes that the structure of unlabeled data is consistent with the structure of the ground truth classes, and hence data points having similar features are likely to be classified identically. This will subsequently allow propagating classes of the labeled data to the unlabeled observations. To achieve this propagation, a graph representation is built based on the features of all data points.

Given a point set $\mathscr{P} = \{p_1, ..., p_l, p_{l+1}, ..., p_n\} \subset \mathbb{R}^d$, each point is represented by a feature vector with length d. The first l points have labels from the label set $\mathscr{L} = \{1, ..., c\}$, and the remaining points are unlabeled. Each data point serves as a node in the graph $G = (V, E)$. The graph can be fully connected or constructed using k-nearest neighbors (NN) to ease the computational burdens. The vertices V equals the point set $\mathscr{P}$. The edges E represent the connections between data points and are weighted by a similarity matrix W, with $W_{ij} = \exp(-\|p_i - p_j\|^2/2\sigma^2)$ if $i \neq j$ and $W_{ii} = 0$. σ^2 is a scaling parameter.

In label spreading, the classification rule of $\mathscr{P}$ is denoted by a $n \times c$ matrix $F = [F_1^T, ..., F_n^T]^T$. F_i^T represents the classification rule for point p_i, resulting in its label $y_i = \arg\max_{u \leq c} F_{iu}$. F is determined by an iterative process and initialized as Y, which is also a $n \times c$ matrix and consistent with the information from the initial labeled points. $Y_{iu} = 1$ if p_i is labeled as $y_i = u \in \{1, ..., c\}$. $Y_{iu} = 0$ if p_i is unlabeled. The label spreading algorithm further defines the cost function associated with F as

$$Q(F) = \frac{1}{2}\left(\sum_{i,j=1}^{n} W_{ij} \left\| \frac{1}{\sqrt{D_{ii}}} F_i - \frac{1}{\sqrt{D_{jj}}} F_j \right\|^2 + \mu \sum_{i=1}^{n} \|F_i - Y_i\|^2 \right), \tag{5.1}$$

where D is a diagonal matrix with the (i,i)-entry equal to the sum of the ith row of W, and $\mu > 0$ is the regularization parameter. D normalizes the weight matrix W and helps to increase the robustness of the label spreading algorithm regarding data noise [144]. Without normalization, nodes with more connections (and thereby potentially high degrees) may dominate the labeling process. In addition, the first term on the right-hand side of Equation 5.1 makes sure that neighboring points in the feature space have similar classification rules. The second regularization term adds constraint in the algorithm so that the final classification does not deviate too much from the initial label assignment. The final classification rule of $\mathscr{P}$ will be achieved by $\arg\min Q(F)$.

In this study, we used the label spreading algorithm implemented in the *scikit-learn* machine learning library in Python. Given the ground truth classes of the boxcore samples (Table 5.1), the label of each MBES data point would be either "sediment" or

"mussel". We treated MBES points within a 3 m radius around each boxcore sampling station as labeled data. Considering both efficiency and the need for enlarging the training dataset, we randomly sampled 5000 points from the unlabeled MBES data. A k-NN graph was then constructed based on these points and the labeled data. Via label spreading, we further propagated the class ("sediment" or "mussel") of the labeled data to those 5000 MBES points.

The generated 5000 pseudo-labels served as additional training data, which helps to make better use of high-dimensional features and alleviate overfitting. In addition, the value of k for building the k-NN graph concerns how many neighboring points are connected to each node in the feature space. Choosing an optimal k, however, is not trivial. We selected the k value by assessing the Silhouette coefficient [132] of the pseudo-labeled data, which measures how well different clusters are separated and helps to ensure enough distinction between different classes of the pseudo-labeled data. The Silhouette coefficient is a normalized metric, with 1 the best, -1 the worst, and 0 indicating overlapping clusters.

5.3.3. Step 3: Training with pseudo-labels and map prediction

The expanded training dataset from step 2 comprised the labeled MBES points near the sampling stations and 5000 pseudo-labeled points. Each point was associated with the selected MBES features and a class ("sediment" or "mussel"). Using these training data, we trained three machine learning models that have been extensively used in seafloor mapping applications [151]: Gradient Boosting (GB), Random Forest (RF), and Support Vector Machine (SVM) with the Radial Basis Function (RBF) kernel. We determined the hyperparameters of these machine learning algorithms through an exhaustive grid search based on a 5-fold cross-validation [152]. Following this, each trained model was used to predict a mussel classification map for the entire study area. We then conducted a qualitative assessment for all these classification maps (see Section 5.3.4).

5.3.4. Step 4: Validation of the classification method

We evaluated our classification method both qualitatively and quantitatively. For the qualitative evaluation, we assessed the classification maps obtained in step 3 based on expert knowledge of the study area, including the interpretation of the bathymetric data and mussel productions of different culture plots (Fig. 5.1). Due to the scarcity of bottom samples, these classification maps were predicted by models trained with all 9 boxcore samples.

Regarding the quantitative validation, we conducted the Leave-One-Station-Out (LOSO) cross-validation using the ground truth data (Fig. 5.2). LOSO was built upon the concept of Leave-One-Out cross-validation [152] that is commonly used in machine learning for small datasets. As explained in Section 5.3.2, we associated ground truth labels of the boxcore samples with the nearby MBES features. Thus, nine sampling stations resulted in nine groups of labeled MBES points. During each round of the LOSO cross-validation, one group (out of nine) was left out and considered as the validation dataset. The other groups of labeled MBES points were used to perform pseudo-labeling and model training. We then predicted the classes of the validation data using the trained

model. Afterward, predictions from nine rounds of LOSO cross-validation were assessed based on their ground truth labels and two metrics, which are accuracy and F1 score.

To investigate the effectiveness of adding pseudo-labeled data in training, we also compared the classification maps (from step 3) and LOSO cross-validation scores (from step 4) with the full-supervised learning results, in which model training only involved the ground truth data. The impact of pseudo-labels was also assessed by differences in the feature importance from the ensemble machine learning methods (GB and RF) between the full- and semi-supervised classification. In GB and RF, feature importance is determined by assessing the impact of each feature on reducing uncertainty across the ensemble of models [153]. Thus, it can be a measure of the significance of each feature to the overall predictive performance of the combined models.

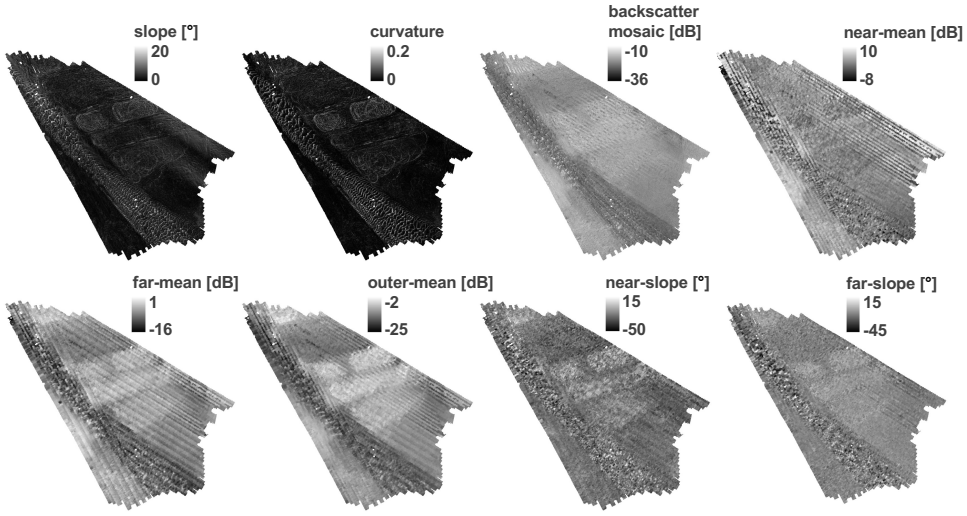


Figure 5.4: Features extracted from the MBES measurements.

5.4. Results and discussion

5.4.1. Feature extraction and selection

Different spatial patterns per region can be observed from the extracted MBES features (Fig. 5.4). Circular patterns caused by the mussel farming activities are highlighted in the geomorphological features (slope and curvature). Regions with sand ripples also show high values, especially in slope. These sand ripples were outside the extent of mussel culture plots and much less likely to contain cultivated mussels. This indicates that solely relying on these two geomorphological features is insufficient to discriminate mussels from sand ripples. Conversely, sand ripples show the lowest values in the backscatter mosaic, far-mean, and outer-mean. The highest values of these three features are located in the mussel cultivation areas. Near-mean shows a more mixed pattern across the study area but highlights the tidal channel in the west. Backscatter mosaic, far-mean, and

outer-mean also show high values for the channel, possibly indicating coarser sediments. In addition, near-slope and far-slope present distinct patterns, with high values in the mussel cultivation area and a lot of variation in the regions of sand ripples.

The aforementioned similarity in the backscatter mosaic, far-mean, and outer-mean is also confirmed by the strong correlation among them (Fig. 5.5). To reduce redundant information, we eliminated far-mean from the feature set. There also exists a large number of striped artifacts in far-mean, so this exclusion will avoid bringing these artifacts to the classification results. The remaining seven features were then considered in pseudo-labeling and model training.

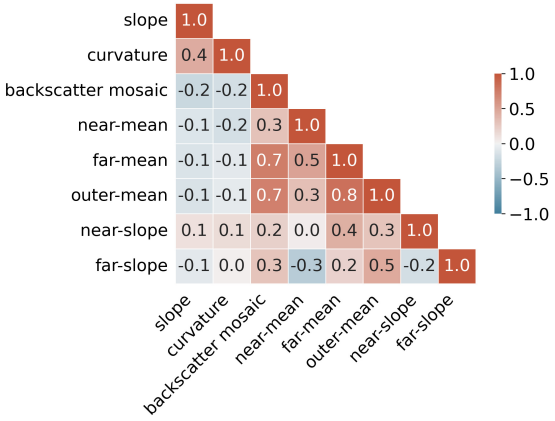


Figure 5.5: Correlation among all extracted MBES features, annotated by Pearson's coefficient.

Using a radius of 3 m around the nine sampling stations (Fig. 5.1), we collected 284 MBES points in total. These points were labeled with the ground truth class of the corresponding boxcore sample, namely "sediment" (160 points) or "mussel" (124 points). Each labeled point was also associated with the seven selected MBES features. We further analyzed the relationship between these features and their ground truth labels using boxplots (Fig. 5.6). It can be observed that the occurrence of mussels has increased the value of all features. It is also evident that the range of far-slope, belonging to two ground truth classes, overlap the most. Moreover, values of outer-mean show the clearest division between the sediment and mussel classes. This indicates that mussels can modify the seabed surface roughness to a large degree [18].

5.4.2. Pseudo-labeling results using all boxcore samples

Using all ground truth data (284 points) and 5000 randomly selected unlabeled points, we performed label spreading and assessed different values of $k \in [10, 5000]$ for building the k-NN graph. The Silhouette coefficient of the pseudo-labeled data reaches the peak when k is around 1200 (Fig. 5.7). The results for k values larger than 2500 are not presented since all pseudo-labels belong to the same class (sediment) in those cases. Therefore, we selected $k = 1200$ as an optimal value for pseudo-labeling, as it maximizes the Silhouette coefficient and at the same time ensures clustering the data into two

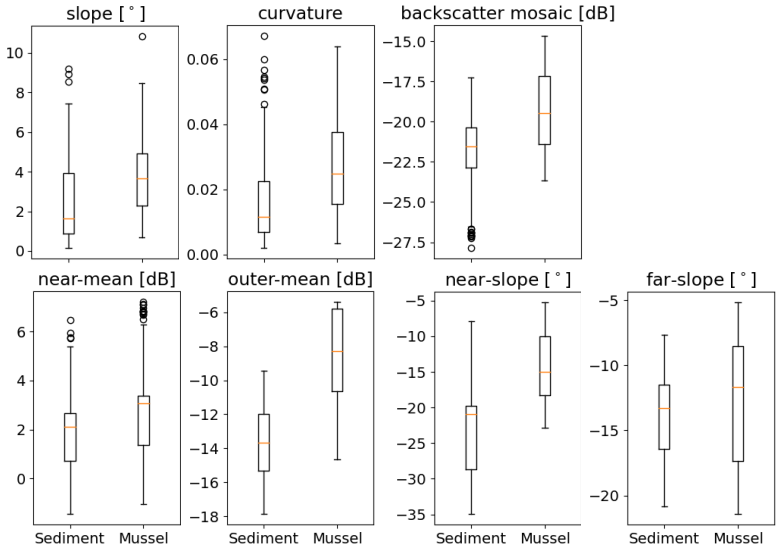


Figure 5.6: Boxplots for the selected MBES features of the labeled data, with the sediment and mussel class plotted separately. Features were collected within a radius of 3 m around each boxcore sampling station.

classes. It is noted that the best Silhouette coefficient is around 0.23, indicating that MBES features of the sediment and mussel class can overlap to some extent. This is what we would expect since the mussel class encompasses cases where the seabed sediment contains only a limited number of mussels.

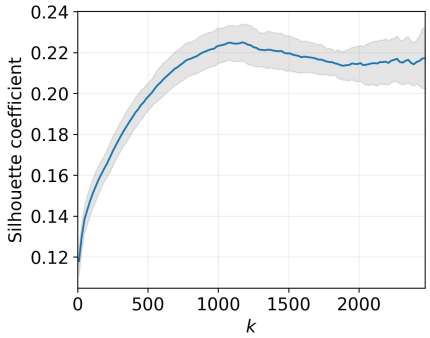


Figure 5.7: Silhouette coefficient of 5000 pseudo-labels with different k in label spreading. The solid curve and shaded area indicate the mean and standard deviation for randomly selecting 5000 points and running label spreading ten times.

We also present the pseudo-labeling results in the feature space using principal component analysis (PCA) [154] of all selected features (Fig. 5.8). From the first three principal components (PCs), which contribute to the total variability of 73% (32%, 21%, and

20%, respectively), it is observed that pseudo-labeled points from the “sediment” and “mussel” class are generally distinct in the PCA space. To easily compare pseudo-labels with the ground truth, we also present them on the two-dimensional space using the first two PCs. The general observation is that the pseudo-labeled points closely follow the pattern of the ground truth points. Overlapping of features from the sediment and mussel class exists to some extent, especially near sampling stations 3 and 7, where only a limited number of living mussels were found.

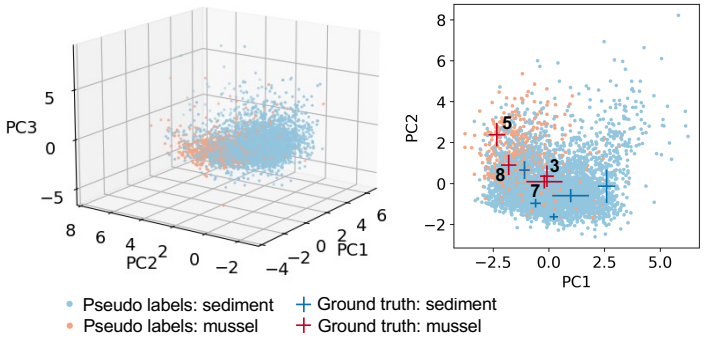


Figure 5.8: Visualization of 5000 pseudo-labeled points (with the optimal $k = 1200$) in the feature space using PCA of the selected MBES features. (Left) 5000 randomly sampled points plotted on the span of the first three PCs, with the point color defined by their pseudo-labels. (Right) Comparison between the ground truth and pseudo-labels using the first two PCs. The cross symbols represent the mean and standard deviation of PC1 and PC2 from the labeled MBES data points. Sampling stations 3, 7, 5, and 8, which contained mussels, are also indicated.

5.4.3. Semi-supervised classification results

Prediction maps of the study area and qualitative evaluation

We trained GB, RF, and SVM in both full- and semi-supervised manner. While only 284 ground truth points were used for model training in full-supervised learning, 5000 pseudo-labeled points generated using the ground truth data were also used as input in semi-supervised learning. Compared to SVM, GB and RF have more tuning parameters, such as the number of trees and the depth of each tree. Through an exhaustive grid search, we chose the optimal number of trees as 60 for full- and semi-supervised GB, with a depth of 3 and 4, respectively. Regarding RF, 5 trees and a depth of 4 were selected for the full-supervised method, while 30 trees and depth of 4 were used in the semi-supervised method.

We further compare the classification maps achieved by the full- and semi-supervised methods (Fig. 5.9). It can be observed from the full-supervised prediction results that the mussel class is overall consistent with areas showing high values in the backscatter mosaic and ARC mean features. Lots of nadir striped artifacts in the ARC features are also included in the mussel class. This indicates that the classification was sensitive to noise, possibly due to overfitting caused by an insufficient number of ground truth samples [155]. The tidal channel located west of the study area is not distinguished

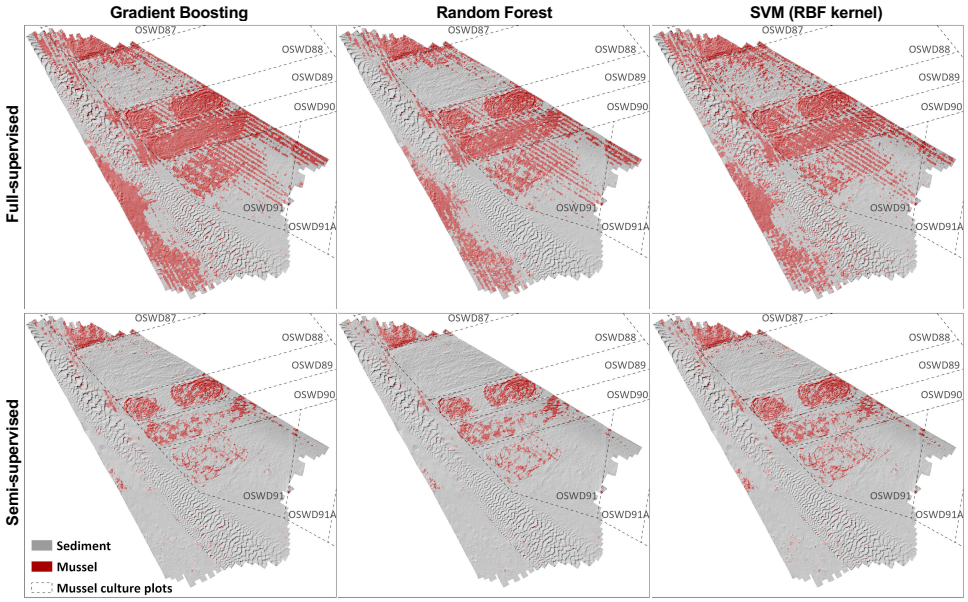


Figure 5.9: Prediction maps of the entire study area from the full- and semi-supervised classification methods, shaded by bathymetry.

from mussels in all three full-supervised predictions. Moreover, many parts of sand ripples are classified as mussels in GB and SVM predictions. Both the mussel cultivation area and sand ripples can be prominent in geomorphological features, but they show different characteristics in the backscatter-related features. These misclassified results indicate that the full-supervised models were unable to handle complex relationships among high-dimensional features and could be overfitted.

As for the semi-supervised methods, classification maps of GB, RF, and SVM show similar spatial patterns in general. The identified mussel class reveals the circular patterns of the mussel bottom culture in detail. Compared to the full-supervised results, misclassification in areas of sand ripples and the channel is largely reduced. Most striped noise is also avoided, except for a small amount in the east. The reported mussel production in culture plots OSWD87 (abundant), OSWD88 (empty), and OSWD89 (abundant) (see Section 5.2) is more consistent with the semi-supervised classification results, but to a lesser extent also with the full-supervised predictions.

From the semi-supervised learning results, we extracted an averaged ARC for both classes (Fig. 5.10). The three machine learning algorithms, GB, RF, and SVM, produced consistent ARC shapes. ARCs of the mussel class show larger BS values than the sediment class, especially in the outer beams. In the nadir, the "mussel" ARCs have a flatter shape than "sediment" ARCs. This indicates the possible impact of mussels on the seabed surface roughness. On the other hand, ARC shape features can be important predictors for identifying mussels on the sediment.

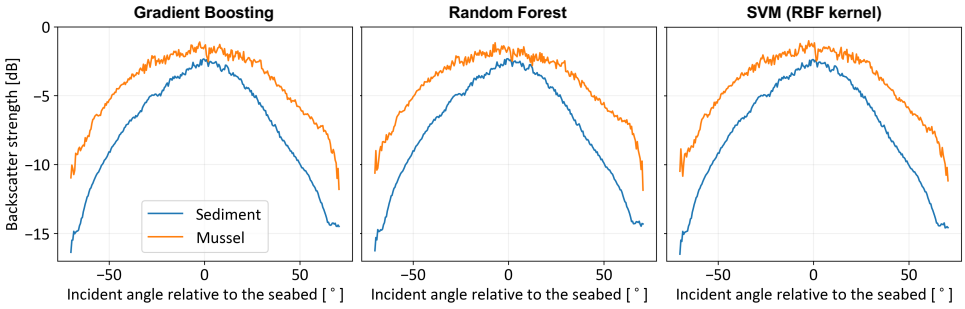


Figure 5.10: Averaged ARC of the sediment and mussel class from the semi-supervised learning results.

Feature importance in full- and semi-supervised classification

As explained in Section 5.3.4, feature importance in tree-based methods (GB and RF) measures the impact of different features on model performance. Comparing the feature importance between the full- and semi-supervised classification will help to assess if machine learning models handle multi-dimensional features better after adding pseudo-labels in training data.

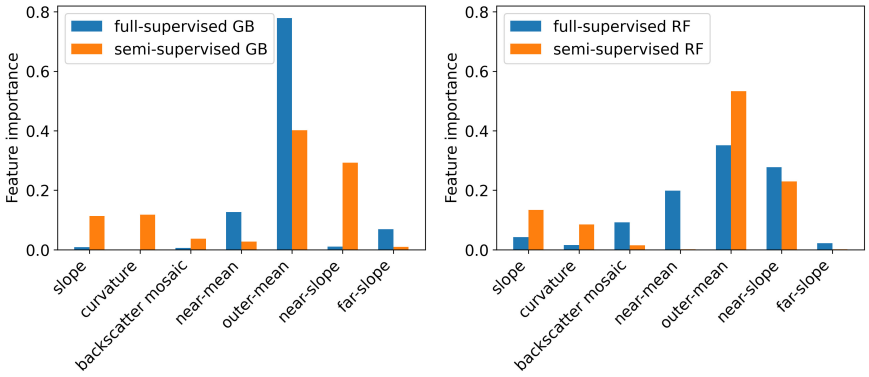


Figure 5.11: Comparison between feature importance of the full- and semi-supervised classification using GB and RF.

Regarding GB, the full-supervised classification depends mostly on outer-mean and near-mean (Fig. 5.11). Geomorphological features, which can provide spatial patterns of mussel cultivation, are overlooked. With limited training data (284 points), it is difficult to account for the interplay among all features. One important feature, including the noise in it, might dominate the classification process. Compared to GB, the importance of different features in RF is more balanced, and near-slope plays a bigger role. During the hyperparameter search of the supervised methods, fewer trees were selected for RF than GB. Simple models are less prone to overfitting, which might explain why RF makes better use of all features and brings less misclassification in the sand ripple area compared to GB.

After using pseudo-labels to enlarge the training data size, the feature importance of both GB and RF exhibit significantly improved consistency. Although outer-mean is still the most influential feature, the importance of slope and curvature increase in both GB and RF classification. Far-slope is found to be the least influential feature in both GB and RF. Semi-supervised methods involve exploiting the data distribution outside the ground truth data, which might help to gain a better understanding of the relations among features and mitigate the risk of the model becoming overfitted on certain features. From a spatial point of view, these pseudo-labeled data also provide information on the study area other than the seabed sampling stations.

Quantitative evaluation: LOSO cross-validation results

LOSO cross-validation helps to estimate the model performance with limited ground truth by considering different train-test splits. Through splitting based on locations of the ground truth sampling stations, it can also reduce the spatial dependency between the train and test data [42, 156].

Table 5.2: Averaged accuracy and F1 score of the LOSO cross-validation[1]

Accuracy	GB	RF	SVM (RBF kernel)
Semi-supervised	0.53 ± 0.000	0.51 ± 0.017	0.54 ± 0.009
Full-supervised	0.72	0.64	0.58
F1 score	**GB**	**RF**	**SVM (RBF kernel)**
Semi-supervised	0.57 ± 0.005	0.54 ± 0.022	0.57 ± 0.005
Full-supervised	0.78	0.72	0.66

Among the full-supervised learning methods, GB delivers the best accuracy and F1 score (Table 5.2). Regarding semi-supervised learning results, we performed label spreading and model training with three sets of randomly selected MBES data points. The accuracy values and F1 scores from these three experiments are generally consistent, with RF showing the largest fluctuations. Although misclassification and noise are shown to be reduced in the semi-supervised prediction maps, the cross-validation scores are lower than the full-supervised methods. This indicates that the full-supervised classification models are better fitted at the ground truth sampling stations.

Details regarding the predictions in LOSO cross-validation are also provided (see Fig. 5.12), further showing that semi-supervised classification using three different machine learning algorithms achieves consistent results. MBES points associated with one sampling station were kept as validation data in each round of LOSO cross-validation. When comparing the percentage of "mussel" predictions with the boxcore sample descriptions, GB delivers the most aligned results among the full-supervised classification methods. With "sediment" seabed samples 1 and 2 used for validation, full-supervised

[1]For semi-supervised methods, the mean and standard deviation of using 3 randomly sampled MBES point sets for label spreading are shown.

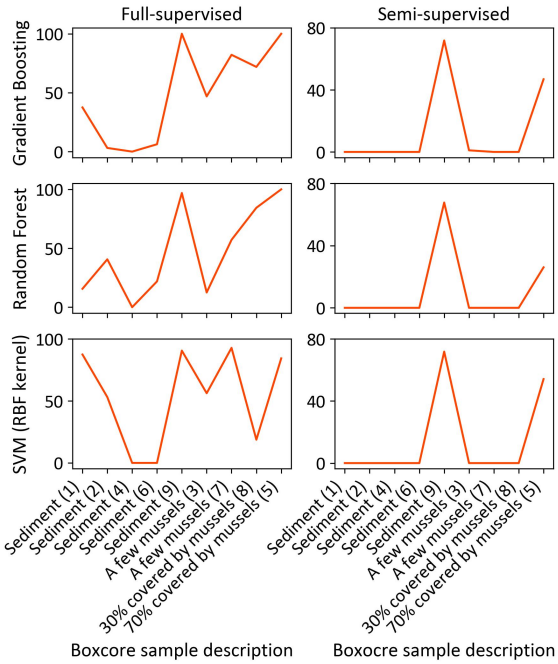


Figure 5.12: Percentage of points predicted as "mussel" in each round of LOSO cross-validation for (Left) full-supervised and (Right) semi-supervised classification. Descriptions of the ground truth sample reserved for validation in each round are presented and sorted based on the mussel coverage. The semi-supervised results are averaged values based on 3 sets of randomly sampled MBES points for label spreading.

methods tend to predict many more "mussel" presence than semi-supervised methods, especially when employing SVM. For the "sediment" sample 9, both full- and semi-supervised classification fail to deliver correct predictions. Although no mussels were found in the boxcore sample 9, this sampling station was located in the mussel culture plot OSWD91, indicating that mussels might exist in the surroundings. However, this is difficult to validate given no production recordings for plot OSWD91. Moreover, semi-supervised methods are not able to predict the mussel presence around sampling stations 3, 7, and 8 when they were defined as the validation set. Only semi-supervised GB predicts a few points as "mussel" near station 3. For stations 3 and 7, this might be attributed to the presence of a few mussels, according to descriptions based on the boxcore sample pictures. Sample 8 contained a certain amount of mussels. When it is not used for pseudo-labeling, critical information regarding the MBES features of the sediment-mussel mixture (Fig. 5.8) can be lost. For sample 5, both full- and semi-supervised classification indicate the mussel presence, but semi-supervised "mussel" predictions are less. In general, semi-supervised methods tend to predict the mussel occurrence more conservatively than full-supervised classification.

Comparison between the predictions and sample descriptions (Fig. 5.12) might show the limitation in quantitative evaluation using traditional seabed samples. With boxcore

samples that usually cover very small seabed areas, it is difficult to accurately label all MBES points around each sampling station. General descriptions based on visual inspection of the samples (like “a few mussels” and “30% covered by mussels”) are not reflected on the hard labels (“sediment” or “mussel”). Moreover, considering that only nine seabed samples were available in this research, the quantitative evaluation of classification for the entire study area is very challenging.

Although semi-supervised classification delivers lower validation scores at the ground truth sampling stations, the overall prediction performance within the study area proves to be better (Fig. 5.9). Combining knowledge from sources other than traditional bottom sampling is therefore important for assessing seabed habitat classification. For instance, the known spatial patterns of cultivated mussels in the Oosterschelde help the evaluation largely, which will be less easy to achieve in the case of natural mussel habitats. Information from environmental conditions and optical measurements at multiple spatial scales can be helpful for natural seabed habitat monitoring [34].

5.4.4. Limitations

In this study, we investigate the application of semi-supervised machine learning in seabed classification by presenting a specific case for identifying mussel cultures. Our study area represents the bottom mussel cultivation in the Netherlands but is relatively small. Moreover, although the surveyed area provides various seabed features regarding geomorphology, it shows a specific spatial pattern of mussel distributions and a preference for certain sediment types. This limited study scope poses challenges to estimating the performance of the proposed method on MBES data from other seabed environments. For instance, naturally growing mussels attach to harder substrates like rocks. They can also present distinct spatial patterns compared to cultivated mussels. Therefore, future work should investigate this semi-supervised method for diverse mussel habitats on different seabeds, by collecting acoustic data from more locations or combining open datasets. This will require enhanced collaboration among researchers and practitioners from various fields. Besides acoustic datasets, ground truth information other than point-based seabed sampling, such as underwater images with a broader spatial extent and field knowledge regarding mussel growth, can assist the training of robust machine learning models and ease the validation process. In addition, when comparing MBES datasets of different study areas, it is necessary to consider backscatter calibration or strategies for combining backscatter data from different surveys.

5.5. Conclusion

With MBES data and a limited number of ground truth bottom samples collected in a mussel cultivation area in the Oosterschelde, the Netherlands, we show the advantage of a semi-supervised machine learning approach for detecting the occurrence of mussels. We use label spreading to generate pseudo-labels, which helps to enlarge the training data size and alleviates overfitting. Trained with the pseudo-labeled data, the prediction maps of three machine learning algorithms (GB, RF, and SVM) show consistency. When using machine learning methods in specific applications, it is highly recommended to compare different algorithms and see if consistent results can be achieved.

Based on evaluation using the abundant ecological knowledge of the cultivation area, the mussel occurrence predicted by the semi-supervised classification method is in line with areas having large slopes and curvatures caused by dredging activities during mussel farming. However, seabed geomorphology cannot be the deterministic feature for mussel detection, since some mussel culture areas are empty and sand ripples can also induce varied geomorphology. In this case, the backscatter-related features are an important indicator. In general, mussels increase *BS* values of the seabed compared to the bare sediment. They may also affect the characteristics of the backscatter angular response curves, especially for the outer beams. During the evaluation stage, we also highlight the importance of studying the spatial patterns of mussels from methods other than traditional bottom sampling, especially in the research for natural mussel habitats.

6

Multi-frequency mapping of living and empty shellfish beds with semi-supervised and unsupervised learning

Across-track angular variations of the backscatter strength (BS) and usually limited overlaps between survey lines complicate the processing of multibeam echosounder (MBES) backscatter data for seabed characterization. In the previous chapter, we extracted angular response curve (ARC) shape features, achieving discriminative seabed information but sacrificing spatial resolution to some extent. To preserve both high spatial and angular density of BS data in benthic habitat mapping, this chapter introduces a two-stage workflow regarding multi-frequency MBES measurements. First, a semi-supervised graph-based completion method helps to generate full-coverage BS data for each incident angle, by leveraging the affinity of soundings regarding geolocations and seabed topography. Afterward, we cluster the resulting multi-frequency multi-angle BS data based on feature reduction using a spectral neural network, which accounts for complex non-linear relationships in high-dimensional data and improves the clustering performance compared to conventional linear reduction and manifold learning methods. Applied to MBES data collected at three frequencies (90, 300, and 450 kHz) for a shellfish habitat in the Dutch North Sea, our workflow is demonstrated effective in distinguishing living and empty shell beds from the sandy seabed and coarse sediments composed of shell fragments, demonstrating that semi-supervised and unsupervised learning can improve seabed habitat mapping and thus contribute to the support of sustainable offshore engineering activities.

This chapter is currently under journal review.

6.1. Introduction

Shellfish play an important role in benthic ecosystems, serving as the food source for many marine species [157]. Some species act as ecosystem engineers, altering the seabed environment and creating biodiversity hotspots [158]. Regular monitoring of the spatial distribution and abundance of shellfish is therefore essential for conserving benthic habitats and facilitating sustainable practices in coastal engineering.

Over the past decades, multibeam echosounders (MBES) have been widely used for seabed habitat mapping [97]. Unlike traditional bottom sampling, which provides sparse point measurements, MBES systems acquire bathymetry and acoustic backscatter data with continuous seabed coverage in a non-invasive way. Especially when occurring in sufficiently high densities, hard shells of bivalves may modify the sediment characteristics and thus induce variations in MBES backscatter, making MBES a promising tool for detecting the shellfish presence (Chapter 5). The spatial continuity of MBES data also allows for identifying boundaries of a shell bed formed by large-scale shellfish aggregations [159].

Recent multi-spectral MBES systems have further improved seabed habitat mapping [68]. By combining multiple frequencies, it is possible to overcome the ambiguity of the backscatter strength (BS) regarding sediment material sizes comparable to the acoustic wavelength and distinguish a wider range of seabed types [41]. Moreover, the penetration depth of acoustic signals varies with frequency [67]. This is particularly relevant for shellfish that bury beneath the sediment surface, such as cut trough shells (*Spisula subtruncata*) and razor shells (*Ensis sp.*) [159]. Multi-frequency backscatter data can thus help to characterize these buried aggregations more accurately.

Nevertheless, BS is highly affected by the incident angle. In practice, limited overlap between survey lines leads to angular and spatial inconsistencies in the across-track direction. To address this problem, angular normalization has been a common preprocessing method, which rescales BS values of most incident angles according to several reference angles and produces seamless backscatter mosaics [93]. An alternative method is to process backscatter data separately for each angle [80]. However, both methods often neglect or discard the backscatter angular information, which is an important descriptor of seabed type. On the other hand, angular range analysis preserves this information by treating an entire swath (or half of a swath) as the classification unit [30], with the cost of the reduced spatial resolution [36].

In addition, ground-truthing for MBES surveys, typically conducted by seabed sampling with boxcores [58] or underwater cameras [13, 160], can be both costly and labor-intensive. Seabed sampling is also potentially destructive to benthic habitats. These restrictions usually result in sparse ground truth samples with limited spatial coverages. This makes supervised MBES classification, which relies on large training data, impractical for large-scale habitat mapping.

To tackle these challenges, we propose a novel seabed classification workflow that improves MBES backscatter analysis by combining domain knowledge of MBES backscatter with modern machine learning and deep learning techniques. The workflow consists of two main components. First, we use a semi-supervised graph-based completion method to predict missing BS values between survey lines at each incidence angle. The graph is constructed using geographic locations, bathymetry, and local seabed

slope, which ensures spatial and topographic smoothness in the produced full-coverage backscatter maps. Second, the resulting multi-frequency multi-angle backscatter data are classified in an unsupervised fashion. We conduct feature reduction for the high-dimensional backscatter data using a spectral neural network. The network is trained with a spectral clustering objective to handle complex data structures and make different classes well-separated in the lower-dimensional space. To further enhance the network's ability to model backscatter variations across angles and frequencies, we also embed incident angle and frequency information in the network through positional encoding. A seabed habitat classification map is then achieved by K-means clustering on the reduced features.

The effectiveness of the proposed backscatter completion and classification method is demonstrated through case studies focusing on shellfish habitat mapping in the Dutch North Sea. The classification results reveal distinct acoustic signatures for living and empty shell beds of *S. subtruncata*, offering valuable insights for long-term monitoring of shellfish habitats using MBES data.

6.2. Related Work

6.2.1. Missing data completion in acoustic seabed mapping

Broadly speaking, missing data completion is a common pre- or post-processing step in seabed mapping. Simple spatial interpolation methods, such as bilinear and nearest neighbor interpolation, have been widely used for achieving consistent geolocations of multiple MBES features or gridding the output classification maps [161]. However, predicting missing values for unsurveyed areas with spatial details requires more advanced approaches. Kriging, as a geostatistical method that estimates values for unknown locations by a weighted average of neighboring observations based on both distance and spatial autocorrelation [162], has also been widely applied for multiple seabed mapping purposes. This includes applications like generating full-coverage MBES backscatter maps for certain incident angles [163] and interpolating measurements of sediment organic carbon content [46]. As an extension of Kriging, Co-Kriging incorporates other variables that are correlated with the target variable to improve the interpolation accuracy. Gaida *et al.* [163] indicated that Co-Kriging can further improve backscatter interpolation, with full-coverage bathymetry as a second variable.

A more direct use of such correlations between the target variable (sparse) and the full-coverage variable (dense) has also been explored. Ji *et al.* [13] filled in missing bathymetry data by training a Gradient Boosting regression model using spectral features of satellite imagery, based on co-located sparse and dense observations. Li *et al.* [164] trained a Random Forest (RF) model to predict sediment mud content from dense bathymetric features, and then applied Kriging to the prediction residuals. RF predictions and Kriging interpolated residuals were summed as the final results. This combined method outperformed machine learning regression or geostatistical interpolation alone. However, building such regression models requires a strong correlation between the sparse and dense variables, which is not always the case.

Instead of using a correlated dense variable, Misiuk and Brown [107] leveraged the collinearity between backscatter intensities across different incident angles to address data

gaps. They applied Multiple Imputation by Chained Equations (MICE), which iteratively selects subsets of predictors (e.g., backscatter from available incident angles) to generate predictions for the target variable (e.g., backscatter at the missing angle). The procedure is repeated until all predictors are visited and the algorithm converges. To achieve this, a prediction model is also required, such as the RF model used in their study, to perform each imputation step. Nevertheless, incorporating the geospatial relationships in MICE remains challenging as MICE has been designed to handle multivariate tabular data rather than spatially structured inputs [165].

6.2.2. Unsupervised learning in acoustic seabed mapping

Unsupervised classification is common in seabed mapping due to the challenges in collecting extensive seabed ground truth. It also helps to fully explore the acoustic data structure and identify nuanced differences in seabed properties. For example, Simons and Snellen [80] modeled beam-averaged backscatter data from a single incident angle as a sum of multiple Gaussian distributions, where each Gaussian represents a distinct seabed type. Gaida *et al.* [106] extended this Bayesian method to multi-frequency backscatter data by assessing the statistical significance of combining acoustic classes across frequency pairs.

Machine learning methods for unsupervised classification, usually known as clustering, have also been widely applied in acoustic seabed mapping. Previous research has applied K-means clustering [35, 47], hierarchical clustering (Chapter 3), and iterative self-organizing data analysis [45, 166, 167] to analyze spatial patterns of various features including bathymetric derivatives, backscatter intensities, and angular range features of backscatter.

To improve the clustering performance and remove redundant information from high-dimensional feature space, it is crucial to use an effective data reduction method. Principal component analysis (PCA) is still the most commonly used technique in acoustic seabed mapping [41, 120, 168]. Through eigenvalue decomposition of the data covariance matrix, PCA efficiently provides uncorrelated features that preserve most variance of the original data. However, PCA might discard low-variance yet informative data structures, and it is sensitive to outliers. Non-linear feature reduction methods, such as manifold learning, remain less explored in seabed characterization applications. While deep learning techniques like autoencoders have been applied to extract features in backscatter images, especially from side-scan sonars [81, 169], their application to multi-angle MBES backscatter data remains limited. This study is a step forward in addressing this gap.

6.3. Material and Dataset

6.3.1. Study area

To collect MBES data for a shell bed, we selected the study area (Fig. 6.1) based on shellfish population data gathered by the annual monitoring campaigns in the Dutch coastal zone [157]. Survey *Spisula* Bed (*Spi*B) was conducted during July 10–13, 2023, with its 1×3 km extent determined according to locations with abundant shellfish

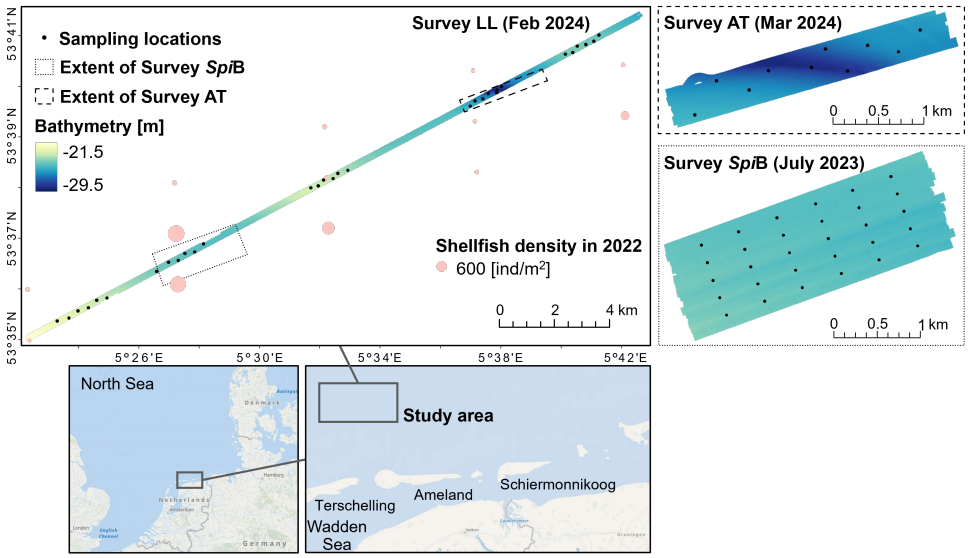


Figure 6.1: Location of the study area in the Dutch North Sea, displayed with the shellfish density in 2022 provided by Wageningen Marine Research and the extent of our three surveys. Bathymetry maps and seabed sampling locations are also presented for all surveys: Long Line (LL) conducted in February 2024, *Spisula* Bed (*Spi*B) in July 2023, and between Ameland and Terschelling (AT) in March 2024.

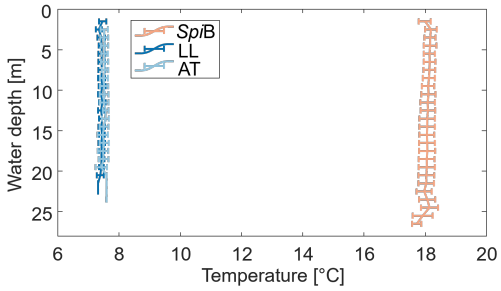


Figure 6.2: Seawater temperature profiles measured during all three surveys.

reported in recent years. A second survey, Long Line (LL), was carried out from February 19 to 22, 2024 along an extended transect (0.2 × 25.5 km) designed to include more seabed variations in MBES measurements. From west to east, LL coincided with the increase and then decrease of the shellfish density. It also includes a prominent trough, with the maximum water depth of 29.5 m. Compared to other surveyed regions, the trough also presents the most varying bathymetry. To increase the ground truth sampling density in this region, we conducted another survey, between Ameland and Terschelling (AT), with a 0.5 × 3 km extent, during March 12–13, 2024. Seasonal differences between the three surveys can be revealed by daily temperature measurements through the water column

(see Fig. 6.2). The seawater temperature difference reached 10°, between surveys in the summer (*Spi*B) and early spring (LL and AT). For the same water depth, temperatures measured in *Spi*B also showed a larger variation.

6.3.2. Spatial variation and seasonal difference of the seabed

We employed consistent bottom sampling techniques for all three surveys to collect seabed ground truth data. At each sampling location (Fig. 6.1), two bottom samples were collected using a cylindrical boxcore sampler with a surface area of 0.078 m^2. One sample was taken for sediment analysis, and the other for macrofauna. To have a broader inspection of the seabed, an underwater video was recorded at each sampling location just above the seabed, which covered a length of about 50 m.

Laboratory analyses of the boxcore samples indicate sandy sediments within the study area. The seabed covered by survey *Spi*B contained well-sorted sediments, with the median grain size (d_{50}) around 210 μm, while d_{50} reached 410 μm in the trough region covered by LL and AT (see Fig. 6.3). Living shells (%bivalve) were abundant in samples from *Spi*B (July 2023), indicating a shell bed region. By contrast, %bivalve of the same region in LL (Feb 2024) was much lower, corresponding to the shellfish seasonal life cycle. Abundant shellfish in *Spi*B also resulted in a higher weight percentage of dead complete shells (%empty shell) and fragmented shells (%shell fragments) than the same region in LL. Regarding the extended transect covered by LL, the trough region contained the most shell fragments. Considering the bathymetric variation within the study area and a lack of shell lag deposits in the shell bed boxcore samples from 2024, it is possible that both sediment and dead shell material were transported from the shell bed to nearby regions between the times of our surveys.

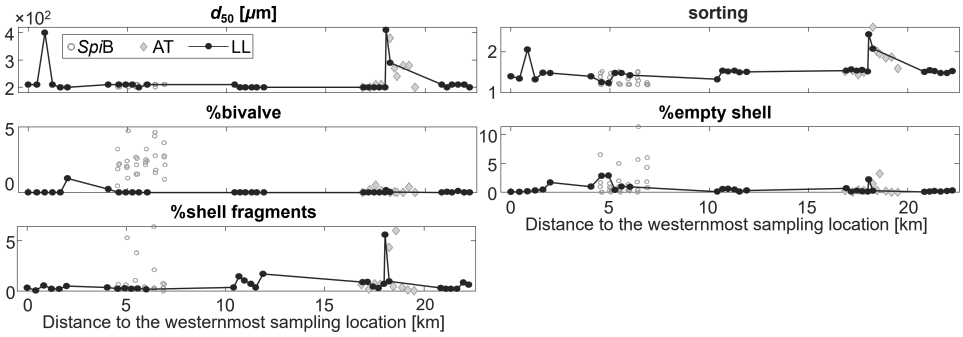


Figure 6.3: Median grain size of the sediment (d_{50}), sediment sorting (d_{60}/d_{10}), %bivalve, %empty shell, and %shell fragments of all sampling locations from boxcore sample analyses.

From each underwater video, we estimated an aggregate surface shell coverage level by visual inspections, dividing the sampling locations into five classes, which are ≈0%, <20%, 20%–50%, 50%–80%, and >80% (see Fig. 6.4). Living or dead shells were not distinguished in this analysis. In general, the shell bed region measured in survey *Spi*B contained the most shells on the sediment surface. Although there was a decrease in this region from July 2023 (*Spi*B) to February 2024 (LL), the amount of surface shells during

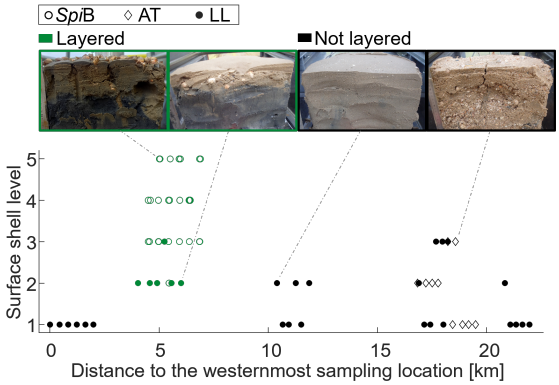


Figure 6.4: Surface shell coverage level of all sampling locations from underwater video analyses (1: ≈0%, 2: <20%, 3: 20%–50%, 4: 50%–80%, 5: >80%), displayed with visual inspection results of sediment layering based on the cross-sectional appearance of all boxcore samples.

LL was still classified as a medium level. Other sampling locations with shell coverage >20% were located in the trough region measured by LL or AT. Surface shells in the trough consist mostly of dead shell fragments. The accumulation of shell fragments here can also provide the habitat and tube-building material for *Lanice conchilega*. Although only a few living *L. conchilega* were found during the survey time, tubes protruding outside the sediment surface were still observed. We considered both tubes and dead shells when estimating the surface shell coverage level in this region.

To summarize the seabed volume properties in our study area, we also categorized the sampling locations into “Layered” or “Not layered”, based on cross-sectional pictures of the boxcore samples (Fig. 6.4). Compared to vertically homogeneous samples composed of sand or dead shells, layered samples were located in the shell bed region, containing shells just underneath the sediment surface and dark-toned anoxic sediment below the shell layer.

6.3.3. Acoustic dataset

During all three surveys, we acquired acoustic data using the multi-frequency multi-beam system R2Sonic 2026 (R2Sonic, Austin, TX, USA), which was operated at three frequencies: 90, 300, and 450 kHz. The corresponding nominal beam opening angles are 2.3°, 0.7°, and 0.5°, respectively. A swath coverage of 130° and the equiangular beam spacing mode were employed, resulting in 256 beams in each ping. R2Sonic 2026 emits acoustic signals at different frequencies subsequently, enabling the acquisition of nearly co-located multi-frequency measurements.

6.4. Methodology

We achieve shellfish habitat mapping with multi-frequency MBES measurements through the following steps: data pre-processing, backscatter completion and evaluation, feature

reduction with a neural network MBMF-SpectralNet, K-means clustering, and validation using seabed samples (see Fig. 6.5).

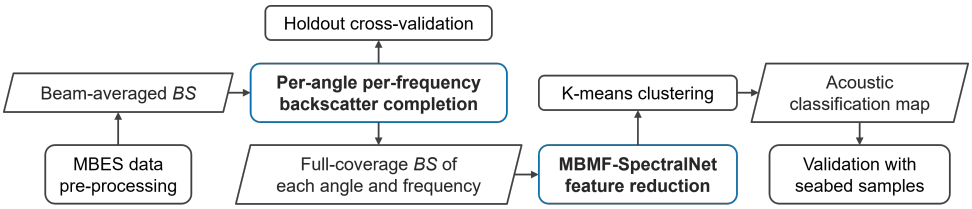


Figure 6.5: Workflow for shellfish habitat mapping with multi-frequency MBES data, with our proposed components (backscatter completion and deep learning-based feature reduction) highlighted.

6.4.1. MBES data pre-processing

The data pre-processing step comprises bathymetric cleaning, backscatter correction, and relative calibration (Section 2.2). In this study, we used the reference area same as that in Chapter 4 for relative backscatter calibration, which is located northeast of the Wadden Sea island Schiermonnikoog (Fig. 6.1). For each frequency, BS achieved by the MBES data pre-processing (Fig. 6.5) was aggregated into 129 angular bins, with the bin centers $\Phi = \{-64^\circ, -63^\circ, ..., 63^\circ, 64^\circ\}$. Incident angle $\theta \in (-65^\circ, 65^\circ)$ from the original MBES measurements was assigned to the bin with center $\phi \in \Phi$ if $|\phi - \theta| \leq 0.5^\circ$. $BS(\phi_j)$, $j \in \{1, 2, ..., 129\}$ for the jth bin is then calculated by

$$BS(\phi_j) = \frac{1}{M} \sum_{m=1}^{M} BS(\theta_m), \tag{6.1}$$

with M the number of θ belonging to the jth bin.

Due to the sandy sediments in our study areas, bathymetry measurements across the three frequencies do not differ much, showing limited differences regarding the penetration depth. We thus gridded bathymetry and slope data from 300 and 450 kHz using a grid size of 5 × 5 m, which were subsequently used in backscatter completion. For a fixed frequency and an incident angle ϕ, grid locations with $BS(\phi)$ values available in a 5-meter radius were filled using the bilinear interpolation results based on these surrounding $BS(\phi)$ values.

The resulting BS values for each incident angle and acoustic frequency serve as input for subsequent backscatter completion and clustering steps.

6.4.2. Backscatter completion per angle and frequency

To resolve across-track inconsistencies in BS that are inherent to incident angles ϕ, while preserving the angular information crucial for seabed mapping, we propose a backscatter completion step based on a semi-supervised graph-based method for each incident angle.

Due to the limited overlap between track lines, the across-track spatial density of MBES backscatter data for a single incident angle ϕ_1 is typically much lower than that

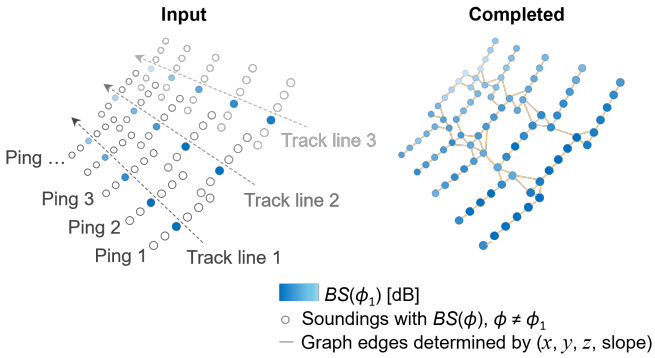


Figure 6.6: Illustration of backscatter completion for incident angle ϕ_1, given a fixed acoustic frequency. x and y [m] are the easting and northing coordinates, while z [m] indicates the bathymetry measurements in the Lowest Astronomical Tide (LAT) datum. This completion will be repeated for all incident angles ϕ.

in the along-track direction (see Fig. 6.6). As a result, conventional spatial interpolation methods such as bilinear interpolation often result in poor spatial continuity across tracks due to this anisotropic data distribution. By contrast, MBES bathymetry data are almost independent of incident angle and deliver continuous measurements for all soundings. Assuming that soundings close to each other in the three-dimensional (3D) space, defined by bathymetry (z) and geographical coordinates (x, y), tend to exhibit similar $BS(\phi_1)$ values, we can exploit this spatial affinity to estimate backscatter $BS(\phi_1)$ at unsampled locations. We achieve this by constructing a graph that captures the local similarities among soundings in the 3D space. Unlike Co-Kriging or other regression-based methods, our approach does not rely on a strong correlation between bathymetry and backscatter. In fact, this correlation does not always exist, considering, for example, that the occurrence of shell fragments or shellfish aggregation underneath the sediment surface might not affect bathymetry but only the backscatter measurements.

Algorithm details

Given n data points in a single-frequency MBES dataset $\mathscr{P} = \{p_1, ..., p_l, p_{l+1}, ..., p_n\} \subset \mathbb{R}^d$, each point is represented by a vector with a length of d. Only the first l points are associated with known backscatter measurements $\{BS_1(\phi), ..., BS_l(\phi)\}$ for $BS(\phi)$, with a specific incident angle $\phi \in (-90, 90)^\circ$. To ensure smoothness in the completed $BS(\phi)$ regarding seabed terrain features, we define $p_i = (x_i, y_i, z_i, \text{slope}_i)$, $i \in \{1, ..., n\}$. Each point p_i serves as a node in the graph $G = (\mathscr{P}, E, W)$, with the graph edges E representing connections between points and weighted by the adjacency matrix $W \in \mathbb{R}^{n \times n}$.

Our backscatter completion method is built upon label propagation [170], which is a semi-supervised learning algorithm predicting the label probability matrix of a partially labeled dataset using a graph constructed by features from both labeled and unlabeled data points. Unlike traditional label propagation, which predicts a probability matrix over multiple classes, our goal is to estimate a single continuous variable $v = BS(\phi)$ at a given incidence angle ϕ. In addition, to reduce the computational burden, we construct

an undirected symmetric k-Nearest Neighbors (NN) graph rather than a fully connected graph. The edge set E then equals $\{(i,j) \mid p_j \in \mathrm{NN}_k(p_i), i \neq j\}$, where $\mathrm{NN}_k(p_i)$ denotes the set of k nearest neighbors of point p_i based on the Euclidean distance in the feature space. The adjacency matrix W is then constructed using a Gaussian kernel as

$$W_{ij} = \begin{cases} \exp\left(-\dfrac{\|p_i - p_j\|^2}{2\sigma^2}\right), & (i,j) \in E \\ 0, & (i,j) \notin E, \end{cases} \tag{6.2}$$

where σ is a scaling parameter. A smaller σ will make the weight decay faster with the increasing distance. We further normalize W with its degree matrix D through

$$S = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}, \tag{6.3}$$

which is a more robust variant compared to $D^{-1}W$ used in the original label propagation algorithm [143]. Before the completion, we initialize v according to

$$v_i = \begin{cases} BS_i(\phi), & i \in \{1, ..., l\} \\ 0, & i \in \{l+1, ..., n\}. \end{cases} \tag{6.4}$$

We then achieve backscatter completion through iterating over the following two steps until v converges:

i) Backscatter propagation: $v \leftarrow Sv$,

ii) Hard clamping: $v_i = BS_i(\phi)$, $i \in \{1, ..., l\}$.

The propagation step transfers backscatter information to the unknown nodes from their connected nodes on the graph, which allows the estimation of $BS(\phi)$ at previously unsampled locations. To ensure fidelity to observed data, hard clamping is applied, which preserves the original $BS(\phi)$ measurements in the prediction. By repeating this completion for each incident angle ϕ in the acoustic dataset (see Section 6.4.1), we can obtain spatially dense backscatter data for the entire angular range.

Performance evaluation

We evaluate the performance of backscatter completion using holdout cross-validation. For a given single-frequency single-angle dataset $\mathscr{P}$, we randomly select m% of data points from $\{p_1, ...p_l\}$ and remove their backscatter measurements. After running the backscatter completion, the predicted $BS(\phi)$ values for the holdout points are compared to their actual measurements using the mean absolute error (MAE) as the evaluation metric. This process is repeated five times for each value of m to account for the sampling variability. By conducting this evaluation for different m ($10 \leq m \leq 99$), we investigate the effectiveness and robustness of our backscatter completion method. Furthermore, we compare the evaluation results among different incident angles and acoustic frequencies to understand the method's performance under different scenarios.

We also use this cross-validation with $m = 20$ for selecting the optimal hyperparameters k and σ in Eq. 6.2. When completing $BS(\phi)$, a larger k or σ can help to incorporate

neighborhood information but might cause over-smoothing effects. Therefore, the optimal k and σ can be determined through grid search, by selecting their values when the MAE stops decreasing with the increasing k or σ. For MBES datasets in this study, we selected $k = 10$ and $\sigma = 0.1$, using graph node vectors $\mathscr{P} = \{p_1, ..., p_l, p_{l+1}, ..., p_n\}$ normalized with the zero mean and unit variance.

6.4.3. Clustering with a spectral neural network

Clustering of the high-dimensional backscatter data collected across various incident angles and frequencies requires an effective feature reduction method. In this research, we employ spectral clustering, which is a popular graph-based clustering algorithm [171, 172] that can capture complex and non-linear relationships between data points. The key strength of spectral clustering lies in its spectral embedding, which projects data in a lower-dimensional space through eigen decomposition of the graph Laplacian matrix derived from pairwise similarities between data points. This spectral embedding serves as a powerful feature reduction technique that makes arbitrarily shaped clusters more clearly separated in the low-dimensional data space. The clustering is then performed using the K-means algorithm, typically with the number of clusters set equal to the dimension of the reduced feature space.

However, spectral embedding of large-scale datasets, such as geospatial data containing millions of points, can be computationally prohibitive and is therefore often impractical in real-world applications. To address this scalability issue, Shaham *et al.* [173] proposed SpectralNet, demonstrating that a deep neural network can be trained to approximate spectral embedding for large datasets. The network can be trained in a stochastic fashion by randomly sampling a minibatch of t data points in each iteration to reduce memory usage. The following loss is minimized during training:

$$\mathscr{L}(\omega) = \frac{1}{t^2} \sum_{i,j=1}^{t} W_{ij} \|y_i - y_j\|^2, \tag{6.5}$$

where ω denotes the learnable parameters of the neural network, $W \in \mathbb{R}^{t \times t}$ is the similarity matrix of points in the minibatch, and y_i is the reduced feature output by the neural network for the ith data point. The architecture of SpectralNet consists of a multi-layer perceptron (MLP), followed by a linear layer performing QR decomposition to orthogonalize the output features. After training, the learned features are clustered using the K-means algorithm.

Angular- and frequency-aware SpectralNet

We build our feature reduction method based on SpectralNet [173]. Given N_f frequencies and N_ϕ incident angles, each point in an MBES dataset is associated with $N_\phi \times N_f$ backscatter values after the backscatter completion step. Assuming the total number of points is n, we then have a 2D input data matrix $X \in \mathbb{R}^{n \times (N_\phi \times N_f)}$ for the clustering step. However, training directly on X may not explicitly utilize the source information (incident angle and frequency) of each MBES point. Thus, to facilitate learning data structure regarding angles and frequencies, we further apply positional encoding to the input features. Originally introduced in natural language processing, positional encoding [174]

injects information about token order into neural networks by adding fixed, position-dependent sinusoidal vectors to word embeddings. Inspired by this idea, we add a fixed encoding that depends on the source information to each backscatter channel in X. This angular- and frequency-aware representation encourages SpectralNet to capture characteristics of BS variations across angles and frequencies, which empirically improves the clustering performance.

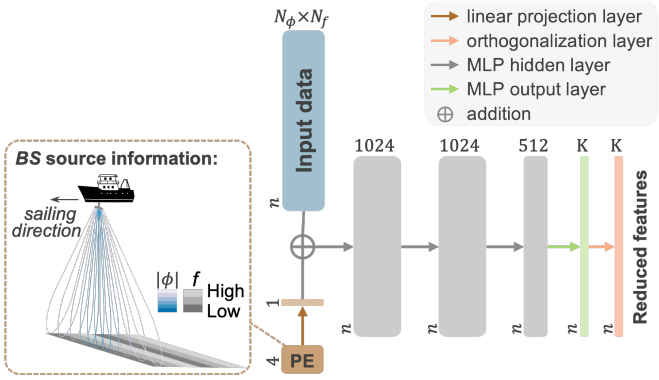


Figure 6.7: Architecture of MBMF-SpectralNet, with n, N_ϕ, N_f, K the number of MBES points, incident angles, acoustic frequencies, and final clusters, respectively. PE represents the positional encoding module accounting for the source (incident angle ϕ and frequency f) of each BS feature.

Given backscatter values BS at incident angles $\phi \in (-90, 90)^\circ$, we encode their unique angular position by a 2D vector $\mathbf{p}(\phi) = [\sin\phi, \cos\phi]$. For each acoustic frequency f [kHz], we define its positional encoding as $\mathbf{p}(f) = [f', f'^2]$, with f' the normalized frequency based on the mean and standard deviation of all available frequencies. We concatenate the angular and frequency encoding into a single 4D vector $[\mathbf{p}(\phi), \mathbf{p}(f)]$, which is then passed through a linear projection layer to yield a scalar value. This scalar is added to the corresponding column of X (see Fig. 6.7), which effectively injects structured positional information into the feature space. We refer to the resulting neural network, which combines the multibeam multi-frequency source information, as MBMF-SpectralNet.

Performance evaluation

We evaluate the clustering performance using the Silhouette coefficient [132], which quantifies the compactness and separation of clusters. The coefficient ranges from -1 to 1, with values close to 1 indicating compact and well-separated clusters. Values near 0 indicate overlapping clusters, and negative values might suggest incorrectly clustered points. Moreover, the Silhouette coefficient can be used to select the optimal number of clusters.

In addition, we validate our clustering results using the ground truth seabed samples. Clusters are compared to each sample property described in Section 6.3.2. To investigate whether the sample property of one cluster is significantly different from the other clusters, we conduct a Mann-Whitney U-test, which is a non-parametric statistical test

suitable for semi-quantitative variables [175]. For a specific cluster and sample property, the Mann-Whitney U-test returns the u-statistic, which measures whether the property value of that cluster is overall higher or lower than the other clusters. The measured relationship can be considered significant if the output p-value fulfills $p < \alpha = 0.05$. Since the u-statistic is affected by the number of samples, we calculate the rank-biserial u [176], which accounts for the sample size and ranges from -1 to 1. Values smaller than 0 indicate the sample property of the tested cluster is lower than the other clusters, while positive values suggest the opposite relationship.

We also apply the non-metric Multi-Dimensional Scaling (nMDS) ordination to visualize the overall structure of all seabed samples. nMDS is a commonly used technique in ecology to represent all sample properties in the 2D space [177]. The resulting nMDS biplot illustrates both the similarity between sampling locations and the correlations among sample properties. We then qualitatively assess whether our clustering results align with the sample structure indicated by the nMDS biplot.

6.5. Results and Discussion

6.5.1. Backscatter completion

In general, our backscatter completion remains robust until the holdout data percentage reaches around 70% (see Fig. 6.8). Moreover, the holdout cross-validation generally indicates a decreasing MAE with increasing incident angles. This trend may be attributed to the uncertainties of MBES backscatter measurements. Due to beamforming, oblique beams typically have a larger beam opening angle than nadir beams, bringing more scatter pixels (footprint of one pulse) in the beam footprint [80] and thus reducing the BS uncertainty. For a fixed incident angle, the MAE of 90 kHz is lower than the higher frequencies, which might also be due to a larger beam opening angle (Sec. 6.3.3).

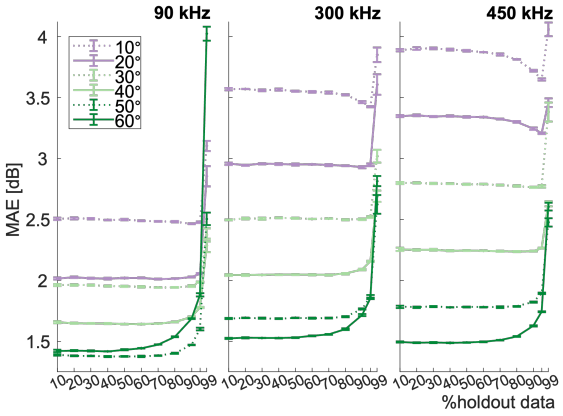


Figure 6.8: Mean absolute error (MAE) in dB resulting from the holdout cross-validation of $BS(\phi)$ completion for MBES data from survey LL, with $k = 10$, $\sigma = 0.1$, for holdout percentages ranging from 10% to 99%. Results of incident angles $\phi \in \{10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ\}$ are presented in color for (Left) 90 kHz, (Middle) 300 kHz, and (Right) 450 kHz.

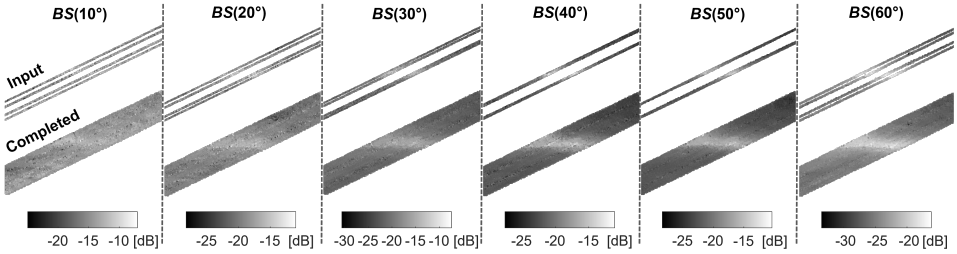


Figure 6.9: Comparison between input and completed $BS(\phi)$ for 300 kHz regarding data from the trough region in survey LL, for incident angles $\phi \in \{10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ\}$.

Maps of completed backscatter data at oblique angles also show clearer patterns and better spatial continuity compared to nadir beams (see Fig. 6.9). This improvement is partly due to more homogeneous and noisy input BS data at nadir, which do not reveal the difference between the deepest trough and its surrounding seabed environment. Completion results for $BS(10^\circ)$ and $BS(20^\circ)$ also present more artifacts compared to larger incident angles. The drop in MAE for holdout percentages larger than 80% regarding 10°–30° might indicate over-smoothing of the completed data instead of a further performance boost. Thus, we only adopted completed $BS(\phi)$ with $|\phi| \geq 30^\circ$ in the subsequent clustering analysis. Regarding 300 and 450 kHz, the MAE values for incident angles 30° and 40° exceed 2 dB, which indicates a limited ability to discern different seabed types [18]. Nevertheless, BS measurements at these angles might still contribute to acoustic classification by providing information on backscatter angular variation.

6.5.2. Clustering with MBMF-SpectralNet

We compared clustering performances of the K-means algorithm based on five feature reduction techniques are compared (see Fig. 6.10): linear reduction with PCA, two popular manifold learning methods including t-distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP), the original SpectralNet, and MBMF-SpectralNet introduced in Sec. 6.4.3. To determine the optimal number of clusters, Silhouette coefficients of 4 to 10 clusters were investigated.

In general, MBMF-SpectralNet yields the highest Silhouette coefficient results for 5–8 clusters. For 4 or 5 clusters, its performance is similar to that of SpectralNet; however, with 6 or 7 clusters, a largely improved Silhouette coefficient is achieved by MBMF-SpectralNet. For 4 to 7 clusters, MBMF-SpectralNet outperforms PCA, UMAP, and t-SNE. With only linear transformation, PCA might not reveal complex and non-linear patterns in data. While UMAP and t-SNE are non-linear, they appear more sensitive to noise, which is also observed from maps of their reduced features. With 8–10 clusters, clustering performances based on PCA, SpectralNet, and MBMF-SpectralNet are more comparable, though all of them exhibit a steadily decreasing trend.

We selected seven as the optimal number of clusters. Although 4 clusters result in the highest Silhouette coefficient, the metric plateaus at 7 clusters regarding MBMF-

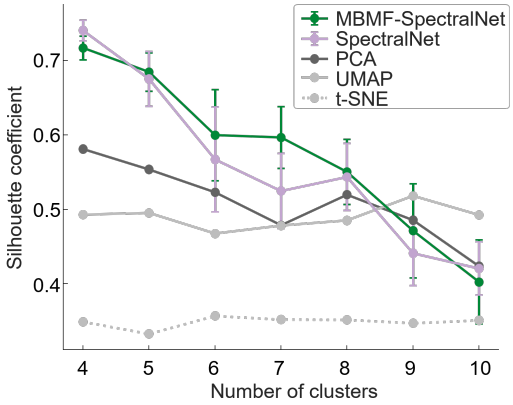


Figure 6.10: Comparison among the Silhouette coefficient for K-means clustering results based on feature reduction using PCA, t-SNE, UMAP, SpectralNet, and MBMF-SpectralNet, with the number of clusters ranging from 4 to 10. Considering the stochastic property of neural network training, the mean and standard deviation from 5 runs are shown for SpectralNet and MBMF-SpectralNet.

SpectralNet, which suggests a stable clustering structure. Furthermore, using 7 clusters may allow for revealing nuanced differences between seabed types.

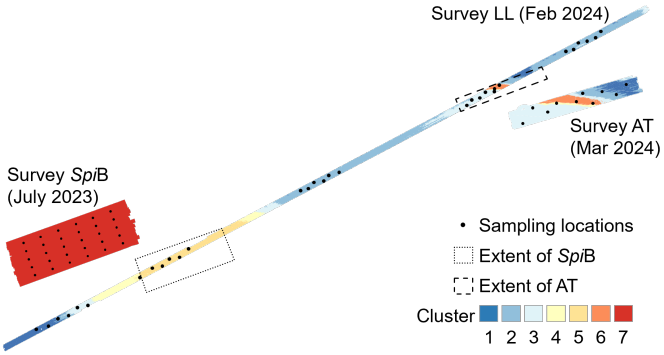


Figure 6.11: K-means clustering maps for seven clusters with MBMF-SpectralNet, based on completed backscatter data from all three surveys.

The 7-cluster maps reveal distinct geoacoustic properties of the shell bed region measured in survey *Spi*B (see Fig. 6.11). Moreover, consistent spatial patterns are observed for the same seabed area covered by survey LL and AT, which is expected since AT was conducted only one month after LL. For the extended seabed measured in LL, the shell bed and the trough region are classified as two different seabed types (clusters 5 and 6). Gradual transitions between clusters are also noticed adjacent to both regions, which suggests continuous changes in seabed properties.

To investigate the contribution of different acoustic frequencies to the clustering results, we calculated an average *BS* angular response curve (ARC) for each cluster (see

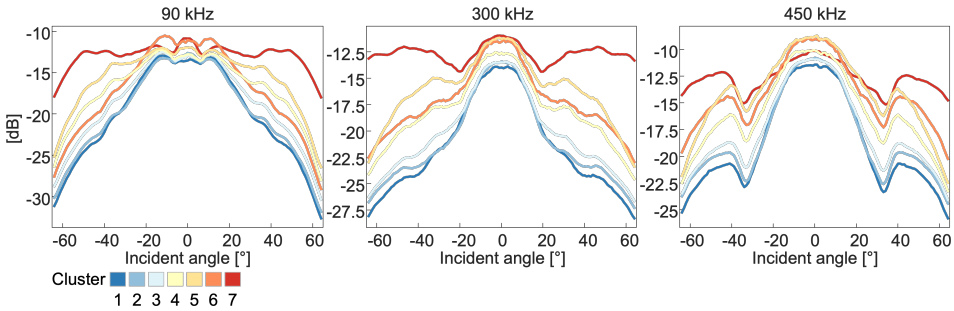


Figure 6.12: Average *BS* angular response curves of each MBMF-SpectralNet cluster for (Left) 90, (Middle) 300, and (Right) 450 kHz.

Fig. 6.12). A generally ascending trend can be noticed from cluster 1 to 3, regarding *BS* of oblique beams for all three frequencies. While clusters 1 and 2 show distinctions in outer beams of 300 and 450 kHz, their separation is less pronounced at 90 kHz, possibly indicating the difference in sediment roughness. By contrast, *BS* of 300 and 450 kHz from cluster 3 show higher values than cluster 2 in the middle angular range, while the difference lies in all incident angles larger than 20° for 90 kHz. This pattern may indicate an increase in both sediment roughness and volume scattering from cluster 2 to 3 [38].

On the other hand, relations among ARCs of clusters 4 to 7, which show generally higher *BS* than clusters 1 to 3, are more complicated and vary more with frequency. For 450 kHz, *BS* of cluster 7 is similar to cluster 5 at 20°–40°, and lower than clusters 5 and 6 in the nadir. By contrast, cluster 7 shows significantly higher *BS* than the other clusters in almost the whole angular range for 90 and 300 kHz. Cluster 7 may represent coarse seabed material that can induce ambiguity in *BS* for frequencies over 300 kHz [60].

For cluster 6, *BS* is lower than that of cluster 5 across 10°–60° at 300 kHz, and lower than both clusters 4 and 5 at angles larger than around 25° for 90 kHz. Since acoustic signals of lower frequencies are more susceptible to seabed volume properties [18], clusters 4 and 5 might indicate more volume scattering or layering of the seabed. On the other hand, the seabed covered by cluster 6 might induce more roughness scattering for the highest frequency, considering the higher *BS* than clusters 4 and 5 in outer beams of 450 kHz.

Moreover, *BS* in the nadir shows limited separations among 7 clusters, especially for 90 kHz. This might be due to the larger *BS* uncertainty, but can also stem from our choice of excluding nadir incident angles in the clustering process.

6.5.3. Validation of clustering results with seabed samples

When comparing our clustering results with seabed properties at the ground truth sampling locations, clear distinctions can be noticed among clusters 5 to 7 (see Fig. 6.13).Although clusters 5 and 7 cover similar spatial extents, they are derived from *BS* measurements taken in different seasons (Fig. 6.11). Both clusters correspond to similar d_{50}, though cluster 7 exhibits much higher %bivalve and surface shell coverage than

cluster 5. A wider spread in dead shell content (%empty shell and %shell fragments) is also observed regarding cluster 7. The Mann-Whitney U-test (Table 6.1) confirms the significance of observing the highest %bivalve and surface shell level in cluster 7 among all clusters, with $u > 0.9$ and $p < 0.05$. A high shell content can positively affect sediment hardness and roughness, which is consistent with the ARC characteristics of cluster 7 (Fig. 6.12). Moreover, cluster 7 is associated with significantly lower sorting values, suggesting better-sorted sediments in the living shellfish habitat.

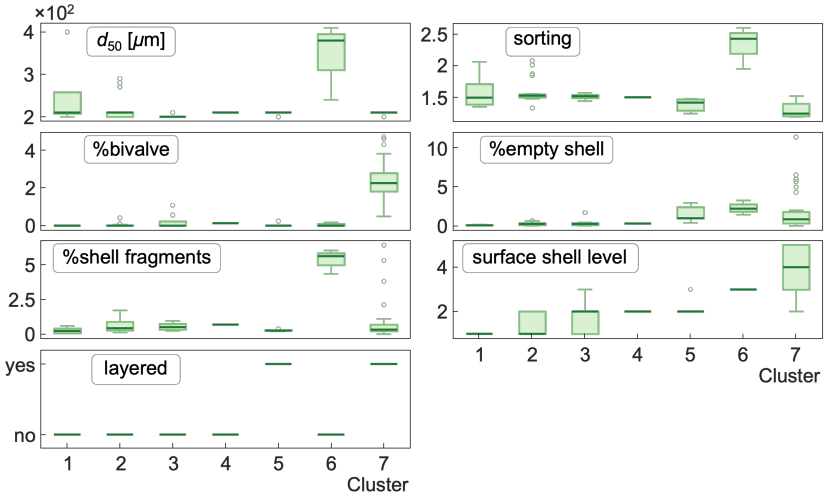


Figure 6.13: Comparison between MBMF-SpectralNet clustering results and each sample property.

Cluster 5 appears to exhibit medium %empty shell within the study area, as indicated by both the visual comparison (Fig. 6.13) and the absolute u-statistic around 0.5 (Table 6.1). Considering the same sediment layering of clusters 5 and 7, the spatial extent of cluster 5 might indicate the boundaries of the empty shell bed. Although no bottom samples were taken in the coverage of cluster 4, adjacent to cluster 5 (Fig. 6.11), bringing difficulties in validation, different ARC characteristics between clusters 4 and 5 (Fig. 6.12) can help to draw this conclusion. "Layered" sediments and the reduced cover of surface shells of cluster 5 may also explain the possibly higher volume scattering than cluster 6, as indicated by their ARCs.

In addition, seabed samples from the trough region (cluster 6) have significantly higher d_{50}, sorting, %empty shell, and %shell fragments compared to the other samples. As shown by the "Not layered" boxcore sample collected in survey AT (Fig. 6.4), cluster 6 can correspond to the coarsest sediment with dead shell accumulations. Nevertheless, surface shells in cluster 7 might still induce stronger acoustic scattering than coarse sediments in cluster 6.

Clusters 1 to 4 correspond to generally lower %empty shell and surface shell level than clusters 5 to 7 (Fig. 6.13). However, significant results are only revealed for clusters 1 and 2 (Table 6.1). Clusters 1 and 2 can also indicate significantly lower %bivalve. Regarding cluster 4, since only one sampling location from survey AT falls in its extent, insignificant

Table 6.1: Rank-biserial u-statistics and p-values from the Mann-Whitney U-test for each sample property except layered, under the null hypothesis H_0: two independent samples, randomly selected from one cluster and the other clusters, come from identical distributions. Significant results with $p \leq 0.05$ are highlighted.

Cluster label	1		2		3		4		5		6		7	
Number of samples	4		18		9		1		6		3		30	
	u	p	u	p	u	p	u	p	u	p	u	p	u	p
d_{50}	0.198	0.467	-0.038	0.795	**-0.645**	**< 0.001**	0.229	0.682	-0.085	0.711	**0.941**	**0.002**	0.083	0.510
sorting	0.299	0.323	**0.534**	**< 0.001**	**0.434**	**0.037**	0.186	0.769	-0.256	0.305	**0.971**	**0.005**	**-0.766**	**< 0.001**
%bivalve	**-0.657**	**0.025**	**-0.592**	**< 0.001**	-0.369	0.068	-0.029	0.980	**-0.554**	**0.022**	-0.294	0.386	**0.995**	**< 0.001**
%empty shell	**-0.731**	**0.015**	**-0.400**	**0.018**	-0.319	0.126	-0.286	0.643	**0.513**	**0.040**	**0.725**	**0.036**	**0.348**	**0.013**
%shell fragments	-0.425	0.159	0.182	0.253	0.122	0.563	0.400	0.510	-0.426	0.088	**0.961**	**0.005**	-0.151	0.282
surface shell level	**-0.750**	**0.011**	**-0.739**	**< 0.001**	-0.371	0.069	-0.194	0.754	-0.129	0.602	0.262	0.443	**0.929**	**< 0.001**

statistical test results are observed. On the other hand, cluster 3 presents lower d_{50} than the other clusters, which is consistent with the higher volume scattering suggested by ARCs of clusters 1 to 3.

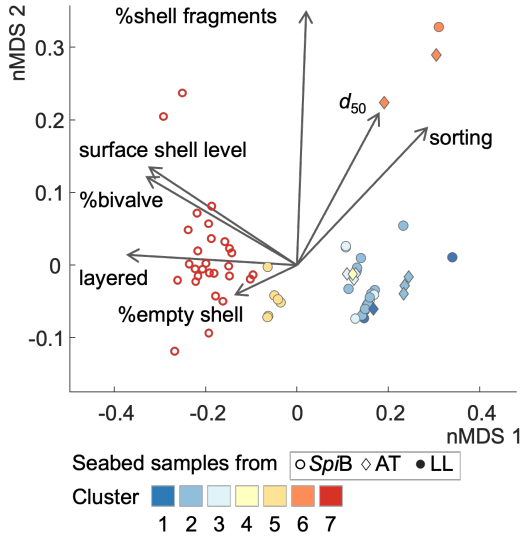


Figure 6.14: Comparison between MBMF-SpectralNet clustering results and ground truth sample properties using the nMDS biplot. Angles between sample property vectors indicate their correlations. Seabed samples with similar properties can be aggregated on the plot.

Apart from analyses for each sample property, the nMDS biplot presents the consistency between the overall sample community structure and our clustering results (see Fig. 6.14). Correlations between sample properties roughly indicate two types of variations within all samples: one is linked to the shell bed properties (%bivalve, %empty shell, and layered) and the other is more linked to the sediment coarseness (d_{50}, sorting, and %shell fragments). The surface shell level is strongly correlated with the living shell content in our dataset, although it can also be affected by shell fragments. Variations in these two different directions can be driven by the extreme property values of seabed samples from *Spi*B (cluster 7) and the trough region covered by LL and AT (cluster 6).

Collected within the extent of *Spi*B, seabed samples from LL belonging to cluster 5 are close to the *Spi*B samples on the biplot. By contrast, samples in clusters 1 to 4 can be seen as a more mixed group. Nevertheless, we can still notice different orientations of the sample groups indicated by clusters 1 to 3 on the biplot. In general, clusters 1 to 3 might indicate soft sediment modified by the shell material to slightly different extents. The one sample in cluster 4 is also close to samples in cluster 3 on the biplot. Comparison between the grouping of samples and *BS* clustering results might be more prominent if the sampling density within the coverage of clusters 1, 3, and 4 was higher.

In summary, comparisons with both individual bottom sample properties and the comprehensive biplot confirm the significant relations between our acoustic classification results and seabed characteristics. This helps to reveal the geophysical and

biological meaning represented by variations in the multi-frequency MBES data.

6.6. Conclusion

In recent years, MBES backscatter measurements have been commonly considered as proxies for seabed properties. However, two remaining challenges hinder the use of MBES backscatter in benthic habitat mapping: the across-track spatial inconsistency caused by different incident angles and the difficulty in acquiring a large amount of seabed ground truth data. This research addresses both challenges by incorporating semi-supervised and unsupervised learning methods in acoustic seabed mapping.

We preserve the high spatial and angular density of MBES backscatter data using a semi-supervised completion method. By leveraging the graph structure and determining connectivity between data points using their geographic locations and local seabed topography, sparse backscatter input from a single incident angle can be completed as a full-coverage and dense map. We mitigate the second challenge by backscatter clustering with a deep learning-based feature reduction method, MBMF-SpectralNet, which combines the spectral clustering objective with the angle and frequency information of each backscatter feature embedded by positional encoding. With multi-frequency MBES data from three surveys conducted in the Dutch North Sea, we show that the resulting low-dimensional features ensure the separation among clusters, improving the K-means clustering performance compared to the conventional linear reduction method PCA, manifold learning techniques such as t-SNE and UMAP, and the original SpectralNet.

We also validate the effectiveness of our seabed classification workflow using ground truth bottom samples. The clustering results align well with the sample community structure, revealing distinct acoustic signatures for soft sandy sediments, coarse dead shell accumulations, an empty shell bed beneath the sediment surface, and a living shell bed covered by abundant shellfish. Through case studies for a specific shellfish habitat, we indicate how artificial intelligence technologies can facilitate acoustic data processing and seabed mapping, even without large training data. Future research may focus on different types of benthic habitats and MBES seabed monitoring with a higher temporal resolution.

7

Recommendations for MBES surveys in marine benthos mapping

Based on the measurements and analyses in previous chapters, this chapter summarizes practical recommendations for future MBES surveys aimed at mapping marine benthos, such as shellfish beds in the Dutch North Sea. To ensure that the acoustic data taken at different times (or seasons) can accurately represent the habitat characteristics and facilitate robust data-driven benthic habitat mapping, a number of survey-design factors must be addressed. Three aspects are discussed in this chapter, including the configuration of multibeam echosounder (MBES) systems, strategies for a cost-effective sonar calibration, and recommendations for the seabed ground truth data collection.

7.1. MBES system

Compared to other sonar systems, the multibeam echosounder (MBES) provides a good balance between spatial coverage, efficiency in data collection, and complexity in geo-referencing. MBES systems are therefore commonly considered in regular surveys for seabed monitoring. For mapping the occurrence of marine benthos, not only the bathymetric variations can be linked to specific habitats, but the backscatter strength (BS) is an informative indicator for the influence of benthic organisms on seabed surface and volume characteristics. Such influence might not be fully revealed by BS at a single frequency, as demonstrated by case studies in Chapter 3 and 6, making it beneficial to employ a multi-spectral system.

Operating at multiple frequencies offers the advantage of resolving BS ambiguities that may occur at a single frequency when differentiating between scattering objects that are comparable in size to or larger than the acoustic wavelength. Furthermore, the use of lower frequencies, which may provide greater penetration depths, can assist in identifying "acoustically hard" marine benthos that are buried within the sediment body. While this thesis primarily utilized the MBES system R2Sonic 2026 for multi-spectral surveys, the general practical considerations are relevant to other systems with similar operational principles. A usable configuration of the R2Sonic 2026 is summarized in Table 7.1.

Table 7.1: Setup of the MBES system R2Sonic 2026 in this thesis

Center frequencies	90, 300, 450 kHz
Along-track beam opening angle	2.3°, 0.7°, 0.5°
Across-track beam opening angle at nadir	2.3°, 0.7°, 0.5°
Swath coverage	130° (-65°/+65°)
Beam spacing mode	Equiangular
Number of beams per ping	256
Nominal pulse length	150 μs
Pulse type	Shaped continuous wave
Maximum ping rate	60 Hz

With the operating frequency range of 90–450 kHz provided by R2Sonic 2026, it is recommended to select several center frequencies that are widely spaced to achieve distinctions among frequencies. For example, similar acoustic classification results of 200 and 300 kHz were found for our study areas in Chapter 3, which is why 200 kHz was not adopted in subsequent seabed surveys. In addition to reducing redundant information, this helps to obtain a higher spatial density of soundings for each remaining frequency. Due to the sequential emission of different frequencies in a multi-spectral system, a high frequency density results in increased distance between pings at an individual frequency. Moreover, when conducting repeated surveys in the same study area or multiple surveys to encompass a research zone, it is crucial to keep the sonar settings consistent.

7.2. Relative calibration

When aiming at a joint use of BS data acquired from various MBES surveys, absolute or relative sonar calibration strategies are necessary to ensure the data comparability across surveys. For an MBES, drift in the transducer sensitivity or beam patterns can happen over time due to aging components. Recently, water temperature was also found to be a potential factor affecting the transducer mechanical characteristics, resulting in an instrumental temperature dependency of the BS measurements [71]. Chapter 4 presented a relevant study that emphasized the importance of accounting for this influence when comparing BS datasets collected with the same MBES system across different seasons. Although absolute backscatter calibration can be an ideal solution, strict requirements for the equipment and controlled water environment bring difficulties in making it a routine procedure.

Relative backscatter calibration, by contrast, may provide a cost-effective alternative. As explained in Chapter 4, the relative calibration can be achieved by using a homogeneous and stable reference area. BS data collected within this area should present homogeneity in successive surveys. A relatively flat seabed is desired, with simple unidirectional geomorphological patterns being acceptable. Stability of the reference area should be represented by long-term bathymetric measurements and better validated with seabed ground truth information. Incorporating a reference area into every survey ensures an up-to-date calibration dataset; however, when study areas lie far from the reference area, the additional transit time can become prohibitive.

Near the Dutch North Sea, well-known and well-established reference areas exist in the Belgian coastal regions, such as the Kwinte area [70], which has already been used by several hydrographic agencies. A practical workflow is to visit the Kwinte area several times per year to capture a range of seawater temperatures in different seasons, using the identical sonar settings as the main surveys. The acquired data make it possible to construct an empirical BS-temperature catalogue, for each operating frequency and various angular ranges. In the absence of an absolutely calibrated system, one may select a convenient reference temperature (T_{ref}) and designate its associated BS measurements of the reference area (BS_{ref}) as the baseline for all subsequent relative calibrations. When a direct visit to the reference area is impractical during the survey period, the developed BS-temperature relationship and BS_{ref} can be utilized for relative backscatter calibration based on *in-situ* measurements of seawater temperature.

7.3. Ground truth data collection

In benthic habitat mapping, seabed ground truth data provide information on sediment composition, benthos occurrence, and presence of shell material and stones. They are essential for validating the map products or interpreting the acoustic backscatter data, especially when absolute backscatter calibration is unavailable. In many cases, the ground truth information is also needed to train and evaluate the data-driven models.

Ground-truthing methods

Along with acoustic surveys, bottom samples gathered by a grab sampler or boxcorer are always used as the seabed ground truth. Compared with grabs, cylindrical boxcores also

allow inspection of layering in the upper 25–30 cm of the sediment body and comparison of seabed compositions between different layers.

For the soft sediments that we encountered in the Dutch North Sea, underwater cameras are a valuable addition for ground-truthing. A drop camera system towed near the seabed can record footage over relatively long distances (about 50 m in this thesis), providing a much broader view of seabed conditions than point-based bottom sampling. However, the image quality is strongly affected by weather and sea state. For fine sediments, wind-driven waves may increase water turbidity near the bottom and degrade visibility. Moreover, underwater cameras cannot reveal what lies within the sediment body, such as buried shellfish beds. In these cases, bottom samples remain necessary.

By contrast, for hard substrates such as rocks, underwater cameras may be the only practical ground-truthing tool, since collecting physical samples is extremely difficult. They are also essential when destructive sampling is restricted, for example, within the Marine Protected Areas (MPAs).

Localization of seabed ground truth

Accurately geo-referencing bottom samples and underwater footage is challenging in the aquatic environment. A boxcorer is typically lowered from the vessel via a cable. In practice, the sample location on the seabed is approximated based on the vessel's Global Navigation Satellite System (GNSS) position and the known offset to the deployment position onboard. Regarding water depths considered in this thesis, which show the maximum value of 25–30 m, this approximation is generally acceptable. In deeper waters, however, the layback of the longer wire and the wire motion in currents can shift the sample location horizontally to a much larger extent. Moreover, in highly heterogeneous benthic habitats, such as the patchy shellfish aggregations, improved localization accuracy of ground truth samples can increase the reliability of interpreting the corresponding acoustic measurements.

The same limitation also applies to drop camera systems. While video footage can be approximately localized using the vessel track, accurately matching individual frames to the acoustic data remains challenging. As a result, Chapter 6 treated each video transect as a single unit and derived an aggregate metric for the entire transect. If the spatial variation within a transect could be resolved more precisely, it would not only improve frame-level correspondence with the acoustic data but also effectively increase the number of ground-truth observations.

To achieve higher localization accuracy, many underwater surveys employ an Ultra-Short BaseLine (USBL) acoustic positioning system, which consists of a transceiver mounted on the vessel and a transponder attached to the underwater device. The transceiver contains an array of transducers. By measuring the travel time of acoustic signals and the phase differences between the array elements, the system determines both the range and angle to the transponder, allowing for computing the 3D position of the underwater device relative to the vessel in real time. In the context of benthic habitat mapping, the integration of the USBL system with ground truth data collection can enable more precise alignment between physical samples, video frames, and MBES measurements, thereby enhancing the reliability and interpretation of acoustic mapping

results.

Selection of ground truth sampling locations

In principle, sufficient ground truth samples are pursued to represent all seabed types encountered during a survey. However, practical constraints, such as survey time and budget, often limit the total number of samples, making it essential to optimize their spatial distribution. A simple random sampling scheme can serve as an initial baseline.

When prior knowledge of seabed topography or composition is available, sampling density should be tailored to the mapping objectives: denser in the heterogeneous or ecologically important zones and sparser in uniform areas. In the absence of this prior information, conducting a preliminary survey, which delivers bathymetry and a simple backscatter mosaic, can help to identify distinct habitat zones. Afterwards, sampling locations can be adjusted accordingly to ensure adequate coverage of potentially different seabed types within the study area.

8

Conclusion and outlook

Seen through the lens of change, the world cannot linger even for a moment. Seen through the lens of the eternal, all things, myself included, are boundless.

Su Shi (1037—1101)

To meet the growing need for efficient, non-destructive characterization of benthic habitats, this thesis has focused on data-driven seabed mapping techniques that exploit multi-frequency acoustic measurements from the multibeam echosounder (MBES). Although the acoustic data hold high potential for inferring the bio-physical properties of benthic habitats, their practical use is hindered by limitations in seabed ground truth collection, the lack of absolute calibration, and the complexities of multibeam backscatter processing. To address these obstacles, four research objectives were formulated in Section 1.3.2. Section 8.1 of this chapter revisits each objective, summarizes the main findings and limitations regarding case studies described in the previous chapters, and discusses the implications for future work. Section 8.2 then outlines broader research directions for benthic habitat mapping beyond the scope of this thesis.

8.1. Conclusion regarding research objectives

i) Mapping the occurrence of "acoustically hard" marine benthos, including epibenthos that modify the seabed surface characteristics and endobenthos that affect the sediment volume properties, and their associated sediment types with multi-spectral MBES measurements.

This thesis has mainly investigated the habitats of three "acoustically hard" marine benthos and shown the effectiveness of using MBES measurements for their characterization. Assemblages of epibenthos such as cultivated blue mussels (*Mytilus edulis*) can modify the sediment surface roughness and present distinct geomorphological patterns (Chapter 5). We employed an MBES system Kongsberg EM2040c at the operating frequency of 300 kHz and conducted a successful binary classification for mussel existence by combining bathymetric and backscatter features. The limitation of this study lies in the small range of sediment types within the cultivation area (from muddy sand to sand). It is of great significance to investigate the difference between mussel assemblages and other coarse sediments in the future, especially with multi-frequency acoustic measurements.

Chapter 3 and 6 put the focus on endobenthos. With a multi-spectral system R2Sonic 2026, we explored acoustic mapping of these habitats with a frequency range of 90–450 kHz. Chapter 3 demonstrated a strong link between the occurrence of sand mason worms (*Lanice conchilega*) and an increase in the backscatter strength (BS) compared to the sandy sediment. Through constructing tubes from sand and shell fragments, sand mason worms can alter the seabed surface roughness and increase the heterogeneity of the shallow subsurface. Although little difference was revealed among 90, 200, 300, and 450 kHz via classification per-frequency, we found that mean BS acquired from three distinct angular ranges can facilitate the discrimination between sand mason worm habitats and coarser sediments composed of higher gravel and dead shell content.

In contrast to the analysis and comparison among single frequencies in Chapter 3, Chapter 6 showed the effectiveness of a combined use of multi-spectral measurements for characterizing benthic habitats of the cut trough shells (*Spisula subtruncata*) and distinguishing them from sandy sediments and coarser seabed materials. By integrating MBES measurements from two seasons, we revealed high BS values at 90, 300, and 450 kHz in summer, caused by shell materials of abundant living shellfish on the seabed

surface. In winter, these frequencies produced different acoustic responses regarding an extended empty shell layer beneath the sediment surface. Cut trough shells tend to inhabit sediments composed of fine sand. When comparing the empty shell bed covered by a thin layer of fine sediments to dead shell materials accumulated in the deeper seabed, different contrasts among 90, 300, and 450 kHz were shown in their angular response curve (ARC) shape, demonstrating the richer seabed information brought by multi-frequency *BS*. Combined with traditional bottom sampling, multi-frequency MBES data make it possible to more accurately define spatial boundaries of the shell bed, which can facilitate ecological studies of these habitats and planning of destructive activities such as sand extraction.

Constrained by the available sonar system, the lowest employed acoustic frequency in this thesis is 90 kHz. Similar *BS* spatial patterns were observed across the frequency range 90–450 kHz for endobenthos such as sand mason worms (Chapter 3), since their dense tube aggregations can protrude out of the sediment surface, and penetration depths among these frequencies might be comparable. Combining acoustic systems that provide lower frequencies may improve sediment penetration and better reveal the subsurface habitats.

ii) Developing data-driven acoustic seabed mapping methods with optimal use of limited ground truth data.

When developing supervised classification models, limited ground truth not only results in a small training dataset but also might lead to overlooking some seabed types within the study area. In this case, this thesis has demonstrated the value of employing unsupervised and semi-supervised classification in acoustic seabed mapping, which allows for exploring the acoustic feature space more fully and combining the ground truth information when necessary.

Three unsupervised classification methods have been investigated, which do not require ground truth data during model development. In Chapter 3, we showed the effectiveness of classifying single-frequency *BS* data with the Bayesian classification and hierarchical clustering combined with principal component analysis (PCA). Bayesian classification is built on the statistical characteristics of *BS* at individual incident angles, provides an efficient way to inspect seabed variations, and can be extended to multi-frequency backscatter data [106]. Although hierarchical clustering was applied to ARC features with a degraded spatial resolution in this thesis, it is a flexible approach to combine various acoustic features in classification. As another commonly used classification method in machine learning (ML), K-means clustering was successfully applied for seabed classification in Chapter 6, given feature reduction that can extract critical information from the high-dimensional multi-frequency data. Compared to the linear feature reduction method PCA, our neural network solution proved more effective, which combines a spectral clustering training objective with *BS* source information on incident angle and frequency. It better captures variations in acoustic features and the non-linear relationships between data points. This deep learning (DL) based feature reduction subsequently improved the K-means clustering performance. While the choice of classification or clustering algorithms is significant, effectively constructing the feature space and integrating domain knowledge are also crucial elements in developing

data-driven techniques for MBES seabed mapping.

Based on the straightforward binary seabed classification of mussels versus bare sediment in Chapter 5, which was indicated by the ground truth samples and expert knowledge, we directly labeled the MBES bathymetry and backscatter features. With only nine seabed samples, the classification was performed in a semi-supervised way. We conducted pseudo-labeling via the label spreading algorithm, which exploited the relationship between the acoustic features of labeled and unlabeled data, largely expanded the training dataset, and enabled robust training of supervised ML classification algorithms.

Apart from classifying MBES data into different seabed types, change detection using *BS* differences (Chapter 4) also benefited from two unsupervised techniques: K-means and Gaussian Mixture Model (GMM) clustering. In addition, we utilized semi-supervised label propagation to generate full-coverage, spatially dense *BS* data at a single incident angle (Chapter 6).We constructed a graph, in which each node represented an individual sounding and was characterized by its geographic coordinates and bathymetric features. This method allowed for propagating single-angle *BS* values from soundings with known measurements to those without, effectively filling spatial gaps across the surveyed area.

Although various techniques help to address the issue of limited ground truth data, model validation can still be challenging in benthic habitat classification or change detection. To address this issue, improving ground truth sampling strategies is essential. For example, sampling density may be modified based on spatial patterns suggested by bathymetry and backscatter data, and consistent sampling locations for long-term seabed monitoring may be selected. Building open datasets for seabed measurements is also an important direction.

iii) Improving the MBES data processing workflow to effectively combine various multi-frequency features, such as bathymetry, seabed backscatter, and *BS* angular responses, from multiple surveys in seabed mapping.

The absence of absolute backscatter calibration brings difficulties to comparing MBES measurements across multiple surveys or combining them in acoustic seabed mapping. Chapter 4 addressed this issue via a relative calibration step before conducting change detection on two MBES datasets. The relative calibration eliminates systematic changes of *BS* measurements due to variations in the seawater environment (e.g., temperature), which can be attributed to drift in transducer sensitivity and other hardware characteristics, by leveraging an unchanged reference area between different survey times. Previous research has also recommended this approach for ensuring the repeatability of an MBES system and making measurements across surveys comparable [70, 71, 94]. Our downstream results for change detection aligned well with the existing ground truth information, demonstrating the necessity and effectiveness of relative backscatter calibration. This relative calibration also allowed for combining data across multiple surveys (see Chapter 6). The limitation of our approach lies in the lack of a reference area that has been monitored for a long time due to practical survey constraints (such as time and budget), which should be taken into account in future research.

To address the challenge of preserving a high spatial resolution when integrating *BS* and ARC features, Chapter 6 proposed to add a backscatter completion step before

acoustic classification. As mentioned in objective ii, this was accomplished by semi-supervised label propagation and resulted in spatially dense *BS* data for individual incident angles, which avoided sacrificing spatial resolution when accounting for ARC characteristics and enriched details in the final classification map.

Although additional steps have been added to the MBES data processing workflow, these strategies guarantee the long-term usability of MBES backscatter datasets and improve the integration of different features, e.g., beam-averaged *BS* and ARC characteristics. On the other hand, this thesis has not taken additional consideration for designing multi-frequency bathymetric features. For the sandy seabeds in our study areas, differences in penetration depth among 90, 300, and 450 kHz are not prominent. Nevertheless, when using a larger frequency range or measuring softer sediments, bathymetric differences among frequencies may be significant [67, 107] and should be included in the data processing workflow.

iv) Providing survey guidelines for future mapping and monitoring of marine benthos, with a focus on shellfish beds in the Dutch North Sea.

Based on the insights from seabed measurements in the Dutch North Sea, we summarized recommendations for future surveys in Chapter 7 from three aspects: MBES system, relative calibration, and ground truth data collection.

As a dynamic and potentially fragile habitat, management and conservation of the shellfish bed require non-destructive monitoring in the long term, making it essential to employ consistent sonar configurations. A multi-spectral system is highly recommended, considering the benefits of using multi-frequency *BS* measurements for shell bed mapping, as demonstrated in previous chapters. Taking the system R2Sonic 2026 as an example, we provided MBES settings that proved effective in this thesis. In addition, due to the strict requirements for equipment and environment regarding absolute backscatter calibration, relative calibration can be a practical alternative for aligning multiple MBES datasets. Near the Dutch North Sea, existing reference areas along the Belgian coast can be considered for relative calibration of MBES systems on an annual basis. Lastly, constraints for extractive activities, survey time, and budget require improved ground truth sampling strategies. Apart from the careful selection of sampling locations, non-extractive techniques such as underwater cameras can provide valuable information, especially when the images or videos are geo-referenced.

8.2. Outlook

This thesis has presented case studies of using multi-frequency MBES data for mapping different benthic habitats and highlighted the advantage of data-driven methods, including statistical modeling, ML, and DL approaches. Despite this, we have mainly focused on the near-shore regions of the Dutch North Sea and three types of marine benthos habitats. Although recent research has investigated diverse benthic habitats worldwide with acoustic mapping, these studies are usually associated with specific regions, distinct MBES systems, and various ways of backscatter data processing. To build a catalogue for acoustic responses of different benthic habitats, it is essential to develop standardized frameworks for backscatter data collection and (relative) calibra-

tion. This is also a key step for constructing benchmark datasets, which can further help to develop seabed mapping algorithms with better generalization abilities.

Although data-driven techniques offer practical solutions for mapping benthic habitats, research on the physical modeling of acoustic reflection and scattering characteristics of marine benthos remains highly significant. First, this provides potential for model inversion of biological habitat properties, which may reduce the need for extractive sampling on the seabed. In addition, physical models involving marine benthos are beneficial for quantifying the uncertainty sources of seabed backscatter measurements. This is valuable for establishing common standards regarding backscatter data acquisition that are comparable to the quality control of bathymetric measurements mandated by the International Hydrographic Organization (IHO).

Beyond the beam-averaged BS used in this thesis, future studies can also exploit the full backscatter time series in acoustic mapping. Analyzing backscattered signals from the water column would allow for resolving gas plumes, aquatic vegetation, and benthos aggregations above the seabed and in the subsurface, extending seabed mapping products from two-dimensional maps to the three-dimensional reconstruction of benthic habitats.

Besides traditional bottom samples, high-resolution and geo-referenced underwater images or videos can offer a broader inspection of the seabed and help to link acoustic classification results to visible seabed structures. Moreover, combining these *in-situ* observations with multi-spectral MBES data and even remote sensing data from other platforms, such as autonomous underwater vehicles (AUVs) and satellites, provides opportunities for multi-scale analyses of marine habitat conditions. Advances in multi-source data fusion techniques will extend marine ecosystem monitoring from local case studies to broader spatial scales, strengthening our ability to quantify the impact of anthropogenic and climate pressures on ocean environments.

Bibliography

[1] C. McLaverty, E. D. Beukhof, K. Bromhall, G. E. Dinesen, A. C. Erichsen and O. R. Eigaard. 'The relative effects of bottom trawling, organic enrichment, and natural environmental factors on coastal seabed communities'. In: *Marine Pollution Bulletin* 209 (2024), p. 117169.

[2] G. A. Díaz-Mendoza, K. Krämer, G. A. von Rönn, C. Heinrich, K. Schwarzer, H.-C. Reimers and C. Winter. 'Hotspots of human impact on the seafloor in the Southwestern Baltic Sea'. In: *Continental Shelf Research* 285 (2025), p. 105362.

[3] S. Garthe, H. Schwemmer, M. Mercker, V. Dierschke, N. Markones, V. Peschko, P. Schwemmer and J. C. Krause. 'Wind farms in proximity to marine protected areas put conservation targets at risk'. In: *Journal for Nature Conservation* 84 (2025), p. 126805.

[4] S. Ajmi, M. Boutet, A.-C. Bennis and J.-C. Dauvin. 'Impacts of offshore wind farms in the English Channel on larval dispersal and connectivity of benthic species using numerical modeling'. In: *Ocean Dynamics* 75.3 (2025), p. 22.

[5] J. B. Teixeira, A. S. Martins, H. T. Pinheiro, N. A. Secchin, R. L. de Moura and A. C. Bastos. 'Traditional ecological knowledge and the mapping of benthic marine habitats'. In: *Journal of Environmental Management* 115 (2013), pp. 241–250.

[6] M. C. De Esteban, P. Feldens, R. Haroun, F. Tuya, A. Gil and F. O. Ferrer. 'Habitat mapping of the Vila Franca do Campo marine reserve (Azores) and recommendations for its improvement'. In: *Estuarine, Coastal and Shelf Science* 303 (2024), p. 108809.

[7] Ministerie van Verkeer en Waterstaat. 'Beleidsregels ontgrondingen in rijkswateren'. In: *Staatscourant 14987* (2010).

[8] R. McGeady, R. M. Runya, J. S. Dooley, J. A. Howe, C. J. Fox, A. J. Wheeler, G. Summers, A. Callaway, S. Beck, L. S. Brown, G. Dooly and C. McGonigle. 'A review of new and existing non-extractive techniques for monitoring marine protected areas'. In: *Frontiers in Marine Science* 10 (2023), p. 1126301.

[9] K. Herkül, A. Peterson and S. Paekivi. 'Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows'. In: *Estuarine, Coastal and Shelf Science* 192 (2017), pp. 57–71.

[10] E. Salameh, F. Frappart, R. Almar, P. Baptista, G. Heygster, B. Lubac, D. Raucoules, L. P. Almeida, E. W. Bergsma, S. Capo, M. D. Michele, D. Idier, Z. Li, V. Marieu, A. Poupardin, P. A. Silva, I. Turki and B. Laignel. 'Monitoring beach topography and nearshore bathymetry using spaceborne remote sensing: A review'. In: *Remote Sensing* 11.19 (2019), p. 2212.

[11] D. Su, H. Gao, A. Yang, J. Wang, X. Mai, X. Liu, F. Yang and Z. Wu. 'Classification of Seabed Sediment by Combining Airborne LiDAR Bathymetry and Multispectral Remote Sensing Images'. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2025).

[12] D. E. C. Rivera, F. F. Diederiks, N. M. Hammerman, T. Staples, E. Kovacs, K. Markey and C. M. Roelfsema. 'Remote Sensing Reveals Multidecadal Trends in Coral Cover at Heron Reef, Australia'. In: *Remote Sensing* 17.7 (2025), p. 1286.

[13] X. Ji, B. Yang, Z. Wei, M. Wang, Q. Tang, W. Xu, Y. Wang, J. Zhang and L. Zhang. 'Benthic habitat sediments mapping in coral reef area using amalgamation of multi-source and multi-modal remote sensing data'. In: *Remote Sensing of Environment* 304 (2024), p. 114032.

[14] B. Misiuk and C. J. Brown. 'Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor'. In: *Estuarine, Coastal and Shelf Science* 296 (2024), p. 108599.

[15] D. G. Tucker and B. K. Gazey. *Applied underwater acoustics*. Pergamon Press, 1966.

[16] X. Lurton. *An introduction to underwater acoustics: principles and applications*. Springer–Verlag Berlin Heidelberg, 2010.

[17] D. R. Jackson and M. D. Richardson. *High-frequency seafloor acoustics*. Springer–Verlag New York, 2007.

[18] Techinical Report APL-UW. *APL-UW High-Frequency Ocean Environmental Acoustic Models Handbook*. Tech. rep. Applied Physics Laboratory, University of Washington, 1994.

[19] R. J. Urick. *Principles of underwater sound*. McGraw-Hill, New York, 1983.

[20] D. G. Simons, M. Snellen and M. A. Ainslie. 'A multivariate correlation analysis of high-frequency bottom backscattering strength measurements with geotechnical parameters'. In: *IEEE Journal of Oceanic Engineering* 32.3 (2008), pp. 640–650.

[21] E. Hutin, Y. Simard and P. Archambault. 'Acoustic detection of a scallop bed from a single-beam echosounder in the St. Lawrence'. In: *ICES Journal of Marine Science* 62.5 (2005), pp. 966–983.

[22] M. Snellen, D. G. Simons and R. Riethmueller. 'High frequency scattering measurements for mussel bed characterisation'. In: *Journal of the Acoustical Society of America* 123.5 (2008), p. 3627.

[23] M. S. Ballard and K. M. Lee. 'The acoustics of marine sediments'. In: *Acoustics Today* 13.3 (2017), pp. 11–18.

[24] K. M. Lee, G. R. Venegas, M. S. Ballard, K. M. Dorgan, E. Kiskaddon, A. R. McNeese and P. S. Wilson. 'Impacts of infauna, worm tubes, and shell hash on sediment acoustic variability and deviation from the viscous grain shearing model'. In: *The Journal of the Acoustical Society of America* 152.4 (2022), pp. 2456–2474.

[25] P. Wille. *Sound images of the ocean: in research and monitoring*. Springer Science & Business Media, 2005.

[26] L. Hamilton, P. Mulhearn and R. Poeckert. 'Comparison of RoxAnn and QTC-View acoustic bottom classification system performance for the Cairns area, Great Barrier Reef, Australia'. In: *Continental Shelf Research* 19.12 (1999), pp. 1577–1597.

[27] M. M. Landero Figueroa, M. J. Parsons, B. J. Saunders, B. Radford, C. Salgado-Kent and I. M. Parnum. 'The use of singlebeam echo-sounder depth data to produce demersal fish distribution models that are comparable to models produced using multibeam echo-sounder depth'. In: *Ecology and Evolution* 11.24 (2021), pp. 17873–17884.

[28] X. Shang, J. Zhao and H. Zhang. 'Obtaining high-resolution seabed topography and surface details by co-registration of side-scan sonar and multibeam echo sounder images'. In: *Remote sensing* 11.12 (2019), p. 1496.

[29] B. D. Van Veen and K. M. Buckley. 'Beamforming: A versatile approach to spatial filtering'. In: *IEEE ASSP Magazine* 5.2 (1988), pp. 4–24.

[30] L. Fonseca and L. Mayer. 'Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data'. In: *Marine Geophysical Research* 28 (2007), pp. 119–126.

[31] D. Marx, M. Nelson, E. Chang, W. Gillespie, A. Putney and K. Warman. 'An introduction to synthetic aperture sonar'. In: *Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing*. IEEE. 2000, pp. 717–721.

[32] T. Mulder, H. Hüneke and A. Van Loon. 'Chapter 1–Progress in deep-sea sedimentology'. In: *Deep-Sea Sediments*. Vol. 63. Elsevier, 2011, pp. 1–24.

[33] H. Álvarez, S. García, A. Perry, J. Blanco, D. Maaholm and R. Aguilar. *Protecting the North Sea: Borkum Stones*. Tech. rep. Oceana, Madrid, 2019.

[34] S. Nieuwhof. 'The use of remote sensing to reveal landscape-scale ecosystem engineering by shellfish reefs'. PhD thesis. University of Twente, 2018.

[35] D. Eleftherakis, A. Amiri-Simkooei, M. Snellen and D. G. Simons. 'Improving riverbed sediment classification using backscatter and depth residual features of multi-beam echo-sounder systems'. In: *The Journal of the Acoustical Society of America* 131.5 (2012), pp. 3710–3725.

[36] P. J. W. Siwabessy, M. Tran, K. Picard, B. P. Brooke, Z. Huang, N. Smit, D. K. Williams, W. A. Nicholas, S. L. Nichol and I. Atkinson. 'Modelling the distribution of hard seabed using calibrated multibeam acoustic backscatter data in a tropical, macrotidal embayment: Darwin Harbour, Australia'. In: *Marine Geophysical Research* 39 (2018), pp. 249–269.

[37] L. Fonseca, C. Brown, B. Calder, L. Mayer and Y. Rzhanov. 'Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures'. In: *Applied Acoustics* 70.10 (2009), pp. 1298–1304.

[38] C. McGonigle and J. S. Collier. 'Interlinking backscatter, grain size and benthic community structure'. In: *Estuarine, Coastal and Shelf Science* 147 (2014), pp. 123–136.

[39] Z. Huang, S. L. Nichol, J. P. Siwabessy, J. Daniell and B. P. Brooke. 'Predictive modelling of seabed sediment parameters using multibeam acoustic data: a case study on the Carnarvon Shelf, Western Australia'. In: *International Journal of Geographical Information Science* 26.2 (2012), pp. 283–307.

[40] R. L. Folk. 'The distinction between grain size and mineral composition in sedimentary-rock nomenclature'. In: *The Journal of Geology* 62.4 (1954), pp. 344–359.

[41] M. Snellen, T. C. Gaida, L. Koop, E. Alevizos and D. G. Simons. 'Performance of multibeam echosounder backscatter-based classification for monitoring sediment distributions using multitemporal large-scale ocean data sets'. In: *IEEE Journal of Oceanic Engineering* 44.1 (2018), pp. 142–155.

[42] B. Misiuk, M. Diesing, A. Aitken, C. J. Brown, E. N. Edinger and T. Bell. 'A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using random forest'. In: *Geosciences* 9.6 (2019), p. 254.

[43] M. Snellen, K. Siemes and D. G. Simons. 'Model-based sediment classification using single-beam echosounder signals'. In: *The Journal of the Acoustical Society of America* 129.5 (2011), pp. 2878–2888.

[44] G. Xu, B. T. Hefner, D. R. Jackson, A. N. Ivakin and G. Wendelboe. 'A Physics-Based Inversion of Multibeam Sonar Data for Seafloor Characterization'. In: *IEEE Journal of Oceanic Engineering* (2024).

[45] Q. Tang, J. Li, D. Ding, X. Ji, N. Li, L. Yang and W. Sun. 'Deep-sea seabed sediment classification using finely processed multibeam backscatter intensity data in the southwest Indian ridge'. In: *Remote Sensing* 14.11 (2022), p. 2675.

[46] C. Brenan, M. Kienast, V. Maselli, C. K. Algar, B. Misiuk and C. J. Brown. 'Seafloor sediment characterization improves estimates of organic carbon standing stocks: an example from the Eastern Shore Islands, Nova Scotia, Canada'. In: *Biogeosciences* 21.20 (2024), pp. 4569–4586.

[47] E. Sklar, E. Bushuev, B. Misiuk, G. Labbé-Morissette and C. J. Brown. 'Seafloor morphology and substrate mapping in the Gulf of St Lawrence, Canada, using machine learning approaches'. In: *Frontiers in Marine Science* 11 (2024), p. 1306396.

[48] Ł. Janowski. 'Advancing Seabed Bedform Mapping in the Kuźnica Deep: Leveraging Multibeam Echosounders and Machine Learning for Enhanced Underwater Landscape Analysis'. In: *Remote Sensing* 17.3 (2025), p. 373.

[49] S. Mestdagh, A. Amiri-Simkooei, K. J. van der Reijden, L. Koop, S. O'Flynn, M. Snellen, C. Van Sluis, L. L. Govers, D. G. Simons, P. M. Herman, H. Olff and T. Ysebaert. 'Linking the morphology and ecology of subtidal soft-bottom marine benthic habitats: A novel multiscale approach'. In: *Estuarine, Coastal and Shelf Science* 238 (2020), p. 106687.

[50] V. Lucieer, N. A. Hill, N. S. Barrett and S. Nichol. 'Do marine substrates 'look'and 'sound' the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle images'. In: *Estuarine, Coastal and Shelf Science* 117 (2013), pp. 94–106.

[51] B. Misiuk, T. Bell, A. Aitken, C. J. Brown and E. N. Edinger. 'Mapping Arctic clam abundance using multiple datasets, models, and a spatially explicit accuracy assessment'. In: *ICES Journal of Marine Science* 76.7 (2019), pp. 2349–2361.

[52] D. Ierodiaconou, J. Monk, A. Rattray, L. Laurenson and V. Versace. 'Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations'. In: *Continental Shelf Research* 31.2 (2011), S28–S38.

[53] R. Che Hasan, D. Ierodiaconou, L. Laurenson and A. Schimel. 'Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping'. In: *PLoS ONE* 9.5 (2014), e97339.

[54] S. Nemani, B. Misiuk, D. Cote, E. Edinger, J. Mackin-McLaughlin, A. Templeton and K. Robert. 'Incorporating functional traits with habitat maps: patterns of diversity in coastal benthic assemblages'. In: *Frontiers in Marine Science* 11 (2024), p. 1141737.

[55] B. R. Wilson, C. J. Brown, J. A. Sameoto, M. Lacharité, A. M. Redden and V. Gazzola. 'Mapping seafloor habitats in the Bay of Fundy to assess megafaunal assemblages associated with Modiolus modiolus beds'. In: *Estuarine, Coastal and Shelf Science* 252 (2021), p. 107294.

[56] M. Lacharité, C. J. Brown and V. Gazzola. 'Multisource multibeam backscatter data: Developing a strategy for the production of benthic habitat maps using semi-automated seafloor classification methods'. In: *Marine Geophysical Research* 39 (2018), pp. 307–322.

[57] P. Porskamp, A. C. Schimel, M. Young, A. Rattray, Y. Ladroit and D. Ierodiaconou. 'Integrating multibeam echosounder water-column data into benthic habitat mapping'. In: *Limnology and Oceanography* 67.8 (2022), pp. 1701–1713.

[58] T. C. Gaida, T. A. van Dijk, M. Snellen, T. Vermaas, C. Mesdag and D. G. Simons. 'Monitoring underwater nourishments using multibeam bathymetric and backscatter time series'. In: *Coastal Engineering* 158 (2020), p. 103666.

[59] J. Berry and C. Nanlal. 'Assessment of the application of each multibeam echosounder data product for monitoring of Laminaria digitata in the UK'. In: *Frontiers in Remote Sensing* 6 (2025), p. 1521958.

[60] D. Eleftherakis, M. Snellen, A. Amiri-Simkooei, D. G. Simons and K. Siemes. 'Observations regarding coarse sediment classification based on multi-beam echo-sounder's backscatter strength and depth residuals in Dutch rivers'. In: *The Journal of the Acoustical Society of America* 135.6 (2014), pp. 3305–3315.

[61] M. Diesing, P. J. Mitchell, E. O'Keeffe, G. O. M. Gavazzi and T. L. Bas. 'Limitations of predicting substrate classes on a sedimentary complex but morphologically simple seabed'. In: *Remote Sensing* 12.20 (2020), p. 3398.

[62] L. Janowski, K. Trzcinska, J. Tegowski, A. Kruss, M. Rucinska-Zjadacz and P. Pocwiardowski. 'Nearshore benthic habitat mapping based on multi-frequency, multibeam echosounder data using a combined object-based approach: A case study from the Rowy site in the southern Baltic Sea'. In: *Remote Sensing* 10.12 (2018), p. 1983.

[63] P. Feldens, I. Schulze, S. Papenmeier, M. Schönke and J. Schneider von Deimling. 'Improved interpretation of marine sedimentary environments using multi-frequency multibeam backscatter data'. In: *Geosciences* 8.6 (2018), p. 214.

[64] K. Haris, B. Chakraborty, B. Ingole, A. Menezes and R. Srivastava. 'Seabed habitat mapping employing single and multi-beam backscatter data: A case study from the western continental shelf of India'. In: *Continental Shelf Research* 48 (2012), pp. 40–49.

[65] W. B. Ryan and R. D. Flood. 'Side-looking sonar backscatter response at dual frequencies'. In: *Marine Geophysical Researches* 18 (1996), pp. 689–705.

[66] J. E. H. Clarke. 'Multispectral acoustic backscatter from multibeam, improved classification potential'. In: *Proceedings of the United States Hydrographic Conference, San Diego, CA, USA*. 2015, pp. 16–19.

[67] T. C. Gaida, T. H. Mohammadloo, M. Snellen and D. G. Simons. 'Mapping the seabed and shallow subsurface with multi-frequency multibeam echosounders'. In: *Remote Sensing* 12.1 (2019), p. 52.

[68] P. S. Menandro, B. Misiuk, J. Schneider von Deimling, A. C. Bastos and C. J. Brown. 'Multifrequency seafloor acoustic backscatter as a tool for improved biological and geological assessments–updating knowledge, prospects, and challenges'. In: *Frontiers in Remote Sensing* 6 (2025), p. 1546280.

[69] C. J. Brown, J. Beaudoin, M. Brissette and V. Gazzola. 'Multispectral multibeam echo sounder backscatter as a tool for improved seafloor characterization'. In: *Geosciences* 9.3 (2019), p. 126.

[70] M. Roche, K. Degrendele, C. Vrignaud, S. Loyer, T. Le Bas, J.-M. Augustin and X. Lurton. 'Control of the repeatability of high frequency multibeam echosounder backscatter by using natural reference areas'. In: *Marine Geophysical Research* 39 (2018), pp. 89–104.

[71] M. Roche, T. I. B. Lønmo, R. Fezzani, L. Berger, S. Deleu, H. Bisquay, A. Gaillot, K. Vanparys, J. Vercaemst, K. Degrendele, F. Barette, L. E. Fonseca, J. Verstraeten, K. E. Nilsen, G. Montereale-Gavazzi, X. Lurton and J.-M. Augustin. 'Instrumental temperature-dependence of backscatter measurements by a multibeam echosounder: findings and implications'. In: *Frontiers in Remote Sensing* 6 (2025), p. 1572545.

[72] A. R. Amiri-Simkooei, L. Koop, K. J. van der Reijden, M. Snellen and D. G. Simons. 'Seafloor characterization using multibeam echosounder backscatter data: methodology and results in the North Sea'. In: *Geosciences* 9.7 (2019), p. 292.

[73] M. Diesing, S. L. Green, D. Stephens, R. M. Lark, H. A. Stewart and D. Dove. 'Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches'. In: *Continental Shelf Research* 84 (2014), pp. 107–119.

[74] P. Porskamp, A. Rattray, M. Young and D. Ierodiaconou. 'Multiscale and hierarchical classification for benthic habitat mapping'. In: *Geosciences* 8.4 (2018), p. 119.

[75] X. Shang, K. Robert, B. Misiuk, J. Mackin-McLaughlin and J. Zhao. 'Self-adaptive analysis scale determination for terrain features in seafloor substrate classification'. In: *Estuarine, Coastal and Shelf Science* 254 (2021), p. 107359.

[76] X. Shang, L. Dong and J. Zhao. 'Optimal scale determination for object-based backscatter image analysis in seafloor substrate classification based on classification uncertainty'. In: *IEEE Geoscience and Remote Sensing Letters* 21 (2024), pp. 1–5.

[77] D. Stephens and M. Diesing. 'A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data'. In: *PLoS ONE* 9.4 (2014), e93950.

[78] B. Misiuk, Y. L. Tan, M. Z. Li, T. Trappenberg, A. Alleosfour, I. W. Church, V. Ferrini and C. J. Brown. 'Multivariate mapping of seabed grain size parameters in the Bay of Fundy using convolutional neural networks'. In: *Marine Geology* 472 (2024), p. 107299.

[79] H. Wang, W. Chen, X. Li, Q. Liang, X. Qin and J. Li. 'CUG-STCN: A seabed topography classification framework based on knowledge graph-guided vision mamba network'. In: *International Journal of Applied Earth Observation and Geoinformation* 136 (2025), p. 104383.

[80] D. G. Simons and M. Snellen. 'A Bayesian approach to seafloor classification using multi-beam echo-sounder backscatter data'. In: *Applied Acoustics* 70.10 (2009), pp. 1258–1268.

[81] P. Zhou, G. Chen, M. Wang, J. Chen and Y. Li. 'Sediment classification of acoustic backscatter image based on stacked denoising autoencoder and modified extreme learning machine'. In: *Remote Sensing* 12.22 (2020), p. 3762.

[82] A. Minelli, A. N. Tassetti, B. Hutton, G. N. Pezzuti Cozzolino, T. Jarvis and G. Fabi. 'Semi-automated data processing and semi-supervised machine learning for the detection and classification of water-column fish schools and gas seeps with a multibeam echosounder'. In: *Sensors* 21.9 (2021), p. 2999.

[83] Y. Wang, Y. Sun, X. Cui, T. Yun, X. Bu and F. Yang. 'A Semi-Supervised Classification Method for Riverbed Benthic Sediments Using Integrated Superpixel Segmentation and Confident Learning Sample Enhancement'. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2025).

[84] C. Lee, S.-J. Park, D. Yoon, B. Y. Yi and M. Lim. 'Enhancing Seabed Sediment Classification with Multibeam Echo-Sounding and Self-training: A Case Study from the East Sea of South Korea'. In: *Frontiers in Earth Science* 13 (2025), p. 1550244.

[85] P. Urban, M. E. Veloso-Alarcón and J. Greinert. 'Echo grid integration: A novel method for preprocessing multibeam water column data to quantify underwater gas bubble emissions'. In: *Limnology and Oceanography: Methods* 21.7 (2023), pp. 377–400.

[86] A. Amiri-Simkooei, M. Snellen and D. G. Simons. 'Riverbed sediment classification using multi-beam echo-sounder backscatter data'. In: *The Journal of the Acoustical Society of America* 126.4 (2009), pp. 1724–1738.

[87] T. Gaida. 'Acoustic mapping and monitoring of the seabed: From single-frequency to multispectral multibeam backscatter'. PhD thesis. Delft University of Technology, 2020.

[88] T. H. Mohammadloo. 'Optimizing hydrographic operations for bathymetric measurements using multibeam echosounders'. PhD thesis. Delft University of Technology, 2020.

[89] T. H. Mohammadloo, M. Snellen, W. Renoud, J. Beaudoin and D. G. Simons. 'Correcting multibeam echosounder bathymetric measurements for errors induced by inaccurate water column sound speeds'. In: *IEEE Access* 7 (2019), pp. 122052–122068.

[90] T. H. Mohammadloo, M. Snellen and D. G. Simons. 'Assessing the performance of the multi-beam echo-sounder bathymetric uncertainty prediction model'. In: *Applied Sciences* 10.13 (2020), p. 4671.

[91] J. Le Deunf, N. Debese, T. Schmitt and R. Billot. 'A review of data cleaning approaches in a hydrographic framework with a focus on bathymetric multibeam echosounder datasets'. In: *Geosciences* 10.7 (2020), p. 254.

[92] G. Lamarche, X. Lurton, A.-L. Verdier and J.-M. Augustin. 'Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—Application to Cook Strait, New Zealand'. In: *Continental Shelf Research* 31.2 (2011), S93–S109.

[93] I. M. Parnum and A. N. Gavrilov. 'High-frequency multibeam echo-sounder measurements of seafloor backscatter in shallow water: Part 2–Mosaic production, analysis and classification'. In: *Underwater Technology* 30.1 (2011), pp. 13–26.

[94] 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'. In: *Marine Geophysical Research* 39 (2018), pp. 55–73.

[95] R. M. Haralick, K. Shanmugam and I. H. Dinstein. 'Textural features for image classification'. In: *IEEE Transactions on Systems, Man, and Cybernetics* 6 (1973), pp. 610–621.

[96] Q. Bai, S. Mestdagh, M. Snellen and D. G. Simons. 'Indications of marine benthos occurrence from multi-spectral multi-beam backscatter data: a case study in the North Sea'. In: *Frontiers in Earth Science* 11 (2023), p. 1140649.

[97] C. J. Brown, S. J. Smith, P. Lawton and J. T. Anderson. 'Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques'. In: *Estuarine, Coastal and Shelf Science* 92.3 (2011), pp. 502–520.

[98] G. Lamarche and X. Lurton. 'Recommendations for improved and coherent acquisition and processing of backscatter data from seafloor-mapping sonars'. In: *Marine Geophysical Research* 39.1-2 (2018), pp. 5–22.

[99] K. Trzcinska, J. Tegowski, P. Pocwiardowski, L. Janowski, J. Zdroik, A. Kruss, M. Rucinska, Z. Lubniewski and J. Schneider von Deimling. 'Measurement of seafloor acoustic backscatter angular dependence at 150 kHz using a multibeam echosounder'. In: *Remote Sensing* 13.23 (2021), p. 4771.

[100] C. J. Brown and J. S. Collier. 'Mapping benthic habitat in regions of gradational substrata: an automated approach utilising geophysical, geological, and biological relationships'. In: *Estuarine, Coastal and Shelf Science* 78.1 (2008), pp. 203–214.

[101] V. E. Kostylev, R. C. Courtney, G. Robert and B. J. Todd. 'Stock evaluation of giant scallop (Placopecten magellanicus) using high-resolution acoustics for seabed mapping'. In: *Fisheries Research* 60.2-3 (2003), pp. 479–492.

[102] Z. Huang, J. Siwabessy, H. Cheng and S. Nichol. 'Using multibeam backscatter data to investigate sediment-acoustic relationships'. In: *Journal of Geophysical Research: Oceans* 123.7 (2018), pp. 4649–4665.

[103] D. G. Simons and M. Snellen. 'A comparison between modeled and measured high frequency bottom backscattering'. In: *Proceedings of the 9th European Conference on Underwater Acoustics (Paris, France)*. 2008, pp. 639–644.

[104] X. Cui, F. Yang, X. Wang, B. Ai, Y. Luo and D. Ma. 'Deep learning model for seabed sediment classification based on fuzzy ranking feature optimization'. In: *Marine Geology* 432 (2021), p. 106390.

[105] E. Alevizos and J. Greinert. 'The hyper-angular cube concept for improving the spatial and acoustic resolution of MBES backscatter angular response analysis'. In: *Geosciences* 8.12 (2018), p. 446.

[106] T. C. Gaida, T. A. Tengku Ali, M. Snellen, A. Amiri-Simkooei, T. A. Van Dijk and D. G. Simons. 'A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data'. In: *Geosciences* 8.12 (2018), p. 455.

[107] B. Misiuk and C. J. Brown. 'Multiple imputation of multibeam angular response data for high resolution full coverage seabed mapping'. In: *Marine Geophysical Research* 43.1 (2022), p. 7.

[108] K. Siemes. 'Establishing a sea bottom model by applying a multi-sensor acoustic remote sensing approach'. PhD thesis. Delft University of Technology, 2013.

[109] W. Revelle. 'Hierarchical cluster analysis and the internal structure of tests'. In: *Multivariate Behavioral Research* 14.1 (1979), pp. 57–74.

[110] S. Menard. *An introduction to logistic regression diagnostics*. California: Sage, 1995, p. 66.

[111] L. Koop, A. Amiri-Simkooei, K. J. van der Reijden, S. O'Flynn, M. Snellen and D. G. Simons. 'Seafloor classification in a sand wave environment on the Dutch continental shelf using multibeam echosounder backscatter data'. In: *Geosciences* 9.3 (2019), p. 142.

[112] C. Heinrich, P. Feldens and K. Schwarzer. 'Highly dynamic biological seabed alterations revealed by side scan sonar tracking of Lanice conchilega beds offshore the island of Sylt (German Bight)'. In: *Geo-Marine Letters* 37.3 (2017), pp. 289–303.

[113] R. M. Runya, C. McGonigle, R. Quinn, J. Howe, J. Collier, C. Fox, J. Dooley, R. O'loughlin, J. Calvert, L. Scott, C. Abernethy and W. Evans. 'Examining the links between multi-frequency multibeam backscatter data and sediment grain size'. In: *Remote Sensing* 13.8 (2021), p. 1539.

[114] T. Schmitt, N. C. Mitchell and A. T. S. Ramsay. 'Characterizing uncertainties for quantifying bathymetry change between time-separated multibeam echosounder surveys'. In: *Continental Shelf Research* 28.9 (2008), pp. 1166–1176.

[115] J. Ellis, T. Maxwell, M. Schratzberger and S. Rogers. 'The benthos and fish of offshore sandbank habitats in the southern North Sea'. In: *Journal of the Marine Biological Association of the United Kingdom* 91.6 (2011), pp. 1319–1335.

[116] H. Holzhauer, B. Borsje, P. Herman, C. Schipper and K. Wijnberg. 'The geomorphology of an ebb-tidal-delta linked to benthic species distribution and functionality'. In: *Ocean & Coastal Management* 216 (2022), p. 105938.

[117] Q. Bai, A. Amiri-Simkooei, S. Mestdagh, D. G. Simons and M. Snellen. 'Unsupervised seabed habitat change detection with multibeam backscatter data using a constrained Gaussian mixture model'. In: *Applied Acoustics* 246 (2026), p. 111222.

[118] D. Stephens and M. Diesing. 'A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data'. In: *PLoS ONE* 9.4 (2014), e93950.

[119] X. Ji, B. Yang and Q. Tang. 'Seabed sediment classification using multibeam backscatter data based on the selecting optimal random forest model'. In: *Applied Acoustics* 167 (2020), p. 107387.

[120] X. Cui, H. Liu, M. Fan, B. Ai, D. Ma and F. Yang. 'Seafloor habitat mapping using multibeam bathymetric and backscatter intensity multi-features SVM classification framework'. In: *Applied Acoustics* 174 (2021), p. 107728.

[121] D. Ierodiaconou, J. Monk, A. Rattray, L. Laurenson and V. Versace. 'Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations'. In: *Continental Shelf Research* 31.2 (2011), S28–S38.

[122] M. Schönke, L. Wiesenberg, I. Schulze, D. Wilken, A. Darr, S. Papenmeier and P. Feldens. 'Impact of sparse benthic life on seafloor roughness and high-frequency acoustic scatter'. In: *Geosciences* 9.10 (2019), p. 454.

[123] L. Khelifi and M. Mignotte. 'Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis'. In: *IEEE Access* 8 (2020), pp. 126385–126400.

[124] M. Hussain, D. Chen, A. Cheng, H. Wei and D. Stanley. 'Change detection from remotely sensed images: From pixel-based to object-based approaches'. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 80 (2013), pp. 91–106.

[125] G. Montereale-Gavazzi, M. Roche, X. Lurton, K. Degrendele, N. Terseleer and V. Van Lancker. 'Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian part of the North Sea'. In: *Marine Geophysical Research* 39 (2018), pp. 229–247.

[126] K. Zhang, X. Lv, H. Chai and J. Yao. 'Unsupervised SAR image change detection for few changed area based on histogram fitting error minimization'. In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–19.

[127] A. C. Schimel, D. Ierodiaconou, L. Hulands and D. M. Kennedy. 'Accounting for uncertainty in volumes of seabed change measured with repeat multibeam sonar surveys'. In: *Continental Shelf Research* 111 (2015), pp. 52–68.

[128] M. Malik, X. Lurton and L. Mayer. 'A framework to quantify uncertainties of seafloor backscatter from swath mapping echosounders'. In: *Marine Geophysical Research* 39 (2018), pp. 151–168.

[129] Y. Bazi, L. Bruzzone and F. Melgani. 'Automatic identification of the number and values of decision thresholds in the log-ratio image for change detection in SAR images'. In: *IEEE Geoscience and Remote Sensing Letters* 3.3 (2006), pp. 349–353.

[130] G. A. Seber. *Multivariate observations*. John Wiley & Sons, 2009.

[131] T. K. Moon. 'The expectation-maximization algorithm'. In: *IEEE Signal Processing Magazine* 13.6 (1996), pp. 47–60.

[132] P. J. Rousseeuw. 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis'. In: *Journal of Computational and Applied Mathematics* 20 (1987), pp. 53–65.

[133] M. Wienhoven, J. Maes, T. van Doorn, H. Schutte and K. Salkovic. *Scenario study for the Dutch part of the North Sea*. Tech. rep. Ecorys, 2023.

[134] R. Mamede da Silva Alves. 'Spatial structure and temporal dynamics of an intertidal population of the marine ecosystem engineering worm Lanice conchilega (Pallas, 1766)'. PhD thesis. Ghent University, 2017.

[135] R. Callaway. 'Juveniles stick to adults: recruitment of the tube-dwelling polychaete Lanice conchilega (Pallas, 1766)'. In: *Hydrobiologia* 503.1 (2003), pp. 121–130.

[136] G. Van Hoey, M. Vincx and S. Degraer. 'Some recommendations for an accurate estimation of Lanice conchilega density based on tube counts'. In: *Helgoland Marine Research* 60.4 (2006), pp. 317–321.

[137] R. Hare. *Error budget analysis for US Naval Oceanographic Office (NAVOCEANO) hydrographic survey systems*. Tech. rep. University of Southern Mississippi, Hydrographic Science Research Center for the Naval Oceanographic Office, 2001.

[138] M. Schönke, P. Feldens, D. Wilken, S. Papenmeier, C. Heinrich, J. S. von Deimling, P. Held and S. Krastel. 'Impact of Lanice conchilega on seafloor microtopography off the island of Sylt (German Bight, SE North Sea)'. In: *Geo-Marine Letters* 37 (2017), pp. 305–318.

[139] L. Koop, K. J. van der Reijden, S. Mestdagh, T. Ysebaert, L. L. Govers, H. Olff, P. M. Herman, M. Snellen and D. G. Simons. 'Measuring centimeter-scale sand ripples using multibeam echosounder backscatter data from the Brown Bank area of the Dutch Continental Shelf'. In: *Geosciences* 10.12 (2020), p. 495.

[140] Q. Bai, A. Amiri-Simkooei, S. Mestdagh, D. G. Simons and M. Snellen. 'Mussel culture monitoring with semi-supervised machine learning on multibeam echosounder data using label spreading'. In: *Journal of Environmental Management* 369 (2024), p. 122250.

[141] S. Asghar, J. Choi, D. Yoon and J. Byun. 'Spatial pseudo-labeling for semi-supervised facies classification'. In: *Journal of Petroleum Science and Engineering* 195 (2020), p. 107834.

[142] R. Sheikhpour, M. A. Sarram, S. Gharaghani and M. A. Z. Chahooki. 'A survey on semi-supervised feature selection methods'. In: *Pattern Recognition* 64 (2017), pp. 141–158.

[143] D. Zhou, O. Bousquet, T. Lal, J. Weston and B. Schölkopf. 'Learning with Local and Global Consistency'. In: *Advances in Neural Information Processing Systems*. Vol. 16. MIT Press, 2003.

[144] A. Ligthart, C. Catal and B. Tekinerdogan. 'Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification'. In: *Applied Soft Computing* 101 (2021), p. 107023.

[145] J. J. Capelle. 'Production efficiency of mussel bottom culture'. PhD thesis. Wageningen University and Research, 2017.

[146] S. J. Blott and K. Pye. 'Particle size scales and classification of sediment types based on particle size distributions: Review and recommended procedures'. In: *Sedimentology* 59.7 (2012), pp. 2071–2096.

[147] M. Mokarram and D. Sathyamoorthy. 'A review of landform classification methods'. In: *Spatial Information Research* 26.6 (2018), pp. 647–660.

[148] R. A. MacMillan and P. Shary. 'Landforms and landform elements in geomorphometry'. In: *Developments in Soil Science* 33 (2009), pp. 227–254.

[149] D. Kowalczyk. 'Casorati curvatures'. In: *Bulletin of the Transilvania University of Brasov, Series III: Mathematics, Informatics, Physics* 1.50 (2008).

[150] B. Misiuk and C. J. Brown. 'Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor'. In: *Estuarine, Coastal and Shelf Science* (2023), p. 108599.

[151] G. Loureiro, A. Dias, J. Almeida, A. Martins, S. Hong and E. Silva. 'A Survey of Seafloor Characterization and Mapping Techniques'. In: *Remote Sensing* 16.7 (2024), p. 1163.

[152] T. Hastie, R. Tibshirani, J. H. Friedman and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer, 2009.

[153] B. H. Menze, B. M. Kelm, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich and F. A. Hamprecht. 'A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data'. In: *BMC Bioinformatics* 10 (2009), pp. 1–16.

[154] M. Greenacre, P. J. Groenen, T. Hastie, A. I. d'Enza, A. Markos and E. Tuzhilina. 'Principal component analysis'. In: *Nature Reviews Methods Primers* 2.1 (2022), p. 100.

[155] A. Safonova, G. Ghazaryan, S. Stiller, M. Main-Knorn, C. Nendel and M. Ryo. 'Ten deep learning techniques to address small data problems with remote sensing'. In: *International Journal of Applied Earth Observation and Geoinformation* 125 (2023), p. 103569.

[156] S. Mastrantonis, B. Radford, T. Langlois, C. Spencer, S. de Lestang and S. Hickey. 'A novel method for robust marine habitat mapping using a kernelised aquatic vegetation index'. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 209 (2024), pp. 472–480.

[157] J. de Fouw, P. W. van Horssen, J. Craeymeersch, M. F. Leopold, J. Perdon, K. Troost, I. Tulp, J. van Zwol and C. J. Philippart. 'Spatio-temporal analysis of potential factors explaining fluctuations in population size of *Spisula subtruncata* in the Dutch North Sea'. In: *Frontiers in Marine Science* 11 (2024), p. 1476223.

[158] J. L. Gutiérrez, C. G. Jones, D. L. Strayer and O. O. Iribarne. 'Mollusks as ecosystem engineers: the role of shell production in aquatic habitats'. In: *Oikos* 101.1 (2003), pp. 79–90.

[159] J. Craeymeersch and E. Velilla. *Shellfish aggregations in the Dutch waters: an exploration towards a definition of a shellfish bed*. Tech. rep. C057/24. Wageningen Marine Research, 2024.

[160] J. H. Damveld, K. J. van der Reijden, C. Cheng, L. Koop, L. R. Haaksma, C. A. Walsh, K. Soetaert, B. W. Borsje, L. L. Govers, P. C. Roos, H. Olff and H. S. J. M. H. 'Video transects reveal that tidal sand waves affect the spatial distribution of benthic organisms and sand ripples'. In: *Geophysical Research Letters* 45.21 (2018), pp. 11–837.

[161] B. Misiuk, V. Lecours and T. Bell. 'A multiscale approach to mapping seabed sediments'. In: *PLoS ONE* 13.2 (2018), e0193647.

[162] M. A. Oliver and R. Webster. 'Kriging: a method of interpolation for geographical information systems'. In: *International Journal of Geographical Information System* 4.3 (1990), pp. 313–332.

[163] T. C. Gaida, M. Snellen, T. A. van Dijk and D. G. Simons. 'Geostatistical modelling of multibeam backscatter for full-coverage seabed sediment maps'. In: *Hydrobiologia* 845 (2019), pp. 55–79.

[164] J. Li, A. D. Heap, A. Potter, Z. Huang and J. J. Daniell. 'Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin'. In: *Continental Shelf Research* 31.13 (2011), pp. 1365–1376.

[165] F. J. Boehmke, E. U. Schilling and J. C. Hays. 'Missing data in spatial regression'. In: *Keynote address, Society for Political Methodology Summer Conference, 23-25, July 2015, University of Rochester*. 2015.

[166] J. Calvert, J. A. Strong, M. Service, C. McGonigle and R. Quinn. 'An evaluation of supervised and unsupervised classification techniques for marine benthic habitat mapping using multibeam echosounder data'. In: *ICES Journal of Marine Science* 72.5 (2015), pp. 1498–1513.

[167] D. Li, C. Tang, C. Xia and H. Zhang. 'Acoustic mapping and classification of benthic habitat using unsupervised learning in artificial reef water'. In: *Estuarine, Coastal and Shelf Science* 185 (2017), pp. 11–21.

[168] J. Wan, Z. Qin, X. Cui, M. Yasir and B. Ma. 'Seafloor habitat mapping by combining multiple features from optic and acoustic data: A case study from Ganquan Island, South China Sea'. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 16 (2023), pp. 7248–7263.

[169] Y. Steiniger, D. Kraus and T. Meisen. 'Survey on deep learning based computer vision for sonar imagery'. In: *Engineering Applications of Artificial Intelligence* 114 (2022), p. 105157.

[170] X. Zhu and Z. Ghahramani. *Learning from labeled and unlabeled data with label propagation*. Tech. rep. Carnegie Mellon University, 2002.

[171] J. Shi and J. Malik. 'Normalized cuts and image segmentation'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.8 (2000), pp. 888–905.

[172] U. Von Luxburg. 'A tutorial on spectral clustering'. In: *Statistics and Computing* 17 (2007), pp. 395–416.

[173] U. Shaham, K. Stanton, H. Li, B. Nadler, R. Basri and Y. Kluger. 'Spectralnet: Spectral clustering using deep neural networks'. In: *Processings of the 6th International Conference on Learning Representations*. 2018.

[174] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin. 'Attention is all you need'. In: *Advances in Neural Information Processing Systems* 30 (2017).

[175] M. P. Fay and M. A. Proschan. 'Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules'. In: *Statistics Surveys* 4 (2010), p. 1.

[176] H. W. Wendt. 'Dealing with a common problem in social science: a simplified rank-biserial coefficient of correlation based on the U statistic.' In: *European Journal of Social Psychology* 2.4 (1972).

[177] P. Legendre and L. Legendre. *Numerical ecology*. Vol. 24. Elsevier, 2012.

Acknowledgements

Two months after submitting this thesis, I still cannot believe that my PhD journey is approaching its end. I was lucky to come across this project just as my passion for research started, and I still feel grateful that I could engage in this topic, which is truly interdisciplinary as well as challenging.

I could not have finished my dissertation in time without the help of many people. First, I would like to express my gratitude to my promotors, Mirjam and Dick. Thanks for your guidance, trust, and encouragement that shaped this thesis and my development as a researcher. I am grateful to Sebastiaan, who has contributed to every aspect of this project. You added a great deal as an ecologist, and without you it would have been much harder for me to accomplish all these research outputs, offshore surveys, and conference trips. I want to acknowledge Alireza as well. Although our collaboration started midway through my PhD, you helped me a lot with developing my research ideas into experiments and papers. I would like to mention Tannaz, who provided important support at the beginning of this project by introducing me to the code and software, and by taking the lead in the first measurement campaign.

Special thanks go to all the project users at Rijkswaterstraat, Deltares, QPS, R2Sonic, Van Oord, Boskalis, and Maritiem Instituut Willem Barentsz. Your input was highly valuable to my research. I also want to acknowledge all surveyors, crew members, and sedimentology and biology experts onboard. Without you, we could not have obtained the amazing datasets that form the basis of this thesis. The seasickness while staring at the computer screen at 2 a.m. was not fun at all, but it was worth it.

From the bottom of my heart, I would like to thank my colleagues in O&E for the walks, laughter, coffee breaks, and all the support along the way. Within and beyond the university, I truly appreciate all the friends I have met in Delft. Looking back six years, I would not have imagined that the decision to study abroad would bring me such genuine friendships that I will cherish for the rest of my life. Finally, I am deeply grateful to my partner and my parents for your continued support and love.

December 2025

Curriculum Vitæ

Qian Bai

02-12-1997 Born in Shanxi, China.

Education

2015–2019 Bachelor of Engineering in Remote Sensing Science & Technology
Wuhan University, China

2019–2021 Master of Science in Geoscience & Remote Sensing
Delft University of Technology, The Netherlands

2021–2026 Ph.D. in Underwater Acoustic Mapping
Delft University of Technology, The Netherlands
Thesis: Multi-frequency Acoustic Mapping of Marine Benthos
Promotors: Prof. dr. D.G. Simons and Prof. dr. ir. M. Snellen

Awards

2021 M.Sc. *cum laude*, Delft University of Technology

2019 Sino-Dutch Bilateral Exchange Scholarship,
Dutch Ministry of Education, Culture and Science

List of Publications

Peer-reviewed journal articles

4. **Q. Bai**, S. Mestdagh, A. Amiri-Simkooei, T. van Dijk, D. G. Simons and M. Snellen. *Multi-frequency seabed shellfish habitat mapping with semi-supervised multibeam backscatter completion and unsupervised spectral learning.* Under review.

3. **Q. Bai**, A. Amiri-Simkooei, S. Mestdagh, D. G. Simons and M. Snellen. *Unsupervised seabed habitat change detection with multibeam backscatter data using a constrained Gaussian Mixture Model.* Applied Acoustics 246 (2026), p. 111222.

2. **Q. Bai**, A. Amiri-Simkooei, S. Mestdagh, D. G. Simons and M. Snellen. *Mussel culture monitoring with semi-supervised machine learning on multibeam echosounder data using label spreading.* Journal of Environmental Management 369 (2024), p. 122250.

1. **Q. Bai**, S. Mestdagh, M. Snellen and D. G. Simons. *Indications of marine benthos occurrence from multi-spectral multibeam backscatter data: A case study in the North Sea.* Frontiers in Earth Science 11 (2023), p. 1140649.

Conference papers

3. **Q. Bai**, S. Mestdagh, A. Amiri-Simkooei, T. van Dijk and M. Snellen. *Accurate mapping of an extended shell bed area in the North Sea using multi-spectral multibeam backscatter data.* ISPRS 3D Underwater Mapping Worshop 2025, Vienna, Austria.

2. **Q. Bai**, S. Mestdagh, M. Snellen and A. Amiri-Simkooei. *Characterizing seabed sediments using multi-spectral backscatter data in the North Sea.* IEEE OCEANS 2023, Limerick, Ireland.

1. S. Mestdagh, **Q. Bai** and M. Snellen. *Mapping biotic and abiotic seafloor habitat characteristics with multi-spectral multibeam backscatter data.* IEEE OCEANS 2023, Limerick, Ireland.