Presence of Sabellaria spinulosa reefs in the northern Brown Bank, Dutch Continental Shelf, North Sea

[MONS ID56: monitoring reef-building species]

ONGERUBRICEERD Releasable to the public ⟩ TNO 2024 R11012
December 2024

TNO innovation for life

Defence, Safety & Security
www.tno.nl
+31 88 866 10 00
info@tno.nl

TNO 2024 R11012 – December 2024

Presence of Sabellaria spinulosa reefs in the northern Brown Bank, Dutch Continental Shelf, North Sea

[MONS ID56: monitoring reef-building species]

Author(s)	Timo Constantin Gaida, Bas Binnerts, Joël Cuperus, Oscar Bos, Helga van der Jagt
Classification report	ONGERUBRICEERD Releasable to the public
Title	ONGERUBRICEERD Releasable to the public
Report text	ONGERUBRICEERD Releasable to the public
Appendices	ONGERUBRICEERD Releasable to the public
Number of pages	85 (excl. front and back cover)
Number of appendices	4
Sponsor	Rijkswaterstaat
Project name	RWS - MONS - Biogenic Reef Monitoring
Project number	060.57605

Our partners

WAARDENBURG Ecology | eurofins AquaSense | WAGENINGEN UNIVERSITY & RESEARCH

Dit onderzoek is uitgevoerd in opdracht van het MONS-programma. MONS vormt de kennisbasis voor het Noordzeeoverleg

NOORDZEEOVERLEG

Medegefinancierd door de Europese Unie

Samenvatting

Sabellaria spinulosa werd gedetecteerd en in kaart gebracht in delen van de noordelijke Bruine Bank, Nederland. De toepassing van de Side-scan sonar (SSS) is succesvol gebleken bij het identificeren van verhoogde *Sabellaria* riffen op de zeebodem. In de videobeelden, verkregen met een op afstand bedienbaar voertuig (ROV), werden individuele tot dichtbevolkte korsten (< 5 cm hoogte) en verhoogde (5 tot 20 cm hoogte) Sabellaria-riffen gedetecteerd. Er werd een kunstmatige intelligentie (AI) classificatie workflow ontwikkeld om verhoogde Sabellaria-riffen semiautomatisch in kaart te brengen met behulp van de verkregen SSS-dataset en ROV-video-opnamen. Het ontwikkelde classificatiemodel, gebaseerd op de support vector machine *classifier*, bracht de verhoogde *Sabellaria* riffen in kaart met een precisie en gevoeligheid van respectievelijk 51,8% en 49,1% procent. Deze aanpak resulteerde in geschatte rifdekkingspercentages tussen 3,2% en 4,6% voor de verschillende beschouwde gebieden.

De kaarten met de aanwezigheid van riffen laten grootschalige ruimtelijke patronen van verhoogde Sabellaria-riffen zien met een voorkeur van *Sabellaria* voor vestiging ten oosten van het diepste deel van de *swale* tussen de getijdenruggen. De uitgevoerde statistische analyse kon geen duidelijke voorkeur van hooggelegen *Sabellaria* voor bepaalde sedimenteigenschappen of zeebodemmorfologie kwantificeren. Een visuele inspectie van de hoge-resolutie (10 x 10 cm) SSS-classificatiekaarten, bevestigd door de video-opnamen van de ROV, wees echter op de voorkeur van verheven *Sabellaria* voor vestiging in de troggen aan de kant van de meer fijnmazige megaripples.

Over het algemeen was de multi-schaal en multi-sensor benadering voor het detecteren en in kaart brengen van *Sabellaria* succesvol. Deze aanpak kan worden uitgebreid naar andere milieugebieden in de Noordzee en de lessen die uit dit project zijn geleerd, kunnen worden meegenomen in volgende campagnes om de kartering van Sabellaria-riffen verder te optimaliseren.

De biodiversiteitsanalyse, uitgevoerd op de verworven boxcore monsters, toonde een significant hogere biodiversiteit in *Sabellaria* monsters vergeleken met niet-*Sabellaria* monsters. Alle geanalyseerde boxcore monsters bevatten gemiddeld meer soorten en hogere abundanties dan de boxcore monsters die werden verkregen en geanalyseerd tijdens het MWTL-programma (Monitoring van de Waterstaatkundige Toestand des Lands) in 2021 [1]. Over het algemeen stond de Bruine Bank tot deze studie bekend als een gebied met een relatief arme soortengemeenschap [1]. De analyse die in dit project is uitgevoerd, heeft aangetoond dat in het diepste deel van de slenk tussen de getijdenruggen een unieke soortengemeenschap voorkomt, die mogelijk het soortenrijkste macrofauna-gebied van Nederland is.

Er werden drie verschillende soorten riffen van *Sabellaria* gevonden in de boxcore monsters: losse klompjes *Sabellaria* (type A), rifstructuren met verhoogde *Sabellaria* (type B) en rifstructuren van *Sabellaria* in het sediment (type C). Type A zijn ofwel jonge exemplaren die in de toekomst riffen kunnen ontwikkelen of waren klompjes van vernielde voormalige riffen. Type B zijn goed ontwikkelde verheven riffen en werden ingedeeld in een *reefiness* score van lage, gemiddelde en hoge bedekking. Type B werd in kaart gebracht met de SSS en de AI-classificator. Type C zijn rifstructuren van *Sabellaria* die niet verheven en deels begraven

zijn. Monsters van type C zijn het soortenrijkst en gemiddeld zijn de soorten zeer talrijk. Deze resultaten tonen aan dat Sabellaria-riffen een uniek ecosysteem met een hoge ecologische waarde in stand houden.

Summary

Sabellaria spinulosa was detected and mapped in parts of the northern Brown Bank, Netherlands. The application of the Side-scan sonar (SSS) has been successful to identify elevated *Sabellaria* reefs on the seabed. In the video footages, acquired with a *remote* operating vehicle (ROV), individual to densely populated crusts (< 5 cm elevation) and elevated (5 to 20 cm elevation) *Sabellaria spinulosa* reefs were detected. An artificial intelligence (AI) classification workflow was developed to semi-automatically map elevated *Sabellaria* reefs employing the acquired SSS dataset and ROV video recordings. The developed classification model, based on the support vector machine (SVM) classifier, mapped the elevated *Sabellaria* reefs with a precision (true positive prediction rate) and sensitivity (true positive rate) of 51.8% and 49.1% percent, respectively. This approach resulted in estimated reef coverage percentages between 3.2 and 4.6 % for the different areas, which were located in the swale between the tidal ridges of the Brown Bank.

The reef presence maps show large-scale spatial patterns of elevated *Sabellaria* reefs, indicating a preference of *Sabellaria* for settlement to the east of the deepest part of the swale between the tidal ridges. The performed statistical analysis could not quantify a clear preference of elevated *Sabellaria* towards certain sediment properties or a seabed morphology. However, a visual inspection of the high-resolution (10 x 10 cm) SSS classification maps, confirmed by the ROV video recordings, indicated the preference of elevated *Sabellaria* for settlement in the troughs towards the stoss side of the finer-scale megaripples.

Low positioning accuracy of the SSS dataset hampered (1) the generation of the final elevated *Sabellaria* reef presence maps with a higher resolution than the obtained 100 x 100 m, (2) the fine-scale comparison (meter scale) with the multibeam echosounder (MBES) data for the statistical habitat analysis and (3) the integration of the MBES data into the classification workflow due to the positional mismatch between the SSS and highly accurate MBES data.

The biodiversity analysis, conducted on the acquired boxcore samples, has shown a significantly higher biodiversity in *Sabellaria* samples compared to non-*Sabellaria* samples. All analysed boxcore samples contained on average more species and higher abundances than in boxcore samples acquired and anyalsed during the MWLT (Monitoring van de Waterstaatkundige Toestand des Lands) programm in 2021 [1]. In general, until this study the Brown Bank was known as an area with a relatively poor species community [1]. The analysis carried out in this project has shown that a unique species community occurs in deepest part of the swale between the tidal ridges, which is possibly the most species-rich macrofauna area in the Netherlands.

Three different types of *Sabellaria* reefs were found in the box core samples: loose *Sabellaria spinulosa* clumps (type A), reef-structures with elevated *Sabellaria spinulosa* (type B) and reef-structures of *Sabellaria spinulosa* in the sediment (type C). Type A are either juveniles and may develope reefs in the future or were clumps of destroyed former reefs. Type B are well developed elevated reefs and were classified into a reefiness score of low, medium and high coverage. Type B was mapped with the SSS and the AI classifier. Type C are *Sabellaria spinulosa* reef-structures which are not elevated and partly buried. Type C samples are the

most species-rich and on average the species are very abundant. These results show that *Sabellaria* reefs sustain a unique ecosystem with high environmental value.

In summary, the multi-scale and multi-sensor approach for detecting and mapping *Sabellaria spinulosa* has been successful. This approach can be extended to other environmental areas in the North Sea and the lessons learned from this project can be incorporated into next campaigns to further optimize the mapping of *Sabellaria* reefs.

Contents

1 Introduction

Protecting and enhancing biogenic reef forming species such as the Ross worm, *Sabellaria spinulosa,* is considered of key importance for North Sea nature restoration. As part of the MONS monitoring program [2], TNO (coordinator) was asked together with Waardenburg Ecology, Eurofins Aquasense and Wageningen Marine Research to investigate the presence of *Sabellaria spinulosa* reefs in the Brown Bank (or Brown Ridge or Bruine Bank (in Dutch)) area, in the Dutch part of the North Sea (MONS ID56, monitoring reef-building species) (Figure 1.1). The interest for this species has to do with its ability to create biogenic reefs that stabilize the seabed and potentially increase biodiversity by providing a habitat for a multitude of other species. Under OSPAR and within the Dutch Marine Strategy, these reefs are recognized to be under threat and in need of protection [2] [3]. The results of this study will help answering the questions (i) to what extent *Sabellaria* reefs are present in the northern Brown Bank area, (ii) to what extent the rest of the Brown Bank area is suitable for *Sabellaria* reef establishment and (iii) what the biodiversity of *Sabellaria reefs* is in this region. Insights from this study will help the Dutch government to develop more effective policy regarding the protection of the North Sea ecosystem.

The cruise report describes the measurements campaign in detail [4]. In summary: during the campaign in August 2023, Multibeam echosounder (MBES) and Side Scan Sonar (SSS) data were acquired in the Areas A, B, E, F and G depicted in Figure 1.2. Area C and D, as shown in Figure 1.1, were not surveyed because of limited survey time and observation made during the trial. Boxcore samples were acquired at selected locations of interest within these areas. In addition, ROV (remotely operated vehicle) video recordings were acquired along transects of interest. In the cruise report, the data acquisition strategy, data coverage and quality of the acquired data are described [4]. In the data factsheet, a description of the public data made available via the Marine Information and Data Centre (IHM) is given [5].

This final project report provides an estimate of the *Sabellaria* reef presence in the Brown Bank obtained from the application of an artificial intelligence (AI) classification method to Side-Scan Sonar data (SSS) and ROV video recordings. The AI classification workflow was developed by TNO within this project and the SSS and ROV data were acquired during a two-week monitoring campaign in August 2023 by RWS, Waardenburg Ecology, Eurofins Aquasense and TNO. The main interest was to find relatively large patches (>~1m^2) of *Sabellaria* exposed on the seabed surface, and hence less on finding buried or smaller reefs (<~1m^2). Furthermore, this report contains an analysis of the preferences of the *Sabellaria* reefs for settlement using the ROV video recordings and the generated maps (i.e. Sabellaria presence map, sediment and bathymetric map). In addition, it contains a detailed analysis of the biodiversity content using the acquired grab samples.

The analysed ROV video data is used to select the required training and testing data in the SSS dataset. The acquired MBES data is used to (i) improve the georeferencing of the SSS data and (ii) to cross correlate the predicted *Sabellaria* hotspot with the acoustic backscatter (as a proxy for the sediment composition). The boxcore samples are used to (i) quantify the biodiversity of *Sabellaria* reefs, (ii) understand the preferred settlement conditions of *Sabellaria* reefs and (iii) correlate the MBES backscatter to the sediment composition.

In the remainder of this introduction, background knowledge on *Sabellaria* is described. Next, the data processing methodology and *Sabellaria* presence maps are described in Chapter 2. In Chapter 3 and 4 the preference of *Sabellaria* for settlement is specified and its biodiversity content is analysed. The results are discussed in Chapter 5. Finally, conclusions and lessons learned are described in Chapter 6.

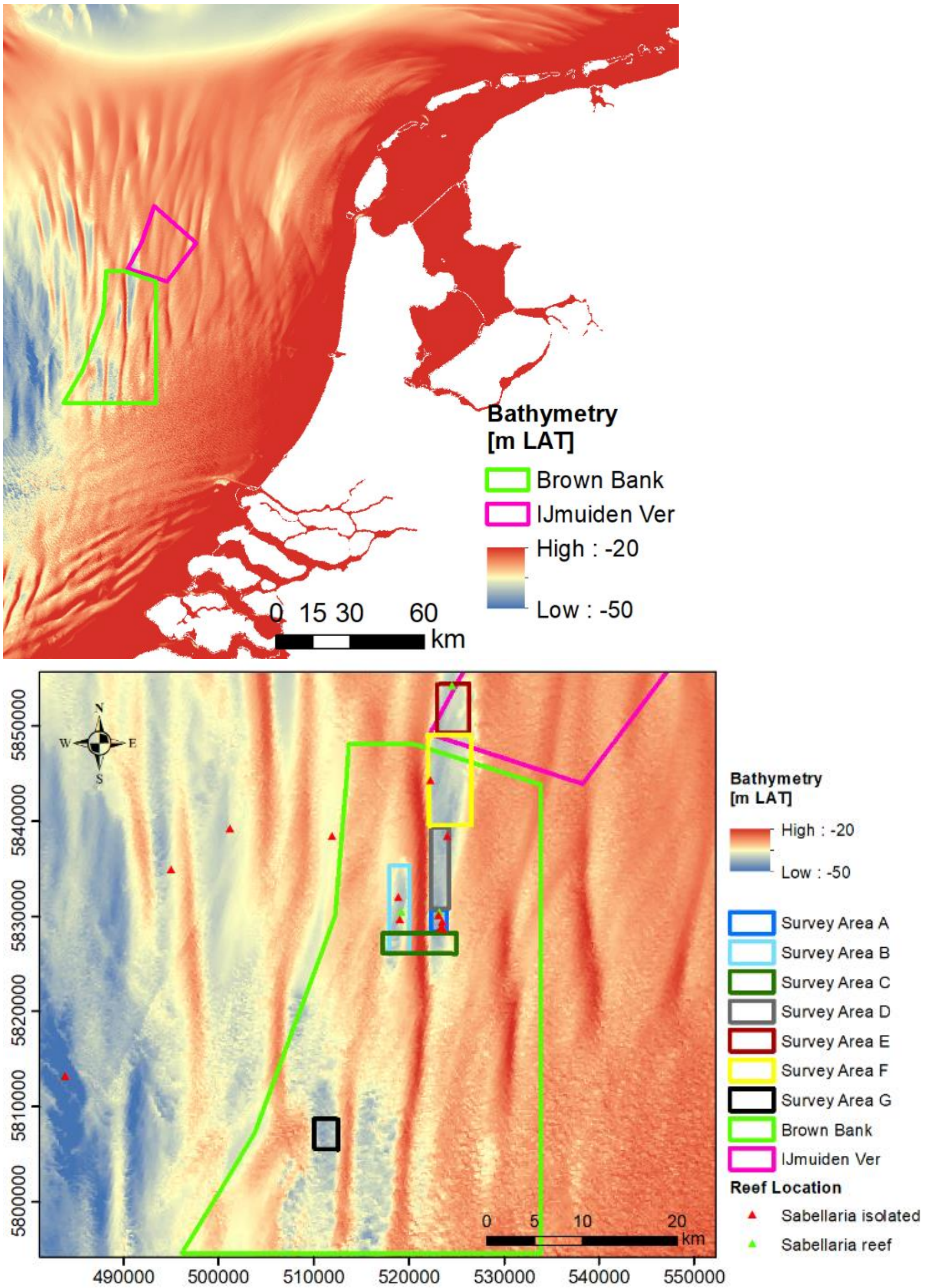


Figure 1.1: The top map displays the geographical location of the Brown Bank and IJmuiden Ver area. The bottom map shows the preliminary planned survey areas (bounding boxes) and *Sabellaria* reef locations found in previous studies (markers).

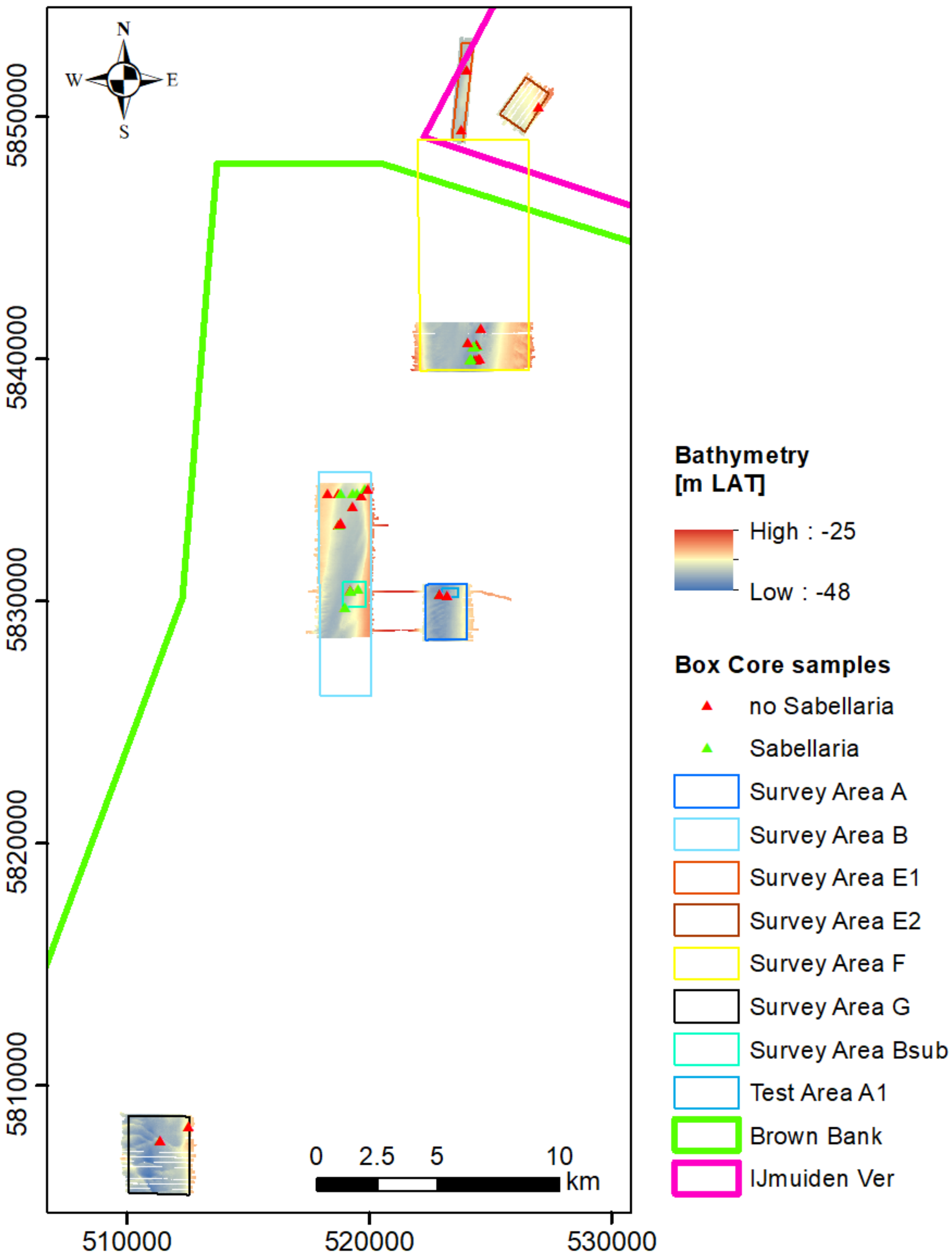


Figure 1.2: Overview map of the northern Brown Bank showing the processed bathymetric data and boxcore sample locations indicating the presence of *Sabellaria* reefs. The Side Scan Sonar data overlaps with the bathymetric data.

1.1 *Sabellaria* spinulosa

The Ross worm (*Sabellaria spinulosa*) is a tube-forming polychaete worm that usually occurs individually or in small aggregations, but that can also form extensive biogenic reefs covering several hectares. To build tubes, *Sabellaria* spinulosa requires a continuous supply of sand grains that are put into suspension by strong water movement [6]. Individual worms can survive 2 - 5 years [7]. Spawning occurs in winter (January-March) with a larval stage of approximately 1-2 months and a dominant settlement period in March [7]. Settlement requires a hard substrate. Individual aggregations may frequently form and collapse, the reefs however can persist in a region for many years ([8] [7] [3]).

Figure 1.3: *Sabellaria* spinulosa reefs at the Brown Bank in 2017 (photo: Oceana).

1.1.1 Protection and restoration

Sabellaria reefs are important biogenic habitats that are in need of protection [9]. The reef subtype found on mixed substrata with sand, granules, pebbles, and cobbles (including biogenic particles such as shells and shell fragments) is defined as covering 30% or more of the area and supporting a distinct epibiota community [9]. In the Netherlands, *Sabellaria spinulosa* is included as a typical species for habitat type H1110 under the Habitat Directive. And under the Dutch Marine Strategy targets are set to restore biogenic reef-building species such as *Sabellaria* [10].

To be able to protect and restore *Sabellaria*, more knowledge is needed on the current distribution and on the conditions that are required for restoration. The research described in this report is instrumental for that. As a first step, habitat suitability maps [11] under the MONS programme and data on *Sabellaria* occurrence have been put together [12]. Next, a background document for a protection plan has been developed for reef building species, such as *Sabellaria* and the European flat oyster [13]. This document includes suggestions for further research and protection measures. Based on this information, new offshore wind farms such as ECOWENDE [14] will take measures to stimulate *Sabellaria* reefs near the Brown Bank area. Further possibilities for *Sabellaria* restoration are taken into consideration under the Dutch North Sea nature restoration programme (Programma Natuurherstel Noordzee).

1.1.2 Distribution in the Dutch North Sea

Habitat suitability of the seabed shows that the Dutch North Sea does not host large suitable areas for reef building (Herman & Van Rees, 2022). In the Brown Bank area however, the abiotic conditions are suitable for *Sabellaria*. In 2017 and 2019 small reefs were found in this area ([15] [16]) (Figure 1.3). In 2021, also small aggregations of *Sabellaria* clumbs on dead *Spisula subtruncata* shells have been found in the Frisian Front area [17] (Figure 1.4). In addition, Sabellaria spinulosa reefs have been observed on Het Friese Front in the WOT surveys (personal observations by Joël Cuperus).
Furthermore, small colonies are found on artificial structures [12], which enhance local scouring that increases bottom shear stress and makes mobile sand available for the worms to build reef structures [11]. In June 2023, the entire Dutch coast was surveyed using the Plankton Imager and, among other things, Sabellaria Spinulosa larvae were caught. Most larvae were found between Scheveningen and Bergen (North Holland) [18].

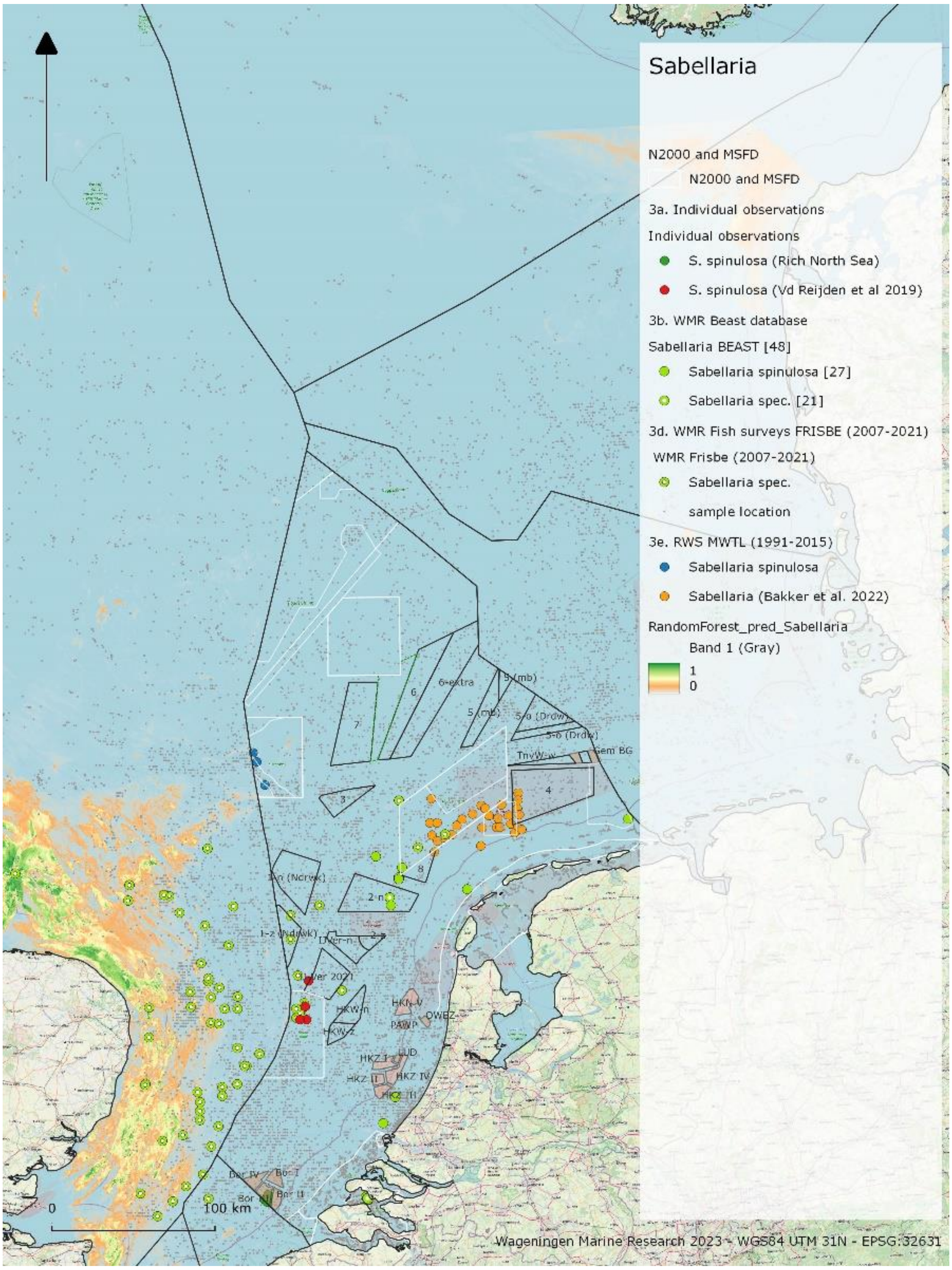


Figure 1.4: *Sabellaria spinulosa* and *Sabellaria* sp. Predicted probability of occurrence of S. spinulosa ('suitability'; on the basis of random forest regression (Herman & Van Rees 2022) and point data of *Sabellaria spinulosa* and *Sabellaria* sp (Bos et al., 2024).

1.1.3 Habitat definitions

EUNIS

In the European system of nature information (EUNIS), two types of *Sabellaria* reefs are described: A5.611 '*Sabellaria spinulosa* on stable circalittoral mixed sediment' and *S. spinulosa* reefs on rock (EUNIS code: A4.22). (https://eunis.eea.europa.eu/habitats/1693).

OSPAR

Under OSPAR [9], the habitat is defined as:
"The tube-building polychaete *Sabellaria spinulosa* can form dense aggregations on mixed substrata and on rocky habitats. In mixed substrata habitats, comprised variously of sand, gravel, pebble and cobble, the *Sabellaria* covers 30% or more of the substrata and needs to be sufficiently thick and persistent to support an associated epibiota community which is distinct from surrounding habitats. On rocky habitats of bedrock, boulder and cobble, the *Sabellaria* covers 50% or more of the rock and may form a crust or be thicker in structure. In some areas, these two variations of reef type may grade into each other. *Sabellaria* reefs have been recorded in depths between 10-50 m Below Chart Datum (BCD) or more. The reef infauna typically comprises polychaete species such as *Protodorvillea kefersteini, Scoloplos armiger, Harmothoe* spp., *Mediomastus fragilis, Lanice conchilega* and cirratulids together with the bivalves *Abra alba* and *Nucula* spp. and tube-building amphipods such as *Ampelisca* spp. Epifauna comprise calcareous tubeworms, pycnogonids, hermit crabs, amphipods, hydroids, bryozoans, sponges and ascidians. *S. spinulosa* reefs are often found in areas with quite high levels of natural sediment disturbance; in some areas of reef, individual clumps of *Sabellaria* may periodically break down and rebuild following storm events. *S. spinulosa* reefs have been recorded from all European coasts except the Baltic Sea, Skagerrak and Kattegat. Areas of dead *Sabellaria* reef indicate the site supported reef habitat in the past and should be reported as this habitat type".

Reefiness

Criteria to assess the reefiness of *Sabellaria* are described by Gubbay (2007) [19] and Hendrick and Foster-Smith (2006) [20]. These criteria assess the reefiness on the basis of a.o. the number of individuals/m2 and the reef elevation above the seafloor. A reliable number of individuals/m^2 cannot be derived from acoustic measurements.

1.1.4 Habitat preferences

The species has a strong preference for turbid waters with a good supply of sand and prefers hard substrates such as bedrock, cobbles and boulders as well as shells or tubes or sites previously used by the species. Furthermore, it occurs in areas with moderately strong to strong (0.5-3 m/s) currents and areas that are (very) exposed to waves [7]. The Brown Bank is characterized by moderately strong currents and therefore fits to this criteria. For the Brown Bank area, Van der Reijden et al. (2019, 2021) ([15] [16]) report that *Sabellaria* are mainly found on the sandy seafloor in the troughs of megaripples (ripples with wavelengths of ~ 10 m), which are superimposed on the sand waves (with lengths of ~ 200 m).

1.1.5 Ecological function and biodiversity

Sabellaria reefs provide shelter and substrate for other species due to their 3D structure. For the Brown Bank area, a higher number of epifauna species were found, compared to adjacent areas. Also species known for their association with hard substrate were observed, such as the *Pholis gunnellus*, *Cancer pagurus* (Figure 1.3) and the *Necora puber*.

Furthermore, the reefs hosted *Pisidia longicornis* that serve as food for associated fish species, and *Scyliorhinus canicula* were observed ([15] [16]).

1.1.6 Threats and impacts

The species is considered to have an r-strategy (a life strategy with high reproduction to take optimal advantage of food and space). This means that recoverability of *S. spinulosa* is considered high due to fast recolonization potential. This is also necessary since the reefs occur at locations where disturbances take place regularly (e.g. storms and pollution) [3]. Physical damage appears to potentially affect the reefs the most, with damage coming from activities such as dredging, trawling, net fishing, potting, and the installation of infrastructure [3].

2 *Sabellaria* presence maps

2.1 Introduction

The main project output is to deliver maps indicating the presence of *Sabellaria* reefs in the Brown Bank. Therefore, a workflow was developed to obtain *Sabellaria* reef maps from raw side scan sonar (SSS) data with the support of ROV video recordings. This workflow is applicable to any SSS sonar survey which includes the acquisition of ROV video recordings or similar ground truth information. However, the trained model was optimized to this specific dataset and the Brown Bank area (see discussion in Chapter 5).
In a first step the SSS data is processed to generate a georeferenced SSS image. The processed ROV video recordings are then used to select training and testing data to train the AI classification algorithm. The trained AI classifier model is evaluated with the selected testing dataset. The trained model is applied to the processed and georeferenced SSS image to classify the data into presence and absence of elevated *Sabellaria* per survey line. In a final step the classified SSS images are merged into a full-coverage reef presence map, indicating the percentage of reef coverage on the seabed (Figure 2.1). Since the model is trained and validated on area A, B, F and G, the model application is currently restricted to these areas.

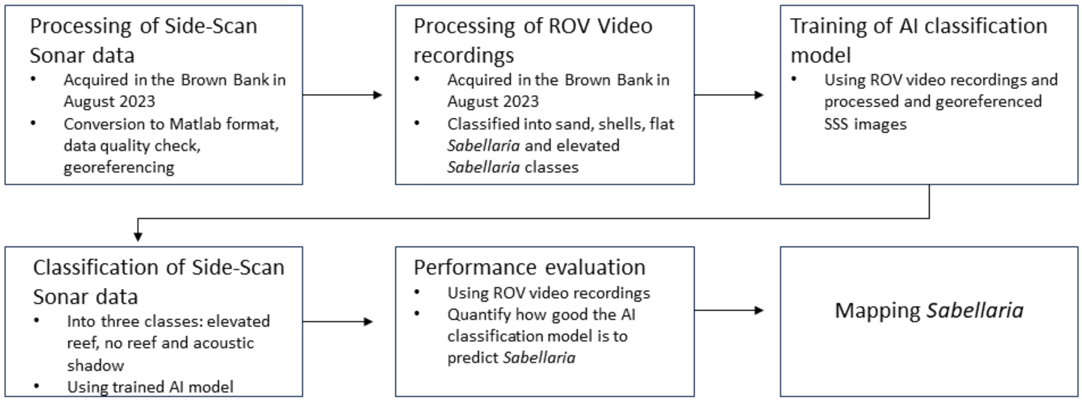


Figure 2.1: Schematic representation of mapping *Sabellaria spinulosa* reefs based on ground truth and Side Scan Sonar data.

2.2 Methodology

2.2.1 Data processing pipeline

Five data processing steps are needed to create maps of the *Sabellaria* coverage. These steps are described below. Steps 2 and 4 are described in more detail in Sections 3.2 and 3.3.

Step 1: loading and pre-processing Side Scan Sonar data:

- In **step 1.1**, the raw sonar data (.sdf) is loaded[1] and converted to a MATLAB format (.mat).
- In **step 1.2** the data is detrended along the range dimension to reduce the effect of variable propagation conditions on the signal. The 75th percentile of the amplitude range curve (also called amplitude time series) is computed over 3000 pings . This curve is then subtracted from each ping to detrend the data.
- In **step 1.3** poor quality data close to the nadir (water column and high incident angle reflections, 0 to 25 degrees) and at longer ranges (poor SNR) is removed (good data quality was observed up to 75 m range). Pings with high deviations from the main heading are removed as well. This is done to remove SSS transects where the ship is turning resulting in poor data quality. Even after these data removal steps the data quality still varies throughout the dataset due strong variation of the two fish altitude (5 to 12 m) caused by manual operation and crabbing of the tow fish caused by ocean currents.
- In **step 1.4** layback correction is applied to improve the position accuracy of the sonar for the data acquired in the first week as no ultra-short baseline (USBL) positioning systems was available. The correct location is important for (i) selecting the training and testing data and (ii) when fusing the different sonar lines into a combined map.
- In **step 1.5** the sonar time series is projected on a regular 10x10cm georeferenced grid in WGS84 UTM 31N coordinates. For projecting the sonar time series on the seabed, the slant range, tow fish flying height, the vessel heading and the tow fish position (from layback correction and USBL locator) were used. A processed and georeferenced SSS image is shown in Figure 2.2.

Step 2: selection of training and testing data

In step 2, training and testing data is selected manually in the areas where labelled ROV video data is available. A manual instead of automatic selection was done because of following reasons: (1) the SSS had a low positioning accuracy (primarily for week 1 when no USBL locator signal was available for the SSS) and therefore do not geographically match with the ROV tracks, (2) the megaripples caused acoustic shadows (blind zones), which shouldn't be used for training and (3) it allowed manual labelling of areas outside (but nearby) the ROV track to obtain enough training and testing data.

For the manual labelling the MBES data with a positioning accuracy of a few centimeters was used to manually correct for SSS data localisation errors (to align it with ROV tracks). The manual labelling was done in three categories being (i) elevated reefs, (ii) no reef, (iii) shadow zones. In the video recordings the elevated *Sabellaria* was identified as individual clumps with an elevation of more than 5 cm up to a high seabed density of more than 50% with an elevation of 20 cm The same SSS track line was not used for training and testing the classifier to avoid overtraining of the classifier. Examples of selected training and testing data are shown in Figure B.1 in Appendix B. A more detailed description of step is provided in Section 2.2.3.

[1] Using a software tool (TNO background) based on https://ge0mlib.com/papers/File_Formats/Jsf_rev1_13.pdf

Step 3: generation of feature layers
In step 3 feature layers, fed into the machine learning algorithm (support vector machine), are generated using the pre-processed SSS lines resulting from step 1. Feature layers are created from the raw data (here the pre-processed SSS line) in order to create a more effective set of inputs for the machine learning algorithm. It helps the algorithm to better predict the classes (i.e. the *Sabellaria* reefs). The rationale behind using a feature layer approach (opposed to just feeding in raw data) was that (i) satisfactory results were obtained using the tested feature layers in literature [21], (ii) it allowed for selecting feature layers based on acoustic and morphological knowledges. For example, the settings for creating the different feature layers from the acoustic SSS image can be chosen in correspondence with the targeted environment like wavelength of megaripples.
A tile size of 1 x 1 m was selected to compute the input layers (hence containing 100 SSS samples per box given the 10x10cm resolution of the gridded data). The following feature layers are generated in step 3.1 to step 3.3.

- In step 3.1 basic feature layers are created being the ground range and mean + standard deviation (STDV) of the signal amplitude (in dB). The ground range is the distance between the tow fish nadir and the ground projected sonar data.
- In step 3.2 Gray level Co-occurrence Matrix (GLCM) metrics are computed (applied to the sonar amplitude image in dB). These metrics contain information about the texture in the image. Texture with different spatial information can be highlighted by computing the GLCM for different pixel distances (i.e. a pixel pair) in the image. The four features were created for pixel pairs of 1 and 5 pixels (i.e. 0.1 and 0.5 m distance between adjacent pixels, see Appendix A. Figure A.2 and Figure A.3). This way patterns with different spatial scales are highlighted.

  GLCM (Gray-Level Co-Occurrence Matrix[2]) metrics: Various statistical texture metrics that considers the spatial relationship. This involves the following four features
  a. **Contrast:** Measures the local variations in the GLCM.
  b. **Correlation:** Measures the joint probability occurrence of the specified pixel pairs.
  c. **Energy:** Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.
  d. **Homogeneity:** Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.
- In **step 3.3** a Gabor filter[3] is applied to obtain directional spatial wavelength information. This filter investigates a specific spatial frequency (i.e. wavelength) content in the image in specific directions in a localized region around the point or region of analysis. Wavelength of 0.5, 1.0 and 2.0 m towards the direction of 45 (SW-NE), 90 (W-E) and 135 (NW-SE) degrees were analysed.

In total 20 features layers were generated which serve as input for step 4. The 20 feature layers are displayed in Appendix A, in Figure A.2 to Figure A.4 and belong to the simulation run 2 described in Table A.1 in Appendix A.

Step 4: train and run supervised clustering
In step 4 the support vector machine (SVM) classifier[4] is trained (using training data), applied (to all data from area A, B, F and G) and validated (using testing data from area A, B, F and G). The locations of area A, B, F and G are shown in Figure 1.2. The SVM classifier is available

[2] https://nl.mathworks.com/help/images/ref/graycomatrix.html
[3] https://nl.mathworks.com/help/images/ref/gabor.html
[4] https://nl.mathworks.com/help/stats/support-vector-machine-classification.html

as a function in MATLAB. Different AI methods were tested in the process and the linear SVM method provided the best results in combination with the selected feature layers (see Appendix A). The evaluation of the performance is done in terms of the confusion matrix[5] generated from the test dataset. A detailed description on the evaluation of the performance of a classification model using the confusion matrix is provided in Section 2.2.4. For each SSS line the feature layers are computed and fed into the trained and chosen classification model. The output is a classified SSS line where each pixel is assigned to one of the pre-defined classes.

Step 5: Generate and analyse maps
In step 5.1, the classified SSS lines following from step 4 are clustered to generate maps of reef presence. Because of the uncertainty of the sonar data and the overlap of the lines; the data of different sonar tracks (at 10x10cm resolution) were aggregated (binned) onto a larger (100x100m) grid, counting the data entries ("pixels") labelled as reef or non-reef (discarding pixels with the acoustic shadow label). This can be used to create a percentage coverage metric for the larger grid. If an investigation of the classification results via a sensitivity study has shown that certain parts of the classified data are poorly classified, these areas can be excluded before the generation of the presence map. An example is the classified data from a ground range where the classifier is poorly trained or less sensitive. One reef presence map was created considering the full dataset and one more conservative map considering only data between a ground range from 30 to 75 m.
In step 5.2 a correlation analysis is done between the *Sabellaria* density derived from the acoustic data, video recordings and boxcore sample and the morphological conditions, with the aim to understand if certain environmental conditions (e.g. bathymetric features or acoustic backscatter) correlate with the presence of *Sabellaria*.

<u>Support vector machine</u>
The support vector machine is a supervised classification method which is a paradigm of machine learning or AI. A supervised classification method uses features layers or also called input data (here the acoustic image and its derivatives) and a pre-defined number of classes (here elevated reef, no reef and acoustic shadow) to train a model. The trained model allows to predict the classes for every location where the feature layers are available. The general concept of SVM is to maximize a hyperplane where the hyperplane is a separator between two classes. A hyperplane between two classes can be found by measuring the margin of the hyperplanes and finding its maximum. The margin is the distance between the hyperplane and the closets data from each class [22]. The closest data is called the support vector which are used to classify the data into the pre-defined classes. The hyperplanes are found by transforming the original feature layers into higher-dimensional space where the classes are easier to separate via a kernel function. The Kernel function can be linear or non-linear and can be chosen as a setting of the classifier. The SVM provides a posterior probability that an observation belongs to a particular class based on the input data. This probability allows to evaluate the classification of a particular area and to reject classifications where the trained model is uncertain if this location belongs to a class with a low probability.

2.2.2 Software implementation

The data processing and classification workflow is written in MATLAB code (development version R2022b). The modules developed within the scope of the MONS project are made available on a GIT repository [23]. Each module can be run individually. The data processing

[5] https://nl.mathworks.com/help/stats/confusionmat.html

pipeline comprises of the following consecutively run modules. An example test function that runs the subsequent module is made available on the GIT repository. The five previously described steps are integrated into the modules.

Module 1: This module loads the raw SSS data (.SDF) and produces pre-processed georeferenced SSS data stored as a .mat file. This module applies all sub steps of step 1.
Module 2: Within this module manual selection of suitable test and training data takes place. The data selection process was done using the MATLAB command line with an interactive interface. To enable reproducibility of the project results, the selected training and test data are stored on the git repository in .mat file format. This module contains step 2.

Note on module 1 and 2 IP: As the key functions deployed in module 1 and 2 are TNO background, the output of module 1 and 2 is made available such that the processing pipeline can be run without the need for these two modules. Commercially available software can also be used to generate similar pre-processed SSS mosaics as generated in module 1.

Module 3: This module computes all possible feature values for the training samples and stores the values in a look-up table. The following modules access the look-up table to extract values for the selected feature layers of the training data to train the classification model. This module contains step 3.

Module 4: Within this module the AI model is trained taking the training data and it feature values from the look-up table. The output of this step is the trained classifier. A version of the code is provided that reproduces the generation of the SVM classifier used to produce the final maps. This module contains steps 4.

Module 5: Within this module the trained AI model is tested using the test data picked in module 2. The trained AI model predicts the classes for the location of the test data based on the computed feature values. Using the predicted class and the true class obtained from the test data, the confusion matrix is computed to evaluate the trained model. This module contains steps 3 and 4.

Module 6: Within this module, the trained AI model is used to classify all SSS data. This module contains step 4.

Module 7: Within this module, the classified SSS data is combined to produce maps of modeled *Sabellaria* percentage coverage. The results are stored as .mat file and .geotiff file format.

2.2.3 Description of training and testing data selection process

The selection of the training and testing dataset for a supervised classification method is a crucial step in the workflow. The training dataset is used to train the classification model to recognize patterns in the input data which belong to a predefined class. The testing dataset should be as independent as possible from the training dataset to properly evaluate the performance of the classification model. Furthermore, the training and testing datasets should be representative for the targeted area.

The first step is to define the number and type of classes. The classes were selected based on a combination of the classified ROV video recordings and the TNO acoustic expert knowledge. The so-called hybrid approach was necessary for the following reasons:

- a varying positioning error of maximum 30 m in the SSS data, which was acquired in the first week of the survey, exists and therefore a geographical mismatch between the SSS and the more accurately positioned ROV video data (1 to 2 m positioning accuracy),
- the sonar data contains acoustics shadows, which could overlap with video recordings,
- allows for discarding of low quality data from the training and testing process (e.g. when fish were in the water column).

Three classes were selected: (1) elevated reef, (2) no reef and (3) acoustic shadow.

1. The elevated reef class comprises *Sabellaria* patches consisting of individual clumps to a seabed coverage of more than 50% with elevations higher than 5 cm (classified as eleavted *Sabellaria* in the video recordings and type B in the biodiversity analysis of the box core samples). The acquired SSS data did not provide enough resolution & contrast to identify *Sabellaria* with a lower elevation (classified as flat *Sabellaria* in the video recordings or type C in the biodiversity analysis).
2. The no reef class contains all areas which were classified in the video recordings as sand, sand with shells/gravel and flat reef.
3. The acoustic shadow is the blind zone of the sonar for that specific track. In the blind zone the acoustic signal is blocked by a morphological feature like a megaripple.

SSS data for which there was an overlap with the labelled video recordings was used for the selection of training and testing data. The MBES bathymetry with a positioning accuracy of a few centimetres was used to support the picking process as peaks of megaripples and other morphological features enabled correction for the inaccurate localisation data of the SSS data.

Training and testing data were selected at the position of the video recordings and further extended in case the pattern in the SSS image extends to a larger area around the video recording position (Figure 2.2). Some areas showing indicative acoustic patterns of elevated reefs but unconfirmed by video recordings were included as well. This had to be done because there was too little video data to generate sufficient training and testing samples. Two methods were deployed to select the samples (i) individually clicking the samples and (ii) drawing a polygon around an area. In case of (ii), the samples were randomly picked within the area bound by the polygon. To improve the independence of the training and testing data, SSS track lines are used only for either training or testing.

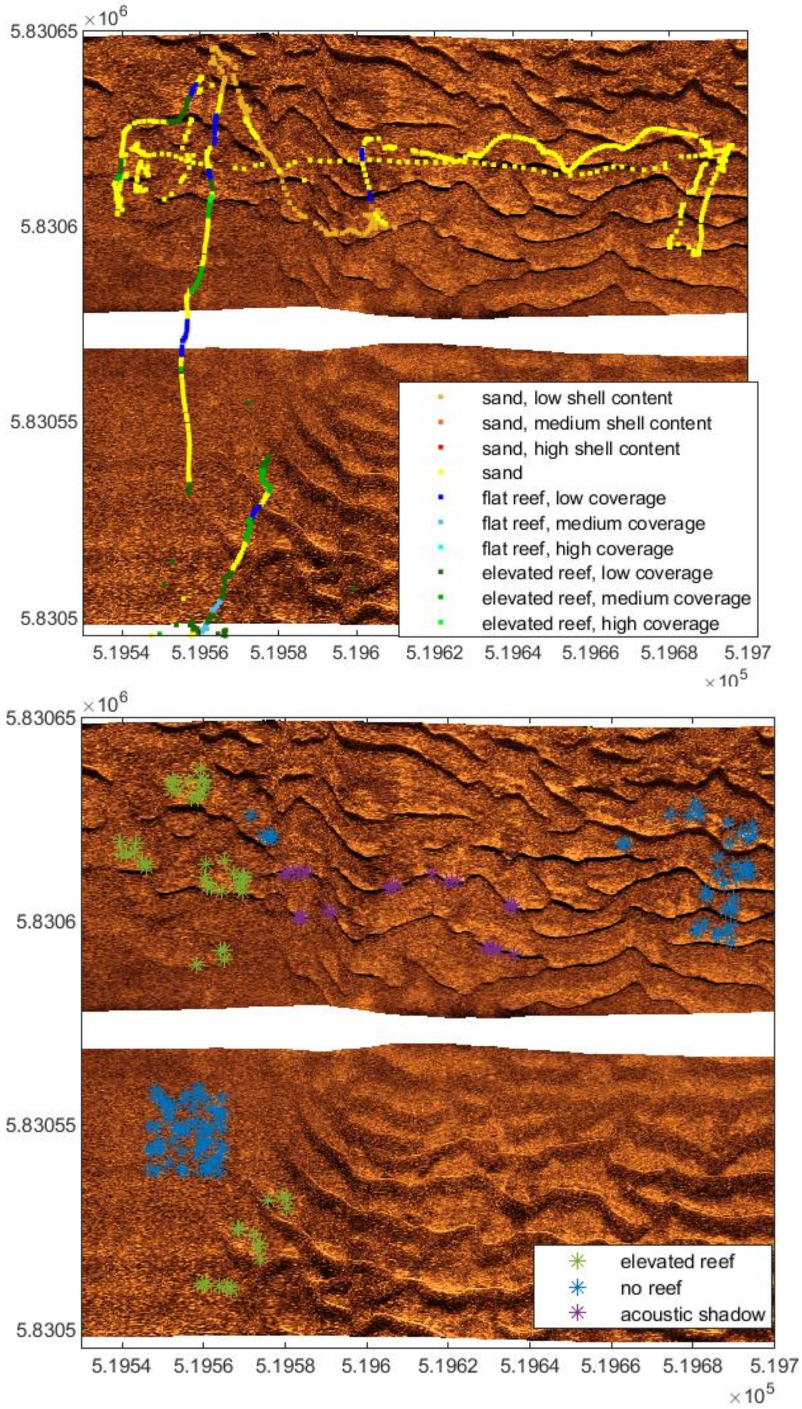


Figure 2.2: SSS image with overlying (top) classified video track and (bottom) picked training samples based on the hybrid approach that considers the geographical offset. SSS image is shown for track Bsub_W_2230831185800.sdf. Coordinates are in UTM31N.

2.2.4 Classifier performance evaluation

A variety of classifiers, such as Support Vector Machine (SVM), naïve Bayes, neural network, decision tree and k-nearest neighbour in combination with a variety of feature layer combinations were tested (Appendix A, Figure A.5 and Table A.1). In addition to the default

MATLAB settings, for the SVM and neural network different classifier settings (i.e. activation function, kernel function, connected layers) were tested as well (Appendix A, Figure A.6). The classifiers were trained on the training sample dataset with 3707 samples selected from area A, B, F and G. The evaluation of the performance of the classifiers was carried out on the test dataset with 2478 samples selected from the same areas A, B, F and G. Both the training and testing samples were selected following the procedure described in Section 2.2.3.

The performance analysis shows that the linear SVM classifier and a set of 20 feature layers provides a good trade-off between high computational speed and good performance. As described in the previous section, the set of feature layers provides the input for the linear SVM. Examples for the selected 20 feature layers are shown in Appendix A (Figure A.2 to Figure A.4). The evaluation of the performance of the classification is based on the confusion matrix. A schematic explanation of the confusion matrix is given in Figure 2.3. Figure 2.4 shows the actual confusion matrix obtained for the developed linear SVM model applied to the test dataset. The overall accuracy is 88.8% which means that only in 11.2% a false class is predicted. Since the main interest is in the performance of the classifier towards the detection of reefs, it is important to analyse the performance per individual class.

For the individual class predictions, a distinction can be made between:
- the true positive rate (Sensitivity) and false negative rate which summarize how well the true class is predicted correctly of falsely;
- the true positive predictive rate (Precision) explains how well the prediction corresponds to the true class.

The confusion matrix shows that elevated reefs are in 51.8% predicted correctly (Sensitivity) but in 48.2% falsely predicted as no reef or acoustic shadow. That means 48.2% of the actual elevated reefs are missed by the classifier. Furthermore, it is important to know how often the classifier predicts elevated reef but the true class is no reef or acoustic shadow (False positive predictive rate). In 50.9% of the test samples the classifier predicts elevated reef but the true class is no reef or acoustic shadow. That's means only 49.1% of the elevated reef class predictions are reliable (Precision - True positive predictive rate). This provides an uncertainty to the estimated reef coverage by the classifier. A more realistic estimate of the *Sabellaria* presence in the Brown Bank is obtained by correcting the *Sabellaria* coverage using the true positive predictive rate (i.e. multiplication of the estimated reef coverage with the precision) of 49.1% . In Figure B.1 (Appendix B) visual examples are shown for positive and false detection of the classifier obtained from the performance evaluation.

The presented scores for the confusion matrix belong to the trained linear SVM classifier trained and validated on samples from Area A, B, F and G. The majority of the training samples, in particular for the elevated reef class, are located in Area B. Testing the trained classifier only on samples of Area B increases the precision score from 49.1 % to 80.7% (see Figure A.7 in Appendix A). It shows that the trained classifier performance differs per surveyed area and it would be beneficial to acquire well-balanced training dataset.

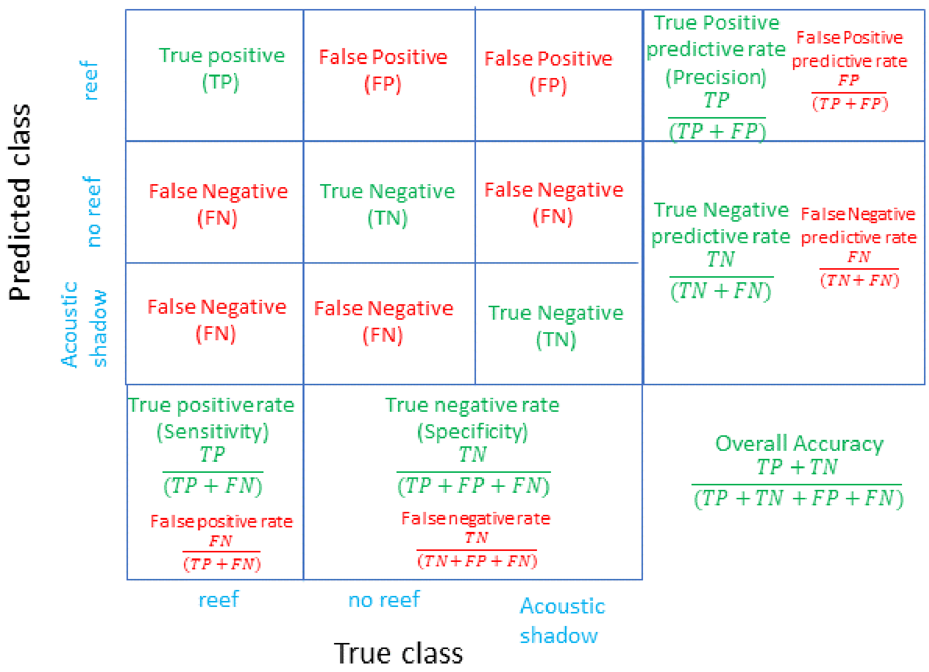


Figure 2.3: Overview of working principle of the confusion matrix. It supports the readability of the values presented in the confusion matrix obtained from the developed model in Figure 2.4.

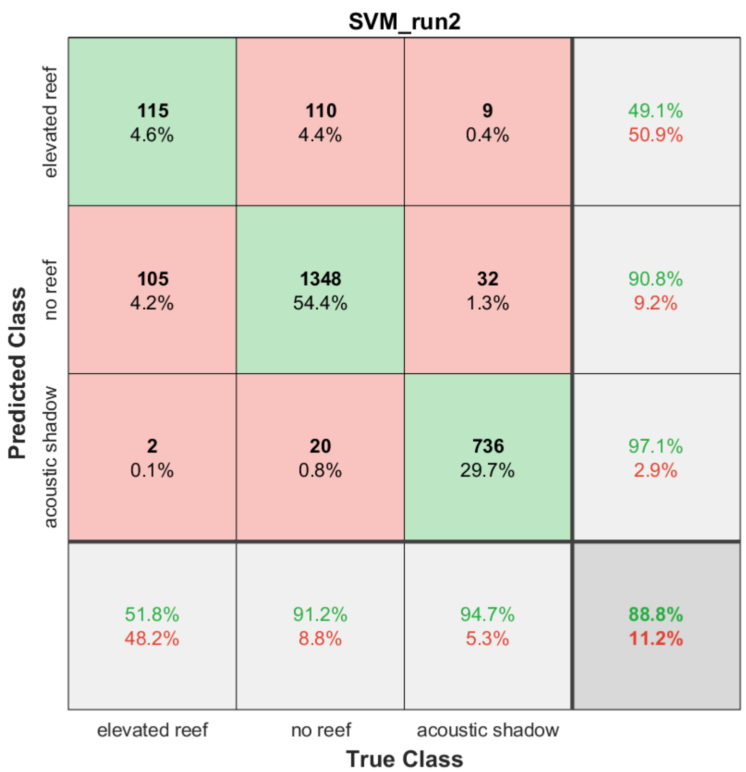


Figure 2.4: Confusion matrix obtained from testing the trained linear SVM classifier on test samples obtained from area A, B, F and G. The SVM was trained on the training samples of area A, B, F and G. The values were obtained via the equations presented in Figure 2.3 and an additional multiplication factor of 100 to obtain percentages.

2.2.5 Discussion on limitation of method

Before presenting the classification results and final *Sabellaria* reef presence maps, the limitations of the methods are discussed. While the availability of ROV video based ground truth data allowed for the selection of validated training and testing data and hence the computation of confidence metrics such as precision and sensitivity, the limited amount of validation data results in a remaining uncertainty that can only be described in qualitative terms. Key points to take into consideration are described below:

- Any AI classification model becomes increasingly inaccurate when applied to datasets deviating significantly from the training dataset. For SSS data this involves both changes in the environment (mainly morphology and sediment composition) and the sonar configurations (Mainly tow fish height, sonar frequency and pulse settings). Also the towing direction can have an effect as was discussed in the trial evaluation report.
- The application of AI classifiers to SSS for habitat mapping is a novel field of research. The complexity of the Brown Bank (large morphological variability) and variability of the towfish altitude complicate the application of AI methods as the environment and operating conditions influence the detectability of *Sabellaria*.
- While the predicted amount of *Sabellaria* is computed correcting for the precision using the test data, it is important to note that the test and training data is biased towards an area where significant *Sabellaria* reefs were found (Area B). To gain more confidence in the prediction made throughout the larger Brown Bank area more data of *Sabellaria* reefs used for training and testing data in other areas would be needed. The predictions for *Sabellaria* hotspots presented in this report could be used to direct future ROV video surveys in the area.
- In case of a need to improve the confidence in the results, it is recommended to acquire more ground truth data in the other areas where significant *Sabellaria* has been predicted.
- The megaripples present in the Brown Bank region challenge the development of a classifier because of the strong effect of the megaripples on sonar data. As described in the trial evaluation report, the *Sabellaria* was best visible in the acoustic data when data was acquired sailing parallel to the megaripples. To avoid introduction of uncertainty in the approach, only SSS data acquired when sailing parallel to the megaripples was analysed. This implied not running the classifier on the data acquired in Area E. Further validation work is needed to be able to assess the accuracy of the developed AI classifier when applied to data acquired when sailing perpendicular to the megaripples.
- The hybrid approach of selecting training and testing samples combines the video recordings with the expert interpretation of the acoustic image to extend the sample dataset from the video track line to larger areas around it when similar acoustic patterns were present. While the video recordings are an accurate representation of the seabed and selecting the samples on this track is precise, selecting the samples with increasing distance from the track line increases the likelihood of inaccurate labelling. It also results in an uncertainty for the performance analysis because not all samples in the testing dataset are fully covered by the ROV video track. This uncertainty is not considered in the confidence metrics.
- The *Sabellaria* reefs are visible in the SSS image via a blotchy pattern consisting of low and high amplitudes. This blotchy pattern is most likely caused by the higher reflectivity of reef patches facing the sonar and the low reflectivity of the acoustic shadow caused by the elevated reef structures on the opposite side. To obtain an acoustic shadow the reef must be elevated and the incident angle of the signal (i.e. angle between incoming sonar signal and the normal of the seabed) must be sufficiently high. An additional important observation is that the *Sabellaria* structure didn't cause a higher backscatter compared to the surrounding sediment. This is also confirmed by the MBES backscatter,

which didn't show a correlation between the acoustic backscatter and *Sabellaria* presence. These is most likely the reason why only elevated *Sabellaria* was detectable and why the classification model cannot be used to map flat *Sabellaria*. In addition, the SSS data and consequently the classification model did not allow to make a distinction between the degree of *Sabellaria* coverage density as defined by the video recordings to low (< 25%), medium (25 to 50%) and high (>50).

- At a late stage of the project testing different kernel functions of the SVM classifier has shown that a gaussian performs better than the employed linear kernel function improving the precision from 49.1 to 52.7 % and the Sensitivity from 51.8 to 62.6% for the elevated reef class (see Figure A.6 in Appendix A). It shows that small improvements can be obtained by optimizing the settings of the SVM classifier. However, this improvement was observed at a late stage of the project and therefore the maps were not reproduced with the updated settings. Also the Precision which is used in the estimation of the reef coverage was only slightly improved.
- The highest precision score for the reef class is obtained using the gaussian kernel function with 52.7% followed by the linear Kernel function with 49.1% and polynomial functions with ~45%. The sensitivity score is also highest for the gaussian function with 62.6% compared to 51.8% using the linear Kernel function. However, the option to select the gaussian function was discovered at a late stage of the project and the classification results were already obtained using the linear function. This result can be considered in the discussion that optimizing classifier settings can achieve an improvement in the classification accuracy.

2.3 Side Scan Sonar classification results

The linear SVM classifier was applied per sonar track line individually and thus the classification is not influenced by the positioning uncertainty of the SSS data (up to 30 m for the first week and up to 10 m for the second week). Merging the classified SSS track lines into a map was carried out afterwards in light of the positional uncertainty (see Section 3.5). Figure 2.5 shows four examples of classified SSS track lines acquired in areas A, B, F and G. The locations of the areas within the Brown Bank area are displayed in Figure 1.2. The example track lines for area G and F have only a small percentage of 1.1% (Area G) and 1.3% (Area F) of the seabed classified as elevated reefs, respectively. The example track line from Area A and B have a higher percentage of reef coverage with 7.3% and 17.1%, respectively. For the computation of the seabed reef coverage in percent, the acoustic shadow class is discarded.

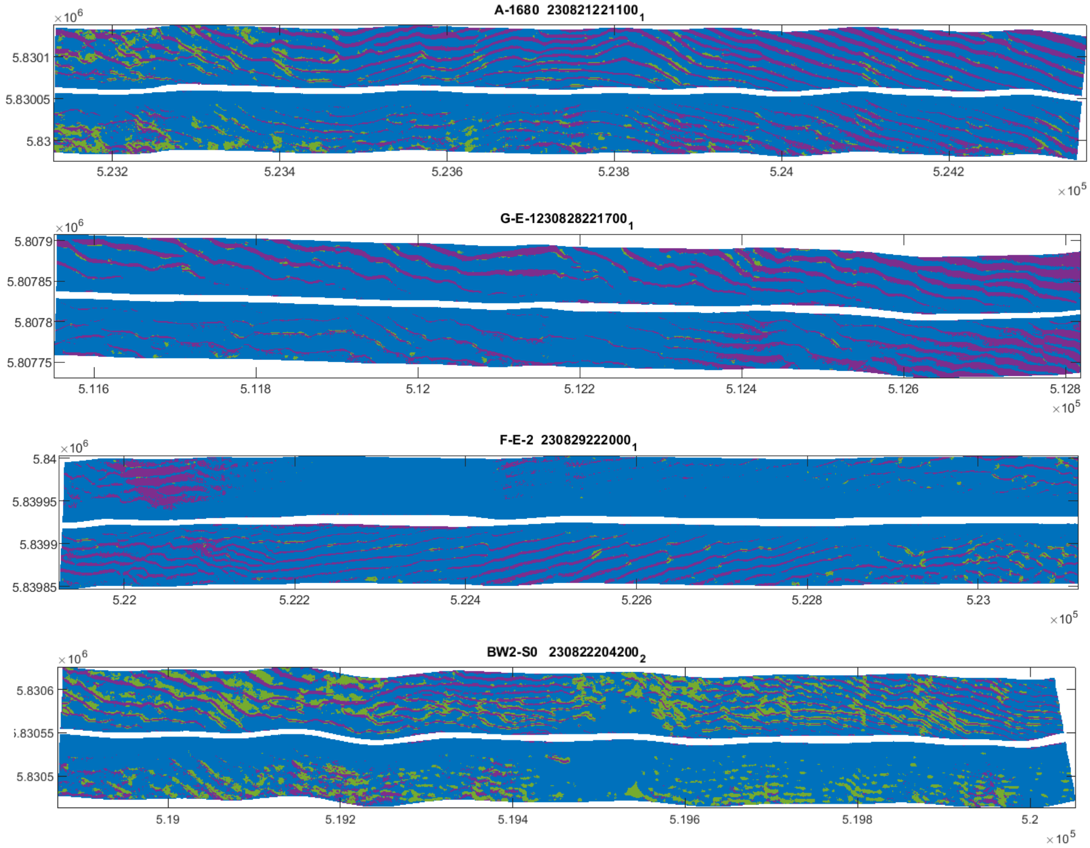


Figure 2.5: Classified SSS track lines with three classes: elevated reef (green), no reef (blue) and acoustic shadow (purple). (Top) SSS line from area A with moderate reef coverage of 7.3%, (middle top) SSS line from area G with low reef coverage of 1.1%, (middle bottom) SSS line from area F with low reef coverage of 1.3%, (bottom) SSS line from area B around a *Sabellaria* hotspot with high reef coverage of 17%. The coordinates are in WGS84 UTM 31N.

The classified SSS images, in particular the line from area B, shows visually that the elevated *Sabellaria* tends to be located on the northside of the megaripples. The deepest part of the trough is located in the blind zone of the sonar (acoustic shadow) but in general, the *Sabellaria* decreases towards the crest, being in alignment with the video recordings, which have shown higher *Sabellaria* presence in the trough and absence of *Sabellaria* on the crest. Furthermore, the figures show that close to the nadir gap (short ground range), there is consistently low *Sabellaria* detected, which is also confirmed by a sensitivity analyis on the effect of the ground range on the classification (ground range is defined as distance from nadir). The sensitivity analysis has shown that the reef coverage percentage increases to a ground range of 30 m before it stabilizes (see Appendix A, Figure A.8). This indicates that the classifier struggles to sucessfully detect *Sabellaria* in the region close to nadir. Most likely the low incident angle at this region does not allow to visulize the small-scale relief of the reefs. This intuitevely leads to fewer training samples being selected between 0 and 30 m ground range. Therefore, the classifier is less well trained at that range. A more detailed description is given in the Appendix A.

2.4 Reef-presence maps

The classified SSS lines are used to compute a reef percentage presence map for the surveyed areas considered in this memo (area A, B, F and G). In Table 2.1, the percentage of reef coverage is used to compute the coverage in km^2. The precision (true positive prediction rate) of 0.49 (49%) of the classifier is used to adjust the actual reef coverage as explained in Section 2.2.4. In the applied method the samples classified as acoustic shadows are discarded as it cannot be determined if these contain *Sabellaria* reefs. Two different reef percentage presence maps were generated: (1) considering the classified data from the full ground range between ~ 5 and 75 m and (2) considering only data between a ground range from 30 to 75 m. In the first map only the nadir filter removing the data from incident angles up to 25 degrees during the pre-processing in step 1 was applied. The second map takes into account that the classifier is less-well trained at shorter ground ranges and struggles to detect reefs in these regions (see Appendix A, Figure A.9).

Table 2.1: Estimated area of elevated *Sabellaria* per area using linear SVM classifier and data mapping procedure.

Area ID	Elevated Sabellaria area (using ground range ~5 to 75m)	Elevated Sabellaria area (using ground range 30 to 75m)
Area A (4.23 km2)	0.16 km^2 (3.8%)	0.24 km2 (5.7%)
Area B (16.14 km2)	0.57 km^2 (3.5%)	0.78 km2 (4.8%)
Area F (9.48 km2)	0.28 km^2 (2.9%)	0.39 km2 (4.1%)
Area G (9.76 km2)	0.27 km^2 (2.8%)	0.37 km2 (3.8%)

Figure 2.6 shows a histogram of the percentage *Sabellaria* coverage in the three studied areas for grid cells of 100 m x 100 m. Figure 2.8 to Figure 2.10 show the spatial distribution of *Sabellaria*. The colour bar shows the percentage coverage on a scale between 0 and 100%, where the dynamic range is clipped to 30% to improve the visualisation. The percentage of *Sabellaria* coverage for the different areas is higher when considering only data from a ground range between 30 to 75 m with values ranging between 3.8 to 5.7% (Table 2.1). The approach to remove the data between 0 and 30 m ground range is more aligned with the results from the sensitivity analysis (see Appendix A, Figure A.8) and can be considered as more robust. The histograms in Figure 2.6 show that area A and B have not only in average a higher reef seabed coverage but also more individual locations (100 x 100m) with a denser reef population than area G and F .
The areas A, B, F and G cover all from west to east the swale between the tidal ridges (see Figure 1.2). The highest *Sabellaria* percentage is mostly located to the east of the deepest location of the swale indicating a preference of *Sabellaria* for settlement. For comparison the bathymetric maps of Area A, B, F and G are shown in the Appendix D. The location and spatial orientation of the highest settlement density of *Sabellaria* in area B detected by the classifier, matches very well with the manual interpretation of the SSS image provided previously in the trial report [4] (see Figure D.1 in Appendix D).

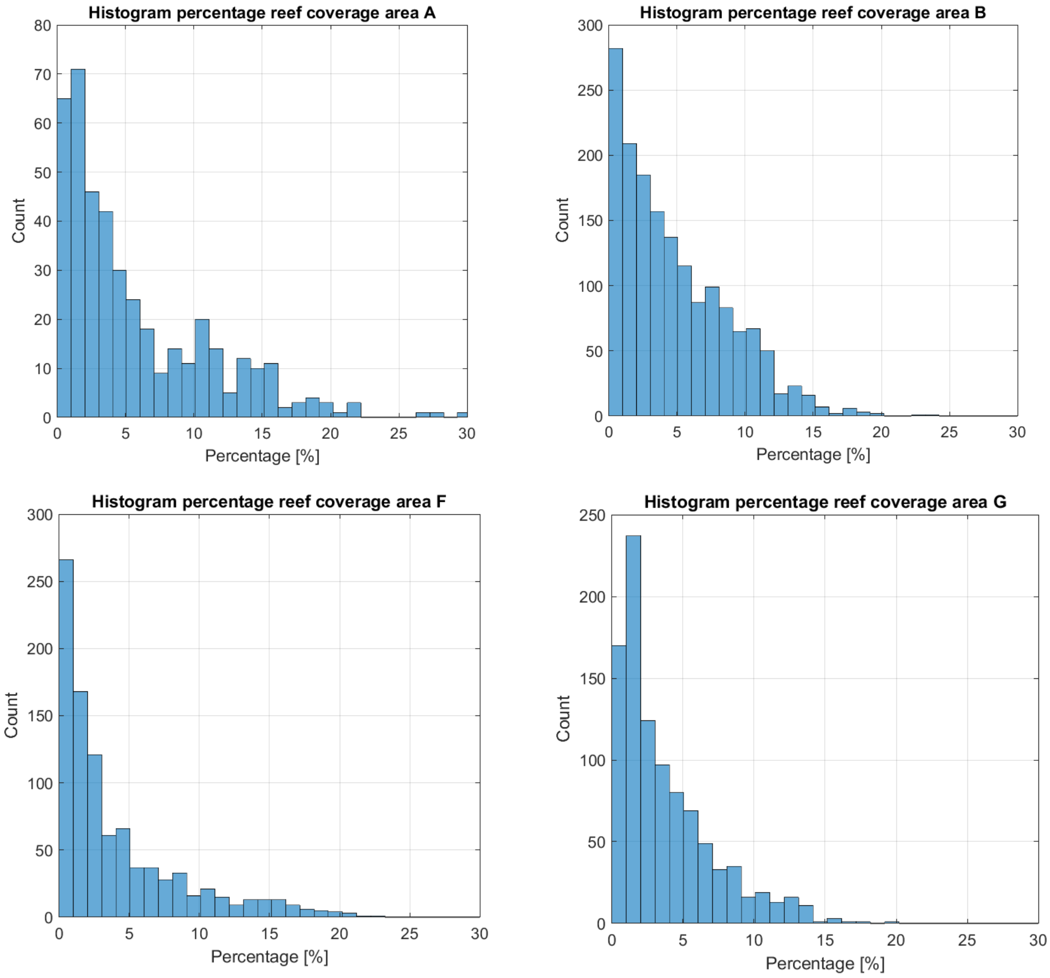


Figure 2.6: Distribution of the percentage *Sabellaria* coverage for area A, B, F and G. Each count represents a 100 x 100m grid cell. The data correponds to the aproach using the limited ground range between 30 and 75m.

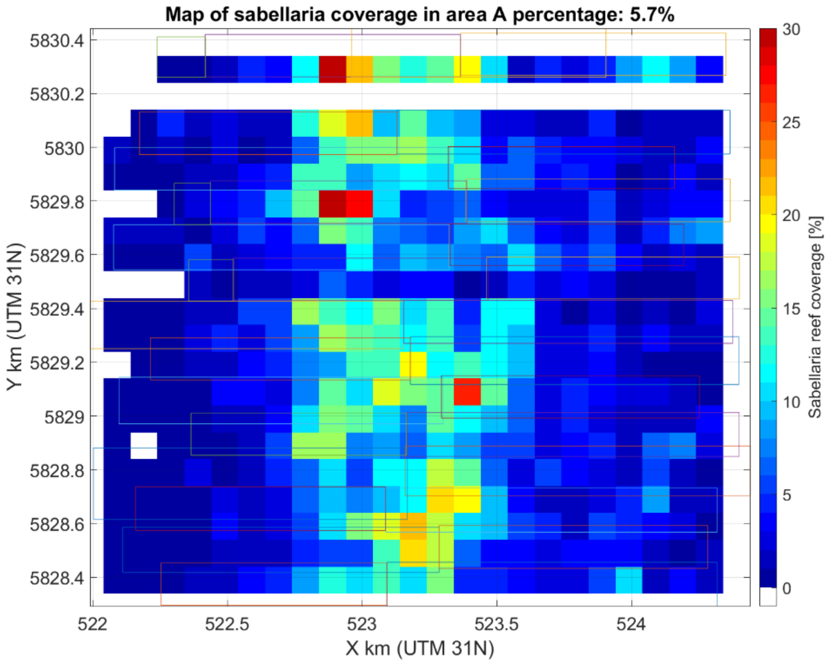


Figure 2.7: *Sabellaria* percentage presence map for area A. Total coverage of the seabed with elevated *Sabellaria* is 5.7%. The cell size is 100 x 100 m. Map is based on data from ground range between 30 and 75 m. The thins lines indicate the seabed coverage per SSS track line.

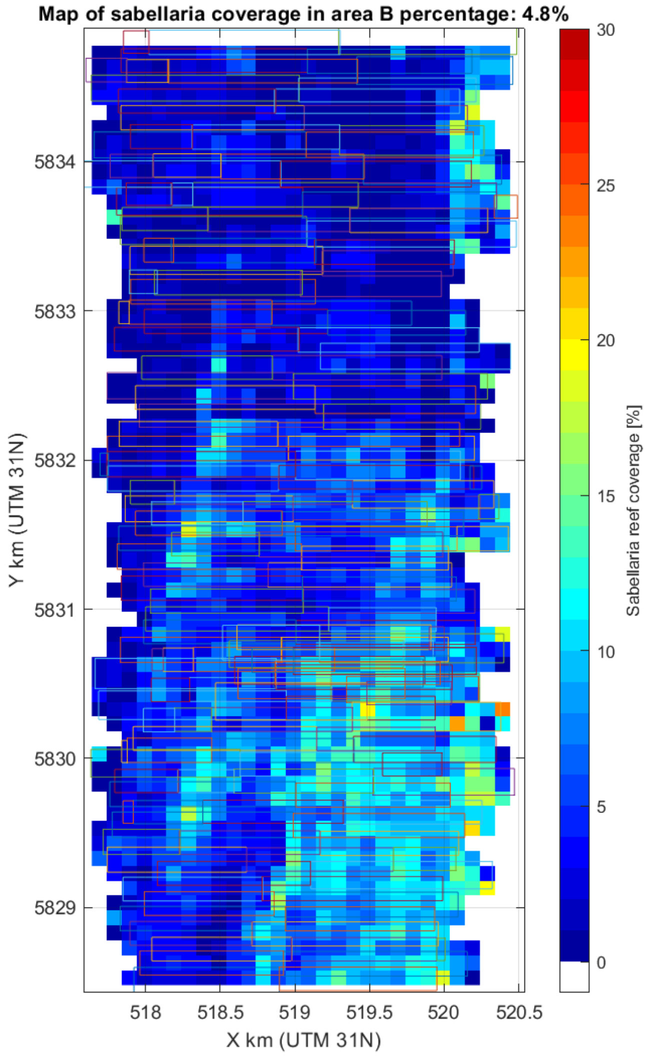


Figure 2.8: *Sabellaria* percentage presence map for area B. Total coverage of the seabed with elevated *Sabellaria* is 4.8%. The cell size is 100 x 100 m. Map is based on data from ground range between 30 and 75 m. The thins lines indicate the seabed coverage per SSS track line.

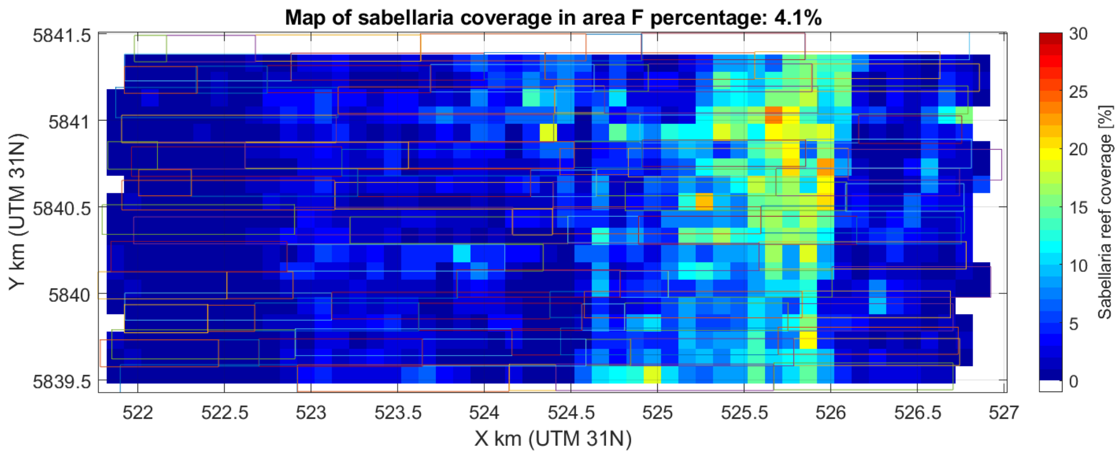


Figure 2.9: *Sabellaria* percentage presence map for area F. Total coverage of the seabed with elevated *Sabellaria* is 4.1%. The cell size is 100 x 100 m. Map is based on data from ground range between 30 and 75 m. The thins lines indicate the seabed coverage per SSS track line.

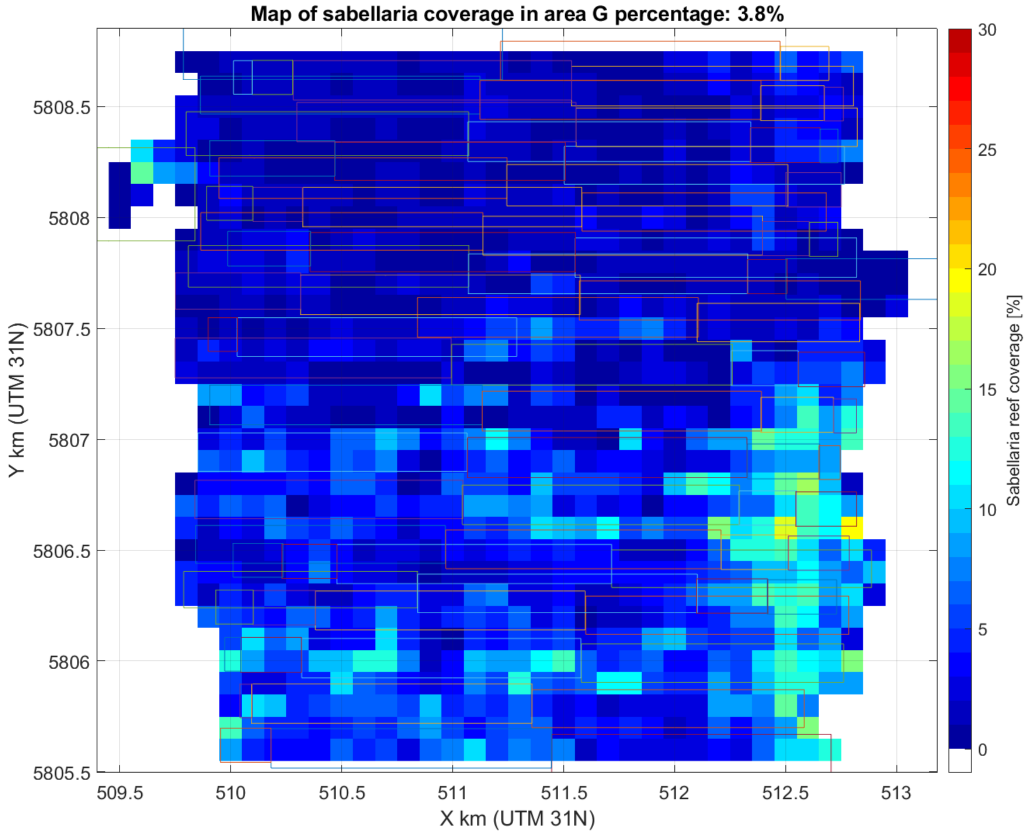


Figure 2.10: *Sabellaria* percentage presence map for area G. Total coverage of the seabed with elevated *Sabellaria* is 3.8%. The cell size is 100 x 100 m. Map is based on data from ground range between 30 and 75 m. The thins lines indicate the seabed coverage per SSS track line.

3 Habitat analysis

3.1 Introduction

A variety of datasets about the seabed from different sensors and devices was collected during the MONS project: boxcore, video, multibeam echo sounder and side scan sonar data. The methodology and results obtained on board are described in the cruise report [4].

Each dataset has advantages and disadvantages towards type of information, spatial scale, positioning accuracy and being a direct or indirect representation of the seabed state (Table 3.1). For the *Sabellaria* habitat analysis the acquired dataset was exploited and different information was combined. The main focus is on the evaluation of the preference of *Sabellaria* towards seabed morphology and sediment composition.

The boxcore samples provide a spatially limited (low coverage) but exact information about the sediment composition and presence of *Sabellaria*. The video recordings provide visual information about *Sabellaria* presence and about the sediment composition of the top surface layer. It covers a larger area but still required a manual interpretation of an ecological expert. The MBES data is a highly accurate and a large-scale measure of the bathymetry providing exact information about the seabed morphology. In addition the MBES data provides an indirect, large-scale measure of the sediment composition via the co-recorded backscatter data. In combination with the boxcore ground truth data, the information about the sediment composition can be extrapolated to a full and large-scale sediment composition map. This can be done with a clustering approach in combination with a correlation analysis using the ground truth data. Finally, the SSS data is suitable to detect elevated *Sabellaria* over a large area. Via a machine-learning algorithm, a large-scale map of the *Sabellaria* presence was created from the SSS data in combination with the ground truth data. This approach is only an indirect measure of the *Sabellaria* presence and contains a higher uncertainty than the ROV video data but it provides large-scale information.

Table 3.1: Overview of characteristics of various datasets acquired during fieldwork.

Dataset	Information type	Type of measure	Spatial coverage	Positioning accuracy
Boxcore samples	Sediment composition, Sabellaria presence	Direct, physical sample	Low (cm2)	Moderate (0,5 to 1 m)
Video recordings	Sabellaria presence	Semi-direct (optical image but requires expert interpretation)	Moderate (m2)	Moderate (1 to 2 m)
MBES data	Seabed morphology, Sediment composition	Direct (morphology), Indirect via physical relationship (sediment)	High (km2)	High (few centimetres)
SSS data	Sabellaria presence	Indirect via pattern recognition	High (km2)	Low (up to 30 m)

3.1.1 ROV video recordings

The video recordings were analysed by Waardenburg Ecology and Eurofins Aquasense into 10 classes: sand, sand with low, medium and high shell coverage, and flat reef as well as elevated reef with low, medium and high coverage. The classes were defined based on the sediments and *Sabellaria* patches observed in the preliminary analysis during the trial and taking into account the assessment criteria as described in Gubbay (2017) [19]. The video classification is reported on a 5 s interval. In the majority of the video tracks just sand (66%) was discovered. Flat reefs were more abundant on the seabed, with 20%, than elevated reefs, with 3%. Around 12 % of the seabed has significant shell coverage.

Table 3.2: Classification of video tracks and track distance per class. Low is defined as < 25%, medium as 25 to 50 % and high as > 50% coverage.

Classification	Total distance	Percentage
all ROV tracks	14994.47 m	100 %
Sand + Shell low	1418.56 m	9.46 %
Sand + shell medium	347.45 m	2.32 %
Sand + shell high	15.40 m	0.10 %
Sand	9913.87 m	66.12 %
Flat low	2476.63 m	16.52 %
Flat medium	285.97 m	1.91 %
Flat high	90.80 m	0.61 %
Elevated low	366.01 m	2.44 %
Elevated medium	75.95 m	0.51 %
Elevated high	3.42 m	0.02 %

3.1.2 Boxcore samples

A grain size analysis was carried out for boxcore samples from 29 locations by Geonius geotechniek B.V. (Figure 1.2) In addition to a visual classification (NEN5104), an identification and description were performed in accordance with NEN-EN-ISO14688-1. The grain size distribution diagrams were made in accordance with the Deltares protocol for sieve curve determination, calculated with salt correction method. For each location the top layer was analysed and the grain size distribution was determined. In 17 locations, a second buried layer was found and in two locations a third layer was detected and both additional layers were analysed as well. In nine samples of the top layer (33%) and in seven samples of a buried layer (41%) *Sabellaria* was found. For some boxcores the aim was to sample the *Sabellaria* reef but for some others the aim was to sample varying patterns in the MBES backscatter data. That means the sampling strategy was neither fully steered towards sampling *Sabellaria* detected in the sonar and video data nor completely random. Therefore, the high percentage of *Sabellaria* detection has to be interpreted with caution.

3.2 Seabed morphology analysis

The seabed morphology analysis is about the evaluation of the preference of *Sabellaria* towards a certain seabed morphology. The analysed video recordings were compared with the MBES data. Both datasets provide a high geographical positioning accuracy with a few cm for the MBES data and 1 to 2 m for the video recordings. From the MBES bathymetry data several second-order derivatives, i.e. slope and BPI (Bathymetric Positioning Index) with different settings, were computed providing additional descriptors of the morphology.

Bathymetric positioning index (BPI)

Bathymetric positioning index (BPI) is a measure of where a location is in vertical space relative to surrounding locations. It subtracts the median bathymetry computed over an inner radius from the median bathymetry over an outer radius. Setting the inner radius to 2 m and the outer radius to 20 m, the trough of megaripples with wavelengths around 5 to 15 m (being representative for the study area on and around the Brown Bank) has negative values, the slope values around zero and the crest positive values. Setting the inner radius to 8 m and the outer radius to 40 m, a focus is on megaripples with a wavelengths around 20 to 30 m, also representative for the Brown Bank area.

The bathymetric values were computed for a radius of 2 m around each video recording sampling point, thereby accounting for the positioning error of 1 to 2 m of the ROV track. The correlation of the video tracks with the bathymetry, slope, BPI (2-20 m), BPI (8-40m) and MBES backscatter are shown in Figure 3.1.

Based on the data there is no preference of flat or elevated reef towards a certain depth and slope. This is also confirmed by the correlation between the SSS-based *Sabellaria* map and the MBES bathymetry in Figure 3.2. The BPIs being indicative for megaripples with a wavelength between 5 and 30 m do not clearly highlight the occurrence of flat or elevated reefs in the troughs. Solely the elevated reef with high seabed coverage indicates some preference to the trough of the smaller megaripples with negative values (BPI 8-20). However, only ~ 10 sample points exist for this class which is highly undersampled compared to the other classes. In contrast to the data-driven results, the visual inspection of the classified SSS image showed rather a tendency of the elevated reefs towards the troughs and slopes of the megaripples. A possible explanation could be that the positioning uncertainty of the video recordings was higher than the expected 1 to 2 m for the ROV track. This uncertainty estimate was based on personal communication with the ROV pilot and not further investigated.

Furthermore, the plots show that the MBES backscatter measurements are not sensitive enough to allow distinction between reef and the surrounding sediment. Possibly, the *Sabellaria* patches are not densely populated enough to have a significant effect on the acoustic backscatter. In general, the range of backscatter values is quite small indicating similar geoacoustic properties (a similar hardness and roughness) between the different video classes. The values for the class “shell with high coverage” and “elevated reef with high coverage” should be considered with caution since both classes have only a limited number of samples (< 20 samples).

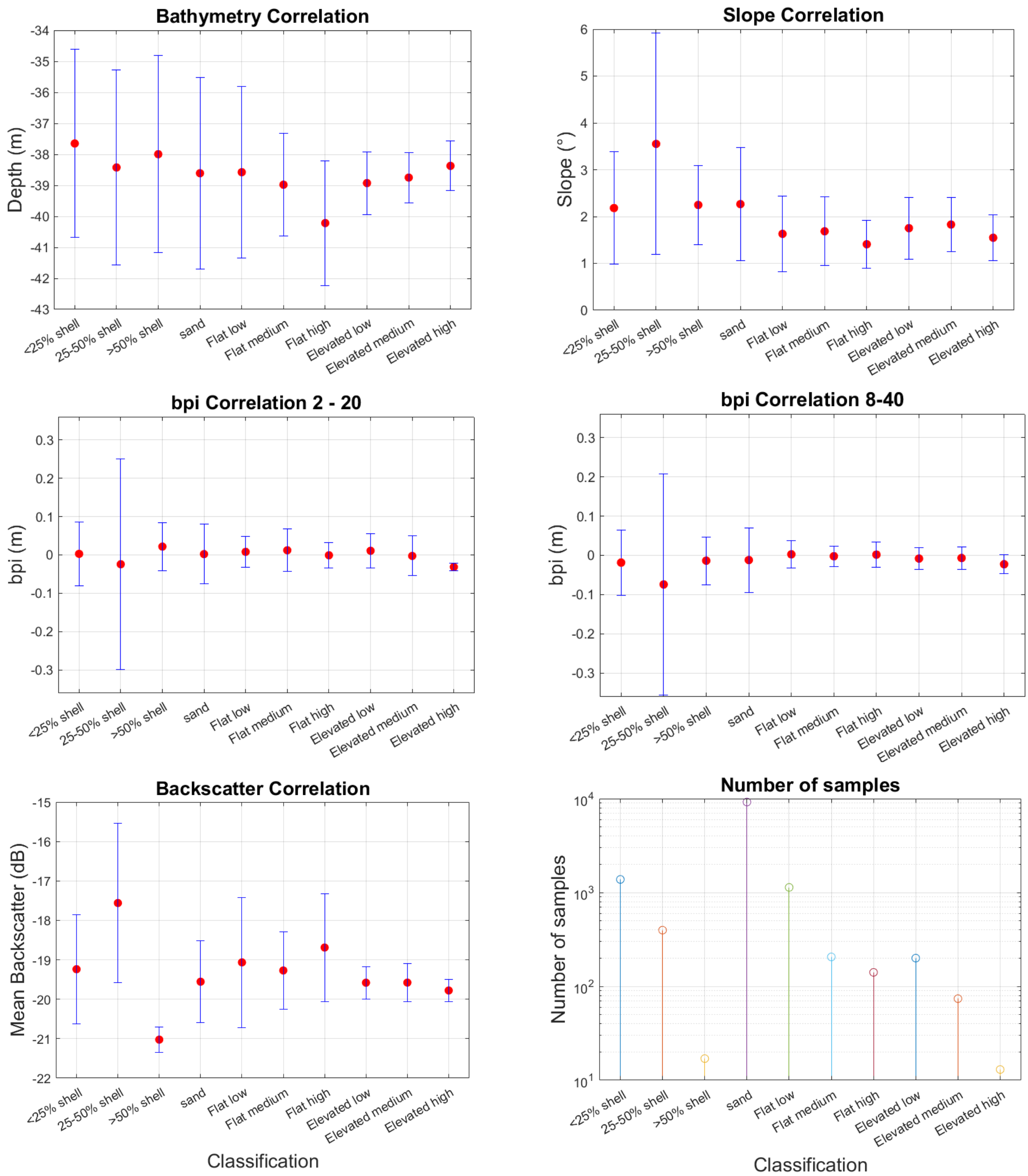


Figure 3.1: Correlation between MBES data and seabed class determined from video recordings. Primary data as MBES backscatter and bathymetry and secondary derivatives as slope and bpi are displayed. Video recordings are classified by WE. The red dots indicate the median and the blue error bars the standard deviation.

The comparison between the SSS-based *Sabellaria* reef-presence map and the MBES bathymetry data shows that *Sabellaria* does not occur at a specific water depth within the surveyed area in the Brown Bank area (Figure 3.2). The acquired and processed data indicate that *Sabellaria* settles without a preference towards a water depth.

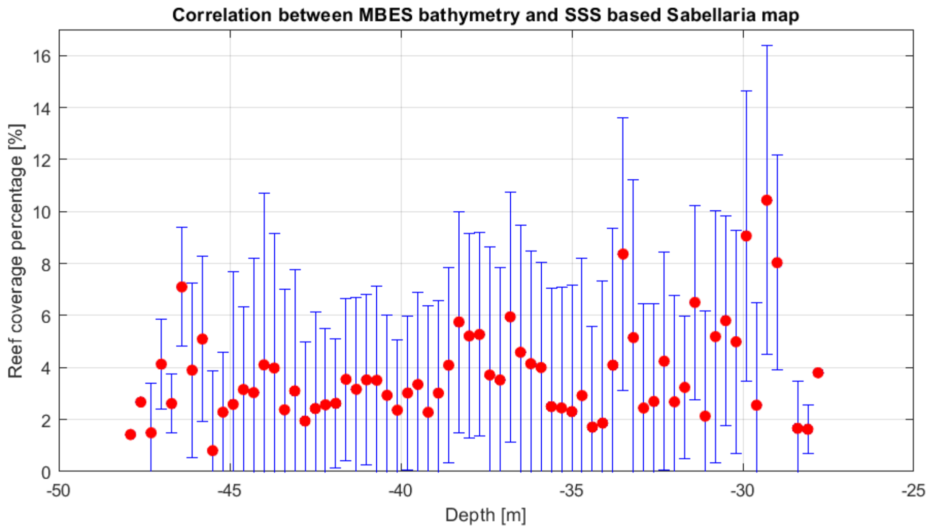


Figure 3.2: Correlation between *Sabellaria* reef seabed coverage percentage obtained from SSS classification (100 m resolution) and water depth from MBES bathymetry (1 m resolution). Red dots indicate median depth and blue error bars indicates standard deviation.

3.3 Sediment composition analysis

To analyse the preference of *Sabellaria* towards a certain sediment type, several combinations of the datasets can be exploited: (1) Boxcore samples solely, (2) Boxcore samples with MBES-based sediment roughness/coarseness map (3) video recordings with MBES-based sediment roughness/coarseness map, (4) SSS-based *Sabellaria* presence map with MBES-based sediment roughness/coarseness map.

3.3.1 Boxcore samples

The grain size analysis of the boxcore samples revealed the median grain size, mud, sand and gravel content. In case sediment layering was present samples were taken from each layer. This information was compared with the presence of *Sabellaria*. In the analysis below no distinction was made between flat and elevated *Sabellaria* as was done in the video classification and later in the biodiversity analysis. Figure 3.3 indicates that the four sediment parameters are not different between the samples containing *Sabellaria* and no *Sabellaria*. Furthermore, Figure 3.3 shows that the buried layer has significantly higher mud content and lower sand content than the bottom layer. In the analysed samples *Sabellaria* occurs in the top and bottom layer as well as in more muddy and more sandy sediments. This limited dataset indicates no clear tenendcy of *Saberllaria* towards a certain sediment type. Since the sample location was selected based on the acoustic data aiming for sampling *Sabellaria* and not randomly, it is difficult to draw general conlsuions whether Sabellaria has a preferred sediment type or infuences the sediment comopsitions.
Even though the boxcore samples provide a highly accurate representation of the sediment composition the dataset is limited. To obtain a more extended comparison with the sediment composition, the remote-sensing dataset over the different areas, the SSS-based *Sabellaria* presence map and the video classification are compared with a MBES-based sediment roughness/coarseness map in the next section.

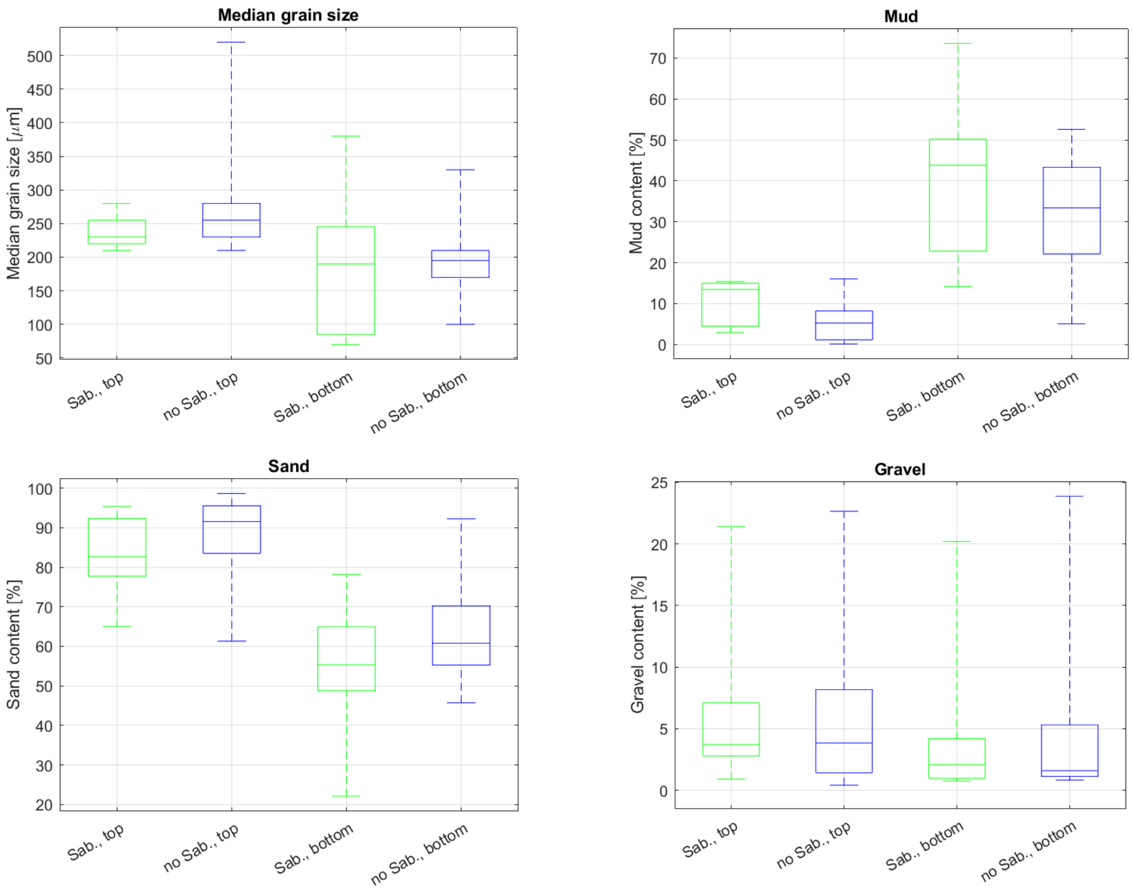


Figure 3.3: Correlation between presence of *Sabellaria* and a variety of sediment parameters obtained from the grain size analysis of the boxcore samples. A distinction between top and bottom layer was made. The box indicates the 25th and 75th percentiles, the horizontal line within a box the median and the whiskers the maximum and minimum values.

3.3.2 Sediment map based on MBES backscatter

A correlation analysis was carried out between the sediment map derived from the MBES backscatter data and the *Sabellaria* information obtained from the boxcore samples and the video recordings. The MBES-based sediment maps for area A and B are shown in Figure 3.4 (Figure C.4 shows maps for area F and G in the Appendix C). Owing to a lacking correlation with the grain size data of the boxcore samples, but still seeing a clear variation in backscatter values between the classes, it was decided to assign only the degree of sediment coarseness and hardness to each class. This decision was based on the general relationship between backscatter and sediment coarseness and hardness described in literature([24] [25]). The lacking correlation could have been caused by (1) shells, shell fragments or biota on the seabed which were not considered in the grain size analysis, (2) small-scale ripples or bioturbation or (3) a varying compaction degree of the sampled sediment. All factors contribute to the backscattering and can mask the effect of the grain size on the backscattering.

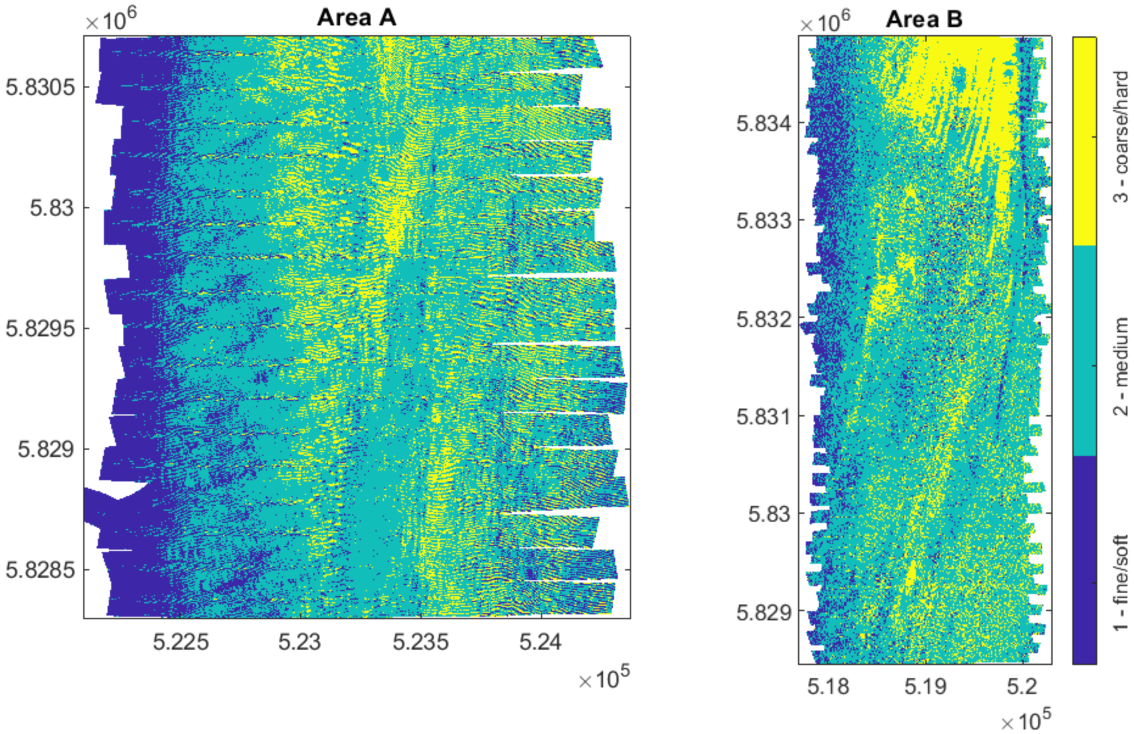
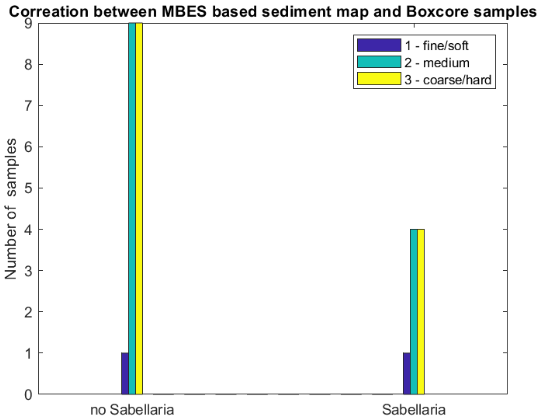


Figure 3.4: Sediment coarseness/roughness map of (left) Area A and (right) Area B. Class 1 corresponds fine/soft sediment, Class 2 to medium coarse/hard sediment and Class 3 to hard/coarse sediment. The maps for area G and F are shown in Appendix C.

Figure 3.5 shows the correlation of the MBES-based sediment map with the presence of *Sabellaria* found in the boxcore samples. *Sabellaria* occurs in all sediment classes with different degrees of roughness and hardness. The percentual presence and absence of *Sabellaria* occurs almost equally among all sediment classes. While this comparison indicates no tendency of *Sabellaria* towards a certain sediment (being in alignment with the results from Section 3.1.2), the comparison with video recordings indicates a tendency of the elevated *Sabellaria* towards finer and softer sediments (i.e. class 1 and 2). Even though the uncertainty in the video classification and MBES-based sediment map is higher than in the boxcore samples, it provides a more extensive dataset over a larger area. Note that the number of samples of the elevated, high-coverage reef class is still limited (<20 samples, see Figure 3.1).



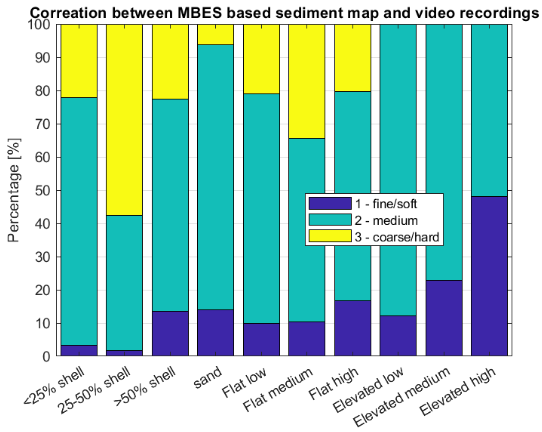


Figure 3.5: Correlation between presence of *Sabellaria* and sediment roughness and hardness. Sediment roughness and hardness classes were derived from MBES backscatter. Presence of *Sabellaria* obtained from Boxcore samples of top layer (left). Relation between presence of flat and elevated *Sabellaria* and additional seabed classes and sediment roughness and hardness (right). Presence of *Sabellaria* and additional seabed classes obtained from ROV video recordings.

The fourth comparison is between the MBES-based sediment map and SSS-based *Sabellaria* reef-presence maps. Figure 3.6 shows the median, 25th and 75th percentiles and minimum and maximum values of the reef seabed coverage per sediment class. Even though the median indicates a tendency of high *Sabellaria* coverage towards coarser and harder sediment, the 25th and 75 percentiles highly overlap between the sediment class. This rather indicates that the presence of *Sabellaria* is statistically independent of the sediment roughness and hardness. However, this results should be seen in light of the limited dataset.

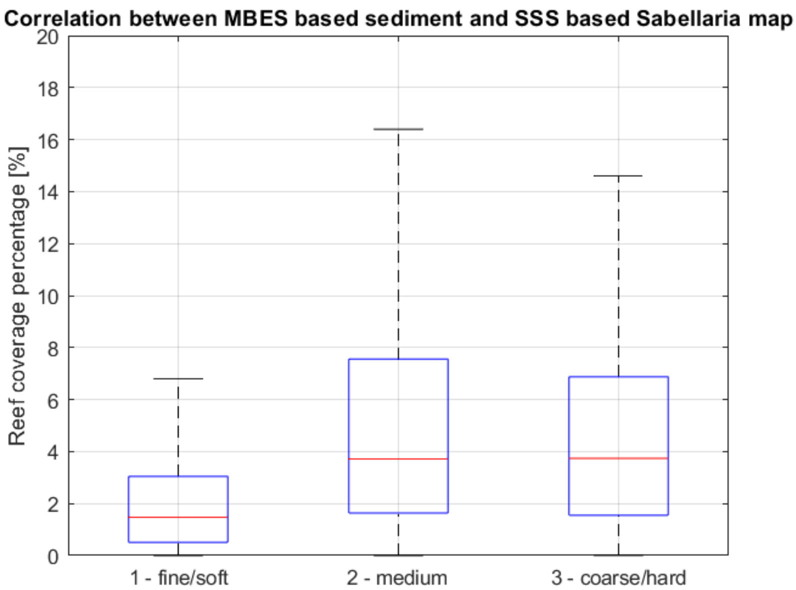


Figure 3.6: Relation between reef percentage obtained from SSS classification (100 m resolution) and sediment roughness/hardness obtained from MBES backscatter classification (5 m resolution). The box indicates the 25th and 75th percentile, the horizontal stripe within each box the median and the whiskers the maximum and minimum value of all values that are not considered as outliers.

3.4 Comparison with previous *Sabellaria* suitability maps

As the maps produced within this work can be used to improve models aimed at predicting the presence of *Sabellaria*, in this section the *Sabellaria* presence maps are compared against model prediction by Deltares on the presence of *Sabellaria* [11]. The input parameters for the model predictions by Deltares were bottom shear stress, water depth, three BPI's with 5, 10 and 75 km inner radius, temperature, salinity, gravel and mud content. Figure 3.7 to Figure 3.10 below show the SSS data driven maps of the elevated *Sabellaria* reef coverage (left) next to the probability estimate for *Sabellaria* reefs (right). It is important to note that the metric displayed in the map is not the same and that more work is needed to allow for a quantitative comparison. Therefore, a direct comparison between the percentages of both maps is not possible. However, a first comparison of the maps illustrates spatial differences in the distribution where the center of the hotspots (highest values) in both maps are between 1 and 3 km apart.

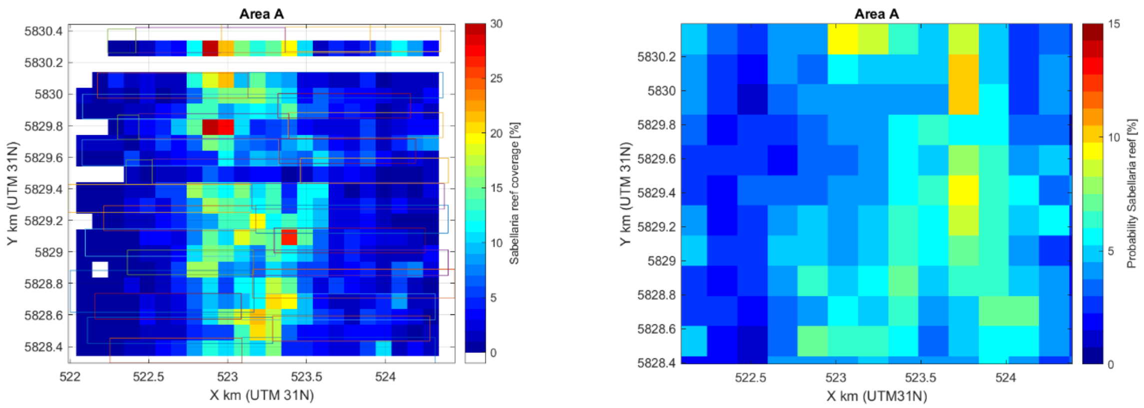


Figure 3.7: *Sabellaria* reef-presence map obtained from TNO AI classifier (left). Probability *Sabellaria* presence map generated by predictive modelling using Random Forest from Deltares (right). Results for Area A are displayed.

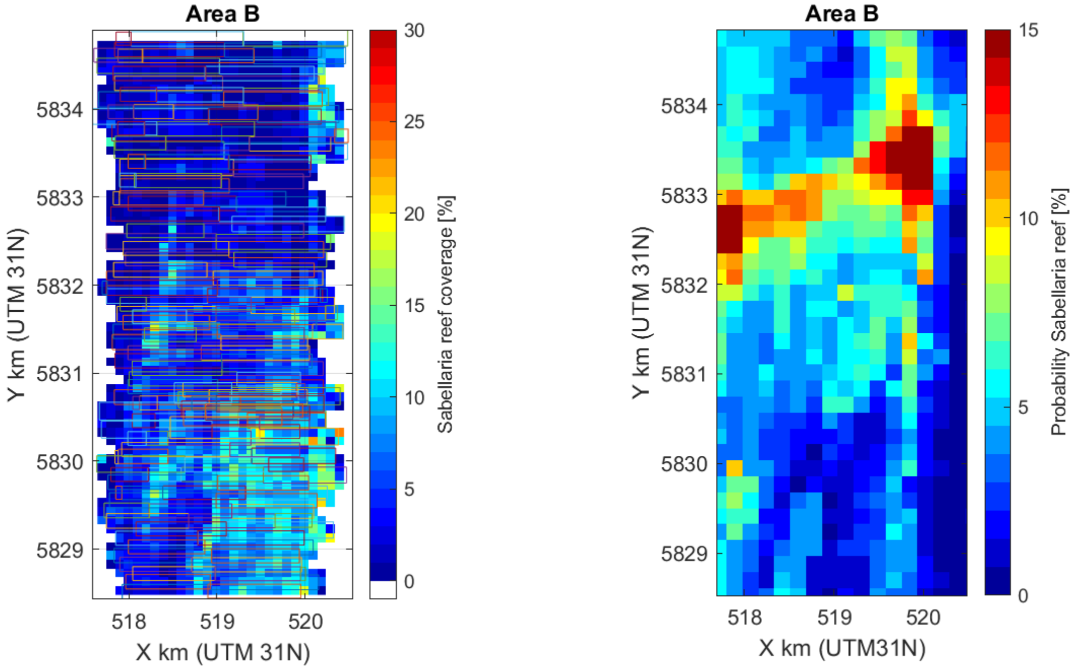


Figure 3.8: *Sabellaria* reef-presence map obtained from TNO AI classifier (left). Probability *Sabellaria* presence map generated by predictive modelling using Random Forest from Deltares (right). Results for Area B are displayed.

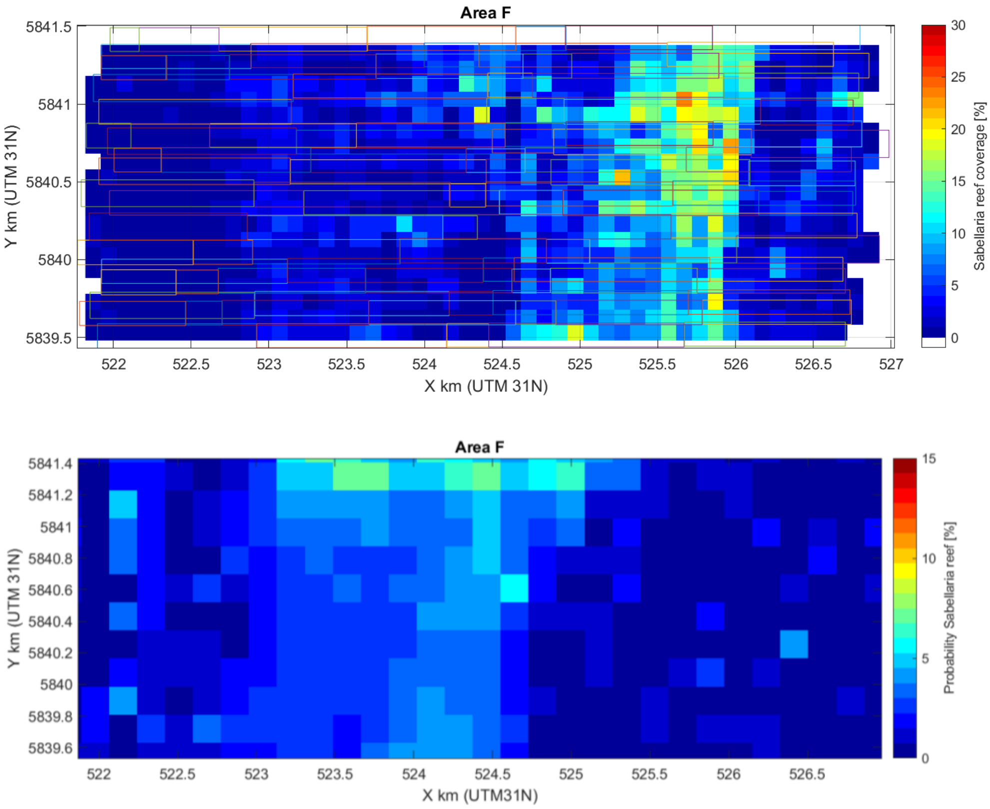


Figure 3.9: *Sabellaria* reef-presence map obtained from TNO AI classifier (left). Probability *Sabellaria* presence map generated by predictive modelling using Random Forest from Deltares (right). Results for Area F are displayed.

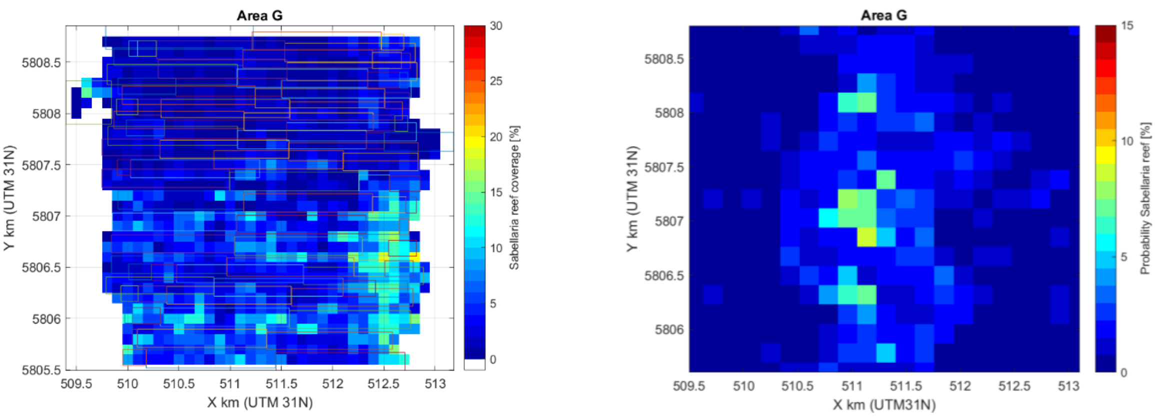


Figure 