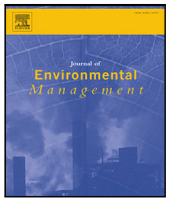
Journal of Environmental Management 369 (2024) 122250

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman



Research article

Mussel culture monitoring with semi-supervised machine learning on multibeam echosounder data using label spreading

Qian Bai [a,*], Alireza Amiri-Simkooei [a], Sebastiaan Mestdagh [a,b], Dick G. Simons [a], Mirjam Snellen [a]

[a] Department of Control and Operation, Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS, Delft, The Netherlands
[b] Department of Ecosystems and Sediment Dynamics, Deltares, Delft, 2629 HV, The Netherlands

ARTICLE INFO

Keywords:
Mussel culture monitoring
Seabed classification
Multibeam echosounder
Seabed geomorphology
Acoustic backscatter
Semi-supervised machine learning

ABSTRACT

High diversity seabed habitats, such as shellfish aggregations, play a significant role in marine ecosystem sustainability but are susceptible to bottom disturbance induced by anthropogenic activities. Regular monitoring of these habitats with effective mapping methods is therefore essential. Multibeam echosounder (MBES) has been widely used in recent decades for seabed characterization due to its non-destructive manner and extensive spatial coverage compared to traditional methods like bottom sampling. Nevertheless, bottom sampling remains essential to link ground truth with acoustic seabed classification. Using seabed samples and MBES measurements, machine learning techniques are commonly employed to model their relationships and generate classification maps of an extended seabed. However, limited ground truth data, resulting from constraints in regulations, budget, or time, may impede the development of robust machine learning models. To address this challenge, we applied a semi-supervised machine learning method to classify seabed sediments of a blue mussel (*Mytilus edulis*) cultivation area in the Oosterschelde, the Netherlands. We utilized nine boxcore samples to generate pseudo-labels on MBES data. These pseudo-labels enlarged the training data size, facilitated the training of three comprehensive machine learning algorithms (Gradient Boosting, Random Forest, and Support Vector Machine), and helped to classify the study site into mussel and non-mussel areas. We found the geomorphological and backscatter-related features to be complementary for mussel culture detection. Our classification results were demonstrated effective through expert knowledge of this cultivation area and brought insights for future research on natural mussel habitats.

1. Introduction

Preserving seabed habitats, such as oyster reefs and shellfish aggregations, is crucial for ensuring sustainable utilization of the ocean environment (Brown and Collier, 2008; Brown et al., 2011). 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, human activities such as fisheries are strictly regulated in marine protected areas (Teixeira et al., 2013; Diesing et al., 2020). Furthermore, in the North Sea, extracting sand has been prohibited within a distance of 100 m from the living shellfish beds since 2008 (Ministerie van Verkeer en Waterstaat, 2010). These obligations, on the other hand, require effective methods for mapping the seabed habitats, especially the occurrence of marine benthos (Ierodiaconou et al., 2011). Traditional seabed biodiversity monitoring relies on bottom sampling (e.g., boxcores, grabs, and dredges), which provides detailed seabed information but is labor-intensive and destructive. Moreover, it can only deliver sparsely distributed measurements. To achieve a spatially continuous map product, assumptions about areas between the sampling stations are often needed (Herkül et al., 2017; Norberg et al., 2019).

Acoustic remote sensing techniques offer efficient alternatives (Van Walree et al., 2005; Simons and Snellen, 2009; Snellen et al., 2011; McGonigle and Collier, 2014). Multibeam echosounder (MBES) has been widely used over the past decades to survey extended parts of the seafloor. MBES emits pings in a wide swath perpendicular to the sailing direction. Moreover, the beam steering technique implemented in MBES helps to distinguish signals backscattered from different directions. Combined with the navigation system and motion sensors on the surveying vessel, large-scale mapping of the seafloor can be realized using MBES. 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 at a certain incident angle as

* Corresponding author.
E-mail address: q.bai@tudelft.nl (Q. Bai).

https://doi.org/10.1016/j.jenvman.2024.122250
Received 17 May 2024; Received in revised form 9 August 2024; Accepted 17 August 2024
Available online 29 August 2024

well as the backscatter angular variation characteristics (Lurton, 2002; Lamarche and Lurton, 2018). Some marine benthos such as oysters and mussels can affect the sediment geoacoustic properties due to the presence of their shells (Hutin et al., 2005). Their aggregations can 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 (Snellen et al., 2008).

While 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. Physical models have been developed to describe the backscatter strength of various sediment types (Jackson, 1994), providing a powerful tool for sediment classification using only acoustic measurements. However, how to account for the occurrence of marine benthos in the physical models is still an ongoing research topic. Some studies assessed the impact of benthos on seafloor backscatter by examining deviations from model predictions based solely on sediment properties (Lee et al., 2022). Ideally, these models also require calibrated backscatter strengths as input. Due to the varied sensitivity of echosounders in different water environments (Roche et al., 2018), absolute backscatter calibration is difficult to achieve for every single survey and requires a deeper understanding of the varying sonar mechanical properties. By contrast, empirical or data-driven methods can be more applicable to uncalibrated backscatter data from a single survey, offering an alternative solution for seabed sediment classification and benthos monitoring.

Data-driven methods can be distinguished based on how to combine ground truth information in seabed classification. For instance, unsupervised methods first identify patterns in acoustic data and assign these patterns to certain classes according to the ground truth. These methods range from statistical approaches based on assumptions for the probability distribution function of backscatter strengths (Simons and Snellen, 2009), to machine learning methods such as K-means clustering (Seber, 2009). On the other hand, supervised methods fit a model between predictor variables and the target, which are used to predict seabed habitat classes as a function of acoustic features (Misiuk and Brown, 2023). Various machine learning and deep learning models have been applied in supervised seabed habitat mapping (Diesing et al., 2014; Cui et al., 2021; Ji et al., 2024). Sophisticated models improve the ability to deal with multi-dimensional features and complex predictor-outcome relations, but often require a large amount of training data (Safonova et al., 2023). In marine applications, however, obligations such as minimizing the seabed disturbance along with restrictions in time and budget usually result in selecting a limited number of ground truth stations, where bottom samples and/or underwater photographs are taken. Limited ground truth can pose challenges to establishing the relationships between MBES data and seabed properties using supervised methods.

Regardless of the limited ground truth samples, 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 (Minelli et al., 2021; Asghar et al., 2020), which is another 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 (Sheikhpour et al., 2017). Semi-supervised machine learning techniques have been successfully adopted in many fields, including classifying MBES data. For instance, Minelli et al. (2021) 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 research, 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, specifically label spreading (Zhou et al., 2003), to MBES bathymetry and backscatter data to facilitate 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 (Ligthart et al., 2021). To further ensure the performance of label spreading, we also proposed to optimize the hyperparameter selection during graph construction by assessing the Silhouette coefficient of pseudo-labels.

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 product 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.

2. Study area and dataset

The study area is located in the Oosterschelde, a tidal basin in the southwest of the Netherlands (Fig. 1). The Oosterschelde has an important economic value due to shellfish culture and has been intensively researched during the past decades (Nieuwhof, 2018). 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 (Capelle, 2017). 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. 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 [dB], were saved in the Kongsberg logging format *.kmall*.

For ground truthing, sediment samples at 9 stations (Fig. 1) were collected using a cylindrical boxcore sampler with a diameter of 30 cm. Grain size distributions were determined afterwards through laboratory analysis. Median grain size of the samples ranged from 114.0 to 264.8 μm (Table 1). In addition, according to the particle size distribution classification (Blott and Pye, 2012), 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 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. 1) was covered by mussel seeds and

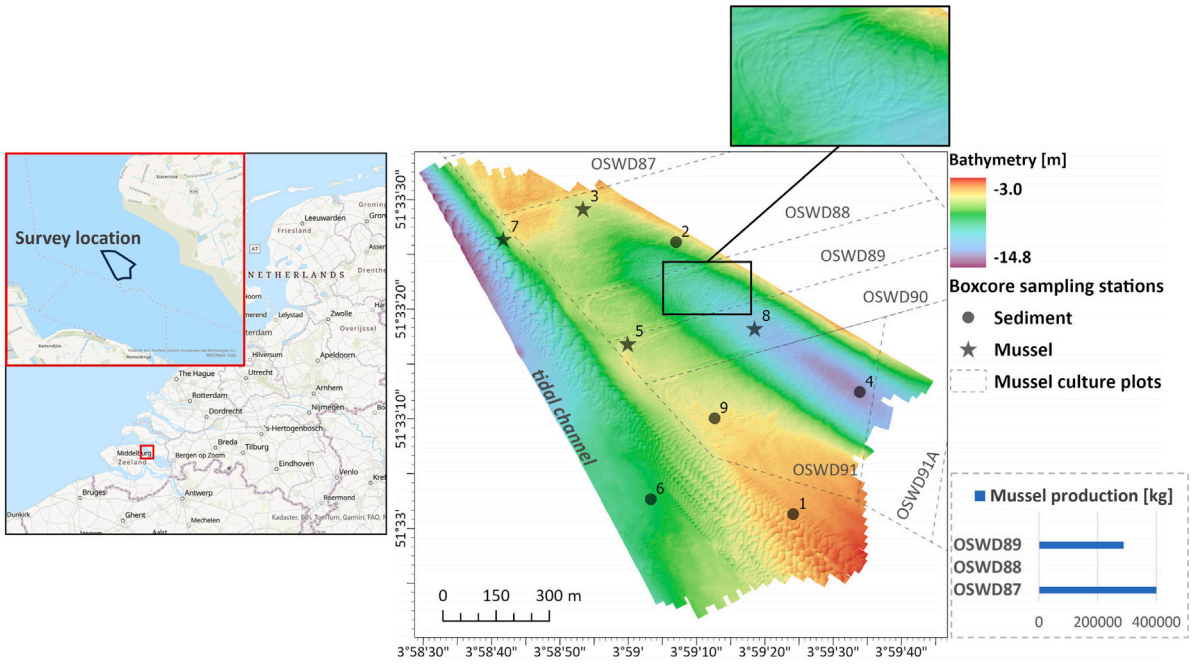


Fig. 1. Study area. (Left) Location of the surveyed area in the Oosterschelde tidal basin, the Netherlands; (Right) Bathymetry map acquired by the MBES, with a grid size of 0.25 × 0.25 m. East of a tidal channel, mussel occurrence of nine boxcore samples and mussel culture plots covered by the study area are also 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) were acquired from the Dutch Mussel Auction.

Table 1
Detailed properties of 9 boxcore samples.

Sampling stations	Median grain size	%mud	Description of mussel occurrence	Class	Boxcore picture example
1	215.4 μm	4.0	No mussels	Sediment	
2	204.9 μm	7.0			
4	179.9 μm	5.8			
6	258.7 μm	0.0			
9	200.2 μm	6.6			
3	170.8 μm	34.2	A few living mussels	Mussel	
7	264.8 μm	0.0			
8	178.2 μm	12.3	30% covered by mussels	Mussel	
5	114.0 μm	42.4	70% covered by mussels	Mussel	

medium-sized mussels of 400,000 kg at the time of our survey, while OSWD88 (Fig. 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, Ministry of Infrastructure and Water Management of the Netherlands, Hydrographic Service of the Royal Netherlands Navy, and the European Marine Observation and Data Network).

3. Methodology

Our workflow for detecting the mussel occurrence from the MBES data comprises three main steps (see Step 1–3 in Fig. 2). Initially, we extracted and selected secondary features from the MBES measurements (Section 3.1). These features, combined with the ground truth samples, were then employed to generate pseudo-labels (Section 3.2). Following this, we trained machine learning models using the expanded labeled dataset and conducted classification across the entire study area (Section 3.3). In Section 3.4, we present the validation procedure of our semi-supervised classification results, including comparisons with the full-supervised method.

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. We processed the MBES bathymetry and backscatter data using both software and in-house MATLAB scripts. Before feature extraction, it is important to clean the data and apply all required corrections.

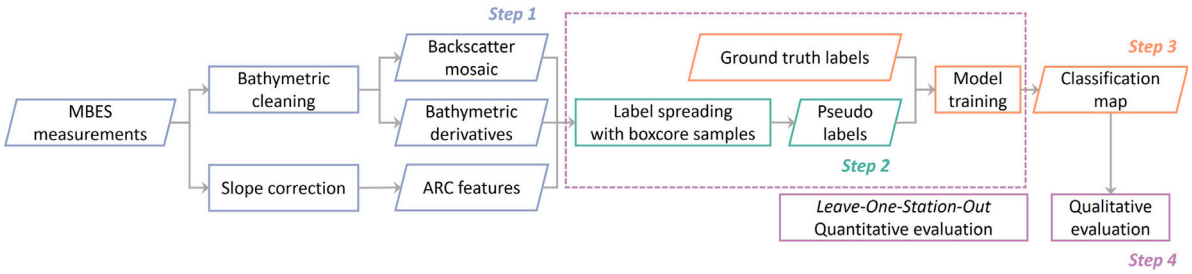


Fig. 2. Workflow of the semi-supervised mussel classification method using the MBES data.

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. 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 (Mokarram and Sathyamoorthy, 2018). 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 (MacMillan and Shary, 2009). 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 (Kowalczyk, 2008), 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.

Backscatter strength is also a powerful discriminator in seafloor classification and benthic habitat monitoring (Misiuk and Brown, 2023). Although absolute backscatter calibration was not available, uncalibrated backscatter strengths collected within our survey can still be comparable. We considered 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 (Janowski et al., 2018), which requires setting a subset of incident angles (usually around 45° (Lamarche et al., 2011)) as reference. Then backscatter strengths at other angles are normalized to this reference.

The angular variation itself, however, is also intrinsic to seabed properties (Lamarche and Lurton, 2018). To account for this variation in our analysis, we extracted the incident angle and backscatter strength of each beam from the *.kmall* files using the CoFFee library in MATLAB. Before building the ARCs, a slope correction (Gaida et al., 2018) was conducted to achieve the actual incident angle ϕ [°] relative to the local seabed according to

$$\cos(\phi) = \frac{\sin(90-\phi_{fl}) + \cos(90-\phi_{fl})\epsilon_{ac}}{\sqrt{1+\epsilon_{al}^2+\epsilon_{ac}^2}} \quad (1)$$

where ϕ_{fl} [°] is the beam angle with respect to the flat seabed. ϵ_{ac} and ϵ_{al} [radians], the across- and along-track slope, were calculated from the bathymetry data with respect to the vessel heading and sonar axis via a 2D finite difference method using a moving average window of 60 pings, covering a distance about 12 m in our study area. We then subtracted a term $10\log A$ from the backscatter strength of each beam to account for the changes in the ensonified footprint area. The footprint area A [m^2] was the smaller value between the pulse-limited

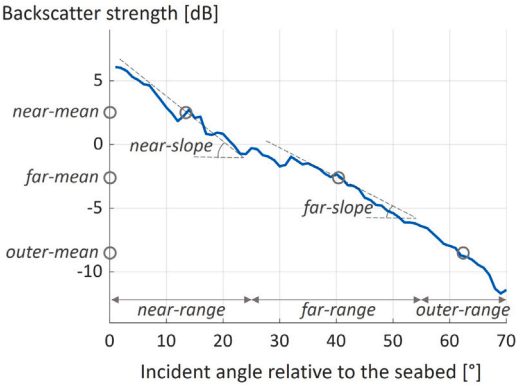


Fig. 3. Illustration of the extracted features from a half-swath ARC.

footprint area A_p and the beam-limited footprint area A_b. A_p and A_b are calculated as

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

and

$$A_b = R^2 \Omega_t \Omega_r, \quad (3)$$

with R [m] the slant range of each sounding, c [m/s] the sound speed, and τ_e [s] the effective pulse length. Ω_t and Ω_r are the beam opening angles for transmission and reception, respectively.

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) (Fonseca and Mayer, 2007). The mean values of the near- and far-range 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. 3). For the outer-range, only the mean value was calculated, since outer-mean can be the most relevant to the critical angle of reflection at the sediment-water interface (Fonseca and Mayer, 2007).

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 4.1).

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. (2003). 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 $\chi = \{x_1, \ldots, x_l, x_{l+1}, \ldots, x_n\} \subset \mathbb{R}^m$, each point is represented by a feature vector with length m. The first l points have labels from the label set $\mathscr{L} = \{1, \ldots, c\}$, 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 (KNN) to ease the computational burdens. The vertices V equals the point set χ. The edges E represent the connections between data points and are weighted by a similarity matrix W, with $W_{ij} = \exp(-\|x_i - x_j\|^2/2\sigma^2)$ if $i \neq j$ and $W_{ii} = 0$. σ^2 is a scaling parameter.

In label spreading, the classification rule of χ is denoted by a $n \times c$ matrix $F = [F_1^T, \ldots, F_n^T]^T$. F_i^T represents the classification rule for point x_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 x_i is labeled as $y_i = u \in \{1, \ldots, c\}$. $Y_{iu} = 0$ if x_i is unlabeled. The label spreading algorithm further defines the cost function associated with F as

$$Q(F) = \frac{1}{2}\left(\sum_{i,j=1}^{n} W_{ij}\left\|\frac{1}{\sqrt{D_{ii}}}F_i - \frac{1}{\sqrt{D_{jj}}}F_j\right\|^2 + \mu\sum_{i=1}^{n}\|F_i - Y_i\|^2\right), \quad (4)$$

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 (Ligthart et al., 2021). 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 Eq. (4) 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 χ 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 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 KNN 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 KNN 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 (Rousseeuw, 1987) 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.

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 (Loureiro et al., 2024): 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 (Hastie et al., 2009). 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 3.4).

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. 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. 2). LOSO was built upon the concept of Leave-One-Out cross-validation (Hastie et al., 2009) that is commonly used in machine learning for small datasets. As explained in Section 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. Afterwards, 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.

Accuracy is the proportion of correctly labeled points, while F1 score is calculated according to

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \quad (5)$$

where Precision measures how many positive predictions are correct. Recall is calculated as the number of true positives divided by the sum of true positives and false negatives. The F1 score is a more objective metric than the accuracy when class imbalance is present. Both accuracy and F1 score have the best value of 1 and the worst value of 0.

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 (Menze et al., 2009). Thus, it can be a measure of the significance of each feature to the overall predictive performance of the combined models.

4. Results and discussion

4.1. Feature extraction and selection

Different spatial patterns per region can be observed from the extracted MBES features (Fig. 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

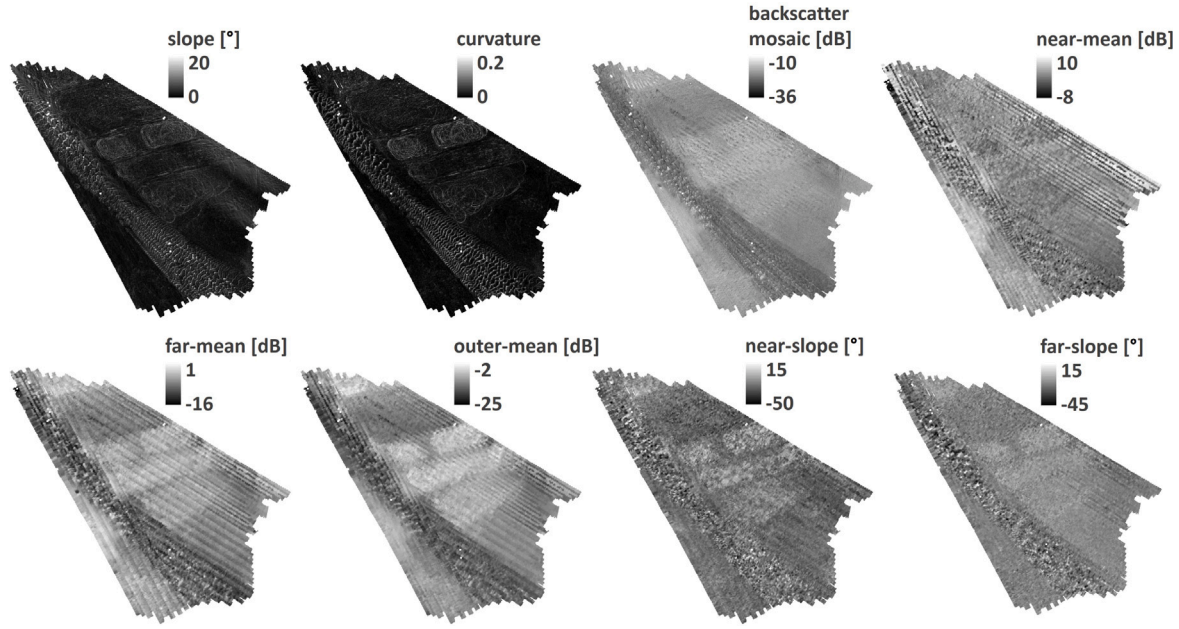


Fig. 4. Features extracted from the MBES measurements.

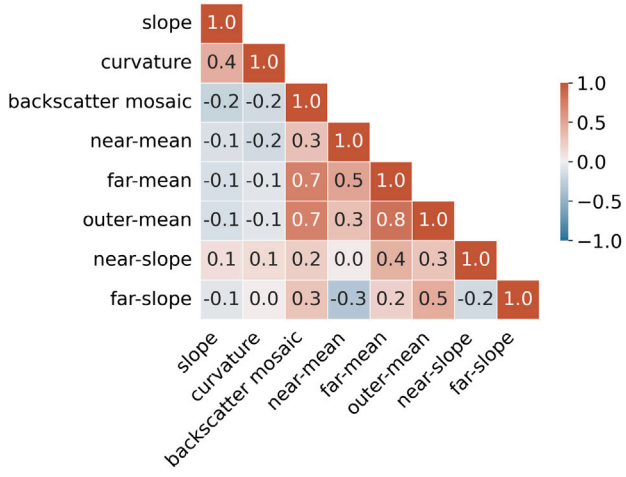


Fig. 5. Correlation among all extracted MBES features, annotated by Pearson's coefficient.

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). To reduce redundant information, we eliminated far-mean from the feature set. There also exist a large number of striped artefacts in far-mean, so this exclusion will avoid bringing these artefacts to the classification results. The remaining seven features were then considered in pseudo-labeling and model training.

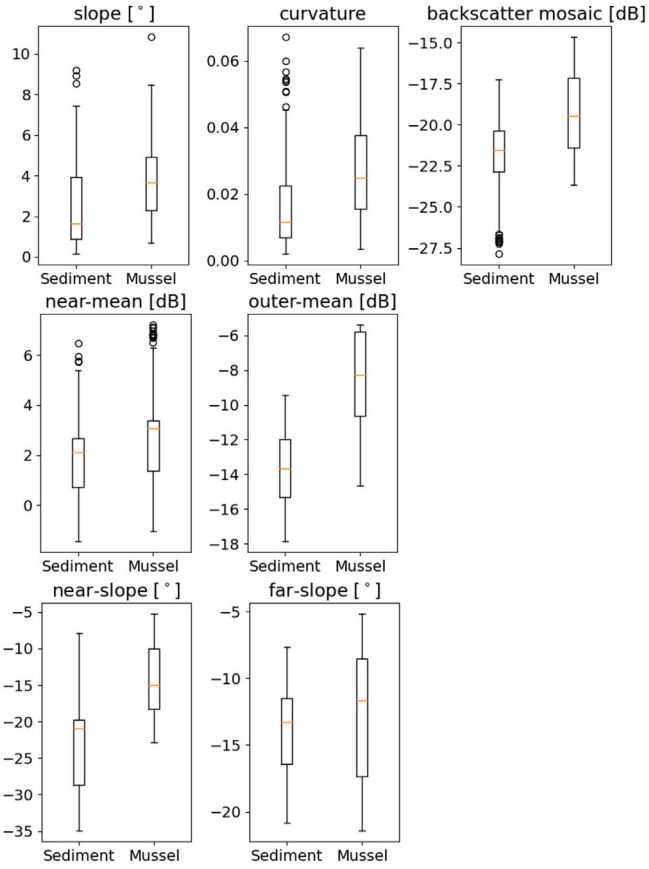


Fig. 6. Boxplots for the selected MBES features of the labeled data, with the sediment and mussel class plotted separately. Features were collected within a radius of 3 m around each boxcore sampling station.

Using a radius of 3 m around the nine sampling stations (Fig. 1), we collected 284 MBES points in total. These points were labeled with

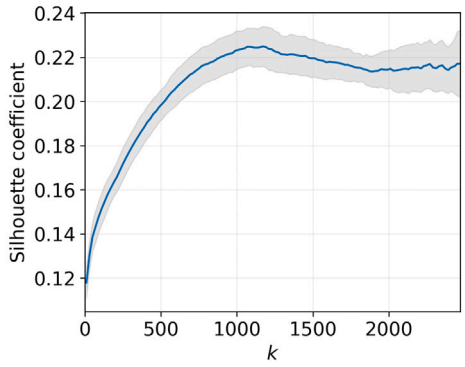


Fig. 7. Silhouette coefficient of 5000 pseudo-labels with different k in label spreading. The solid curve and shaded area indicate the mean and standard deviation of randomly selecting 5000 points and running label spreading ten times.

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. 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 (Jackson, 1994).

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 KNN graph. The Silhouette coefficient of the pseudo-labeled data reaches the peak when k is around 1200 (Fig. 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 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.

We also present the pseudo-labeling results in the feature space using principal component analysis (PCA) (Greenacre et al., 2022) of all selected features (Fig. 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.

4.3. Semi-supervised classification results

4.3.1. 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 depth of 3 and 4, respectively. Regarding RF, 5 trees and 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 compared the classification maps achieved by the full- and semi-supervised methods (Fig. 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 artefacts 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 (Safonova et al., 2023). The tidal channel located west of the study area is not distinguished 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 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. 10). The three machine learning algorithms, GB, RF, and SVM, produced consistent ARC shapes. ARCs of the mussel class show larger backscatter strengths 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.

4.3.2. Feature importance in full- and semi-supervised classification

As explained in Section 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.

Regarding GB, the full-supervised classification depended mostly on outer-mean and near-mean (Fig. 11). Geomorphological features, which can provide spatial patterns of mussel cultivation, were 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 was more balanced, and near-slope played 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 made better use of all features and brought 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 exhibited significantly improved consistency. Although outer-mean was still the most influential feature,

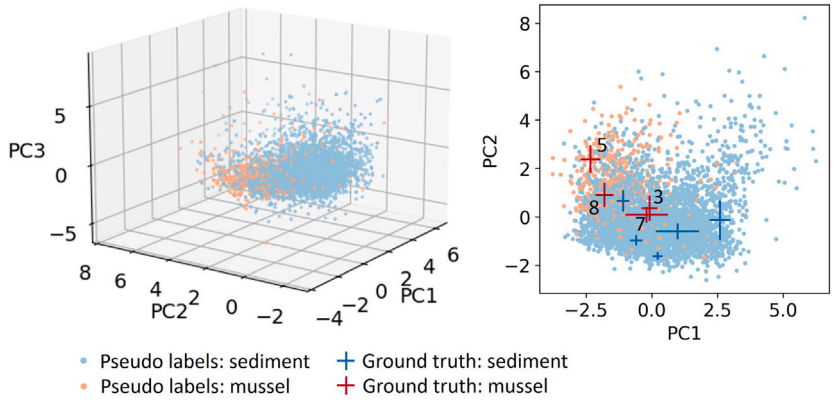


Fig. 8. Visualization of 5000 pseudo-labeled points (with the optimal k = 1200) in the feature space using PCA of the selected MBES features. (Left) 5000 randomly sampled points plotted on the span of the first three PCs, with the point color defined by their pseudo-labels; (Right) Comparison between 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. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

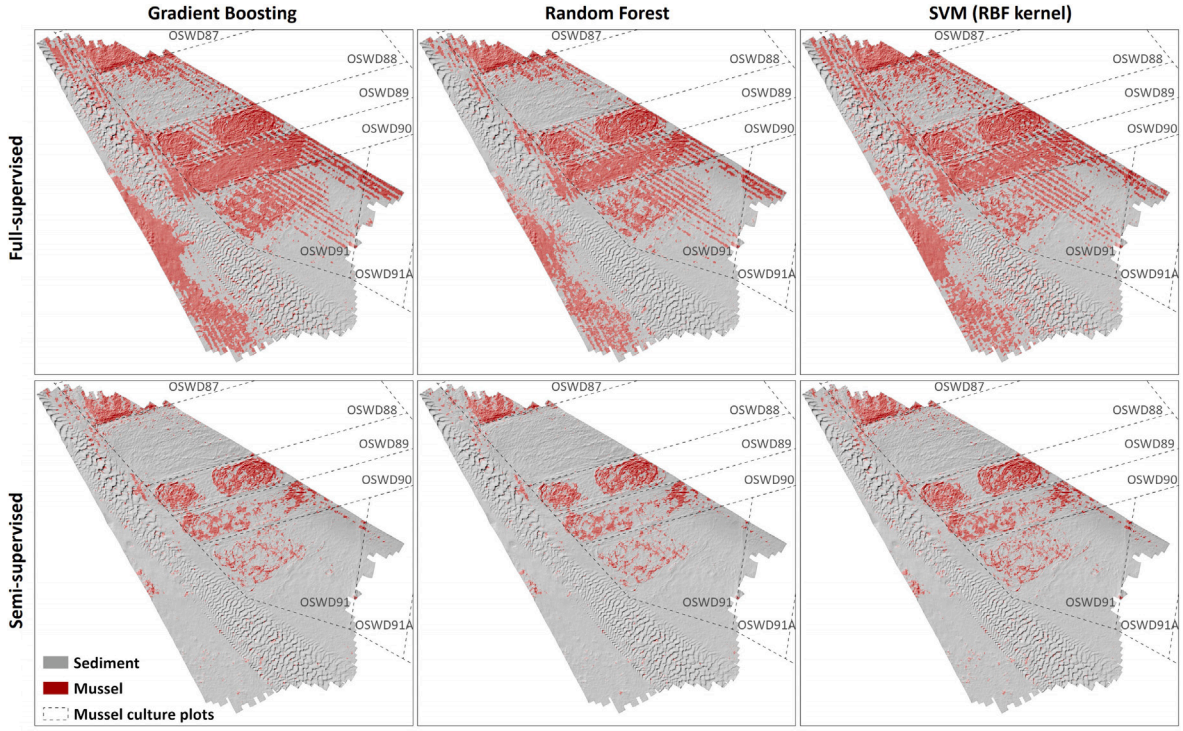


Fig. 9. Prediction maps of the entire study area from the full- and semi-supervised classification methods, shaded by bathymetry.

the importance of slope and curvature increased in both GB and RF classification. Far-slope was 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.

4.3.3. 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 (Misiuk et al., 2019; Mastrantonis et al., 2024).

Among the full-supervised learning methods, GB delivered the best accuracy and F1 score (Table 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 were generally consistent, with RF showing the largest fluctuations. Although misclassification and noise were shown to be reduced in the semi-supervised prediction maps, the cross-validation scores were lower than the full-supervised methods. This indicates that the full-supervised classification models were better fitted at the ground truth sampling locations.

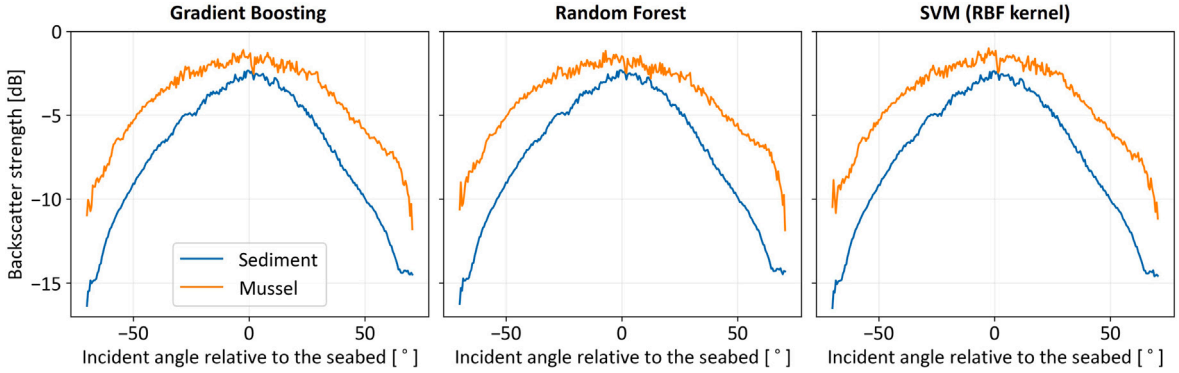


Fig. 10. Averaged ARC of the sediment and mussel class from the semi-supervised learning results.

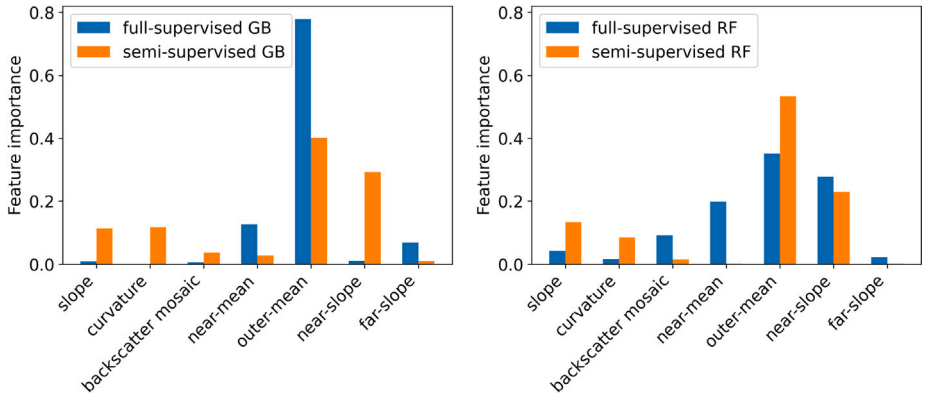


Fig. 11. Comparison between feature importance of the full- and semi-supervised classification using GB and RF.

Table 2
Averaged accuracy and F1 score of the LOSO cross-validation. Mean and standard deviation of semi-supervised results with 3 sets of randomly sampled MBES points for label spreading are presented.

Accuracy \ F1 score	Gradient Boosting	Random Forest	SVM (RBF kernel)
Semi-supervised methods	0.53 ± 0.000 / 0.57 ± 0.005	0.51 ± 0.017 / 0.54 ± 0.022	0.54 ± 0.009 / 0.57 ± 0.005
Full-supervised methods	0.72 / 0.78	0.64 / 0.72	0.58 / 0.66

Details regarding the predictions in LOSO cross-validation are also provided (see Fig. 12), further showing that semi-supervised classification using three different machine learning algorithms achieved 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 delivered the most aligned results among the full-supervised classification methods. With 'sediment' seabed samples 1 and 2 used for validation, full-supervised methods tended 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 failed 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 were 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 predicted 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 was not used for pseudo-labeling, critical information regarding the MBES features of the sediment-mussel mixture (Fig. 8) could be lost. For sample 5, both full- and semi-supervised classification indicated the mussel presence, but semi-supervised 'mussel' predictions were less. In general, semi-supervised methods tended to predict the mussel occurrence more conservatively than full-supervised classification.

Comparison between the predictions and sample descriptions (Fig. 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 delivered lower validation scores for the ground truth sampling stations, the overall prediction performance within the study area was proven better (Section 4.3.1). 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 helped 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

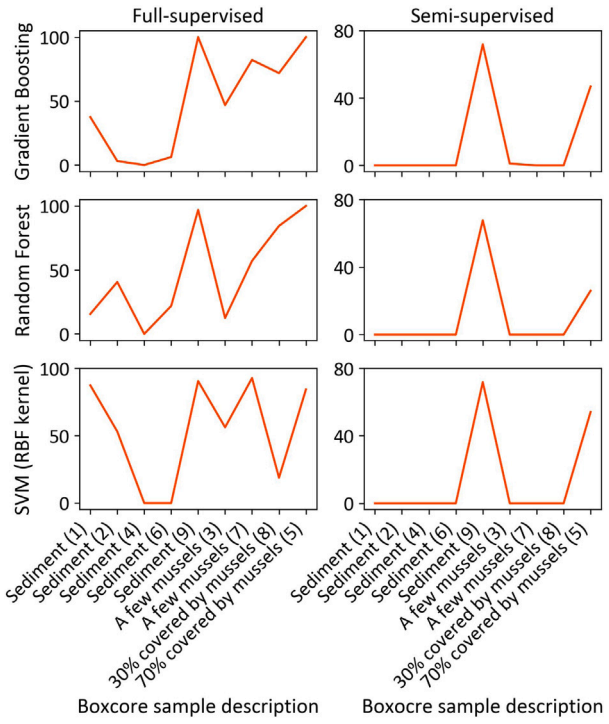


Fig. 12. Percentage of points predicted as 'mussel' in each round of LOSO cross-validation for (Left) full-supervised and (Right) semi-supervised classification. Descriptions of the ground truth sample reserved for validation in each round are presented and sorted based on the mussel coverage. The semi-supervised results are averaged values based on 3 sets of randomly sampled MBES points for label spreading.

scales can be helpful for natural seabed habitat monitoring (Nieuwhof, 2018).

4.4. Limitations and future work

In this research, we investigated 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. 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 showed the advantage of a semi-supervised machine learning approach for detecting the occurrence of mussels. We used label spreading to generate pseudo-labels, which helped to enlarge the training data size and alleviated overfitting. Trained with the pseudo-labeled data, the prediction maps of three machine learning algorithms (Gradient Boosting, Random Forest, and Support Vector Machine) 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 the backscatter strength 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.

CRediT authorship contribution statement

Qian Bai: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alireza Amiri-Simkooei:** Writing – review & editing, Supervision, Methodology, Investigation. **Sebastiaan Mestdagh:** Writing – review & editing, Validation, Supervision. **Dick G. Simons:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Mirjam Snellen:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research is part of the NWO-funded project 18698 "Multi-spectral Multi-beam Imaging for Mapping the Occurrence of Marine Benthos". The authors express their gratitude to the Ministry of Infrastructure and Water Management of the Netherlands (Rijkswaterstaat) for the support in data collection and seabed sample analysis. Special thanks to Douwe van den Ende from Wageningen Marine Research for providing the information on mussel culture plots and essential contacts. Peter van den Boomgaard from the Dutch Mussel Auction is acknowledged for providing the mussel production information of the study area. The company QPS is thanked for the hydrographic software licenses. R2Sonic and Van Oord are also recognized for providing sonar equipment. Finally, the authors thank Boskalis and Deltares for their technical support.

References

Asghar, S., Choi, J., Yoon, D., Byun, J., 2020. Spatial pseudo-labeling for semi-supervised facies classification. J. Pet. Sci. Eng. 195, 107834.

Blott, S.J., Pye, K., 2012. Particle size scales and classification of sediment types based on particle size distributions: Review and recommended procedures. Sedimentology 59 (7), 2071–2096.

Brown, C.J., Collier, J.S., 2008. Mapping benthic habitat in regions of gradational substrata: an automated approach utilising geophysical, geological, and biological relationships. Estuar. Coast. Shelf Sci. 78 (1), 203–214.

Brown, C.J., Smith, S.J., Lawton, P., Anderson, J.T., 2011. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuar. Coast. Shelf Sci. 92 (3), 502–520.

Capelle, J.J., 2017. Production Efficiency of Mussel Bottom Culture (Ph.D. thesis). Wageningen University and Research.

Cui, X., Yang, F., Wang, X., Ai, B., Luo, Y., Ma, D., 2021. Deep learning model for seabed sediment classification based on fuzzy ranking feature optimization. Mar. Geol. 432, 106390.

Diesing, M., Green, S.L., Stephens, D., Lark, R.M., Stewart, H.A., Dove, D., 2014. Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Cont. Shelf Res. 84, 107–119.

Diesing, M., Mitchell, P.J., O'Keeffe, E., Gavazzi, G.O.M., Bas, T.L., 2020. Limitations of predicting substrate classes on a sedimentary complex but morphologically simple seabed. Remote Sens. 12 (20), 3398.

Fonseca, L., Mayer, L., 2007. Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. Mar. Geophys. Res. 28, 119–126.

Gaida, T.C., Tengku Ali, T.A., Snellen, M., Amiri-Simkooei, A., Van Dijk, T.A., Simons, D.G., 2018. A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data. Geosciences 8 (12), 455.

Greenacre, M., Groenen, P.J., Hastie, T., d'Enza, A.I., Markos, A., Tuzhilina, E., 2022. Principal component analysis. Nat. Rev. Methods Prim. 2 (1), 100.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2, Springer.

Herkül, K., Peterson, A., Paekivi, S., 2017. Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows. Estuar. Coast. Shelf Sci. 192, 57–71.

Hutin, E., Simard, Y., Archambault, P., 2005. Acoustic detection of a scallop bed from a single-beam echosounder in the St. Lawrence. ICES J. Mar. Sci. 62 (5), 966–983.

Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L., Versace, V., 2011. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. Cont. Shelf Res. 31 (2), S28–S38.

Jackson, D.R., 1994. APL-UW High-Frequency Ocean Environmental Acoustic Models Handbook. Technical Report, Applied Physics Laboratory, University of Washington.

Janowski, L., Trzcinska, K., Tegowski, J., Kruss, A., Rucinska-Zjadacz, M., Pocwiardowski, P., 2018. 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. Remote Sens. 10 (12), 1983.

Ji, X., Yang, B., Wei, Z., Wang, M., Tang, Q., Xu, W., Wang, Y., Zhang, J., Zhang, L., 2024. Benthic habitat sediments mapping in coral reef area using amalgamation of multi-source and multi-modal remote sensing data. Remote Sens. Environ. 304, 114032.

Kowalczyk, D., 2008. Casorati curvatures. Bull. Transilvania Univ. Brasov Ser. III: Math. Inform. Phys. 1 (50).

Lamarche, G., Lurton, X., 2018. Recommendations for improved and coherent acquisition and processing of backscatter data from seafloor-mapping sonars. Mar. Geophys. Res. 39 (1–2), 5–22.

Lamarche, G., Lurton, X., Verdier, A.-L., Augustin, J.-M., 2011. Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—Application to Cook Strait, New Zealand. Cont. Shelf Res. 31 (2), S93–S109.

Lee, K.M., Venegas, G.R., Ballard, M.S., Dorgan, K.M., Kiskaddon, E., McNeese, A.R., Wilson, P.S., 2022. Impacts of infauna, worm tubes, and shell hash on sediment acoustic variability and deviation from the viscous grain shearing model. J. Acoust. Soc. Am. 152 (4), 2456–2474.

Ligthart, A., Catal, C., Tekinerdogan, B., 2021. Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification. Appl. Soft Comput. 101, 107023.

Loureiro, G., Dias, A., Almeida, J., Martins, A., Hong, S., Silva, E., 2024. A survey of seafloor characterization and mapping techniques. Remote Sens. 16 (7), 1163.

Lurton, X., 2002. An Introduction to Underwater Acoustics, second ed. Springer, Berlin, Heidelberg.

MacMillan, R.A., Shary, P., 2009. Landforms and landform elements in geomorphometry. Dev. Soil Sci. 33, 227–254.

Mastrantonis, S., Radford, B., Langlois, T., Spencer, C., de Lestang, S., Hickey, S., 2024. A novel method for robust marine habitat mapping using a kernelised aquatic vegetation index. ISPRS J. Photogramm. Remote Sens. 209, 472–480.

McGonigle, C., Collier, J.S., 2014. Interlinking backscatter, grain size and benthic community structure. Estuar. Coast. Shelf Sci. 147, 123–136.

Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC Bioinform. 10, 1–16.

Minelli, A., Tassetti, A.N., Hutton, B., Pezzuti Cozzolino, G.N., Jarvis, T., Fabi, G., 2021. 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. Sensors 21 (9), 2999.

Ministerie van Verkeer en Waterstaat, 2010. Beleidsregels Ontgrondingen in Rijkswateren. Technical Report.

Misiuk, B., Brown, C.J., 2023. Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor. Estuar. Coast. Shelf Sci. 108599.

Misiuk, B., Diesing, M., Aitken, A., Brown, C.J., Edinger, E.N., Bell, T., 2019. A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using random forest. Geosciences 9 (6), 254.

Mokarram, M., Sathyamoorthy, D., 2018. A review of landform classification methods. Spat. Inf. Res. 26 (6), 647–660.

Nieuwhof, S., 2018. The Use of Remote Sensing to Reveal Landscape-Scale Ecosystem Engineering by Shellfish Reefs (Ph.D. thesis). University of Twente.

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., et al., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol. Monogr. 89 (3), e01370.

Roche, M., Degrendele, K., Vrignaud, C., Loyer, S., Le Bas, T., Augustin, J.-M., Lurton, X., 2018. Control of the repeatability of high frequency multibeam echosounder backscatter by using natural reference areas. Mar. Geophys. Res. 39, 89–104.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.

Safonova, A., Ghazaryan, G., Stiller, S., Main-Knorn, M., Nendel, C., Ryo, M., 2023. Ten deep learning techniques to address small data problems with remote sensing. Int. J. Appl. Earth Obs. Geoinf. 125, 103569.

Seber, G.A., 2009. Multivariate Observations. John Wiley & Sons.

Sheikhpour, R., Sarram, M.A., Gharaghani, S., Chahooki, M.A.Z., 2017. A survey on semi-supervised feature selection methods. Pattern Recognit. 64, 141–158.

Simons, D.G., Snellen, M., 2009. A Bayesian approach to seafloor classification using multi-beam echo-sounder backscatter data. Appl. Acoust. 70 (10), 1258–1268.

Snellen, M., Siemes, K., Simons, D.G., 2011. Model-based sediment classification using single-beam echosounder signals. J. Acoust. Soc. Am. 129 (5), 2878–2888.

Snellen, M., Simons, D.G., Riethmueller, R., 2008. High frequency scattering measurements for mussel bed characterisation. J. Acoust. Soc. Am. 123 (5), 3627.

Teixeira, J.B., Martins, A.S., Pinheiro, H.T., Secchin, N.A., de Moura, R.L., Bastos, A.C., 2013. Traditional ecological knowledge and the mapping of benthic marine habitats. J. Environ. Manag. 115, 241–250.

Van Walree, P.A., Tęgowski, J., Laban, C., Simons, D.G., 2005. Acoustic seafloor discrimination with echo shape parameters: A comparison with the ground truth. Cont. Shelf Res. 25 (18), 2273–2293.

Zhou, D., Bousquet, O., Lal, T., Weston, J., Schölkopf, B., 2003. Learning with local and global consistency. In: Advances in Neural Information Processing Systems. Vol. 16, MIT Press.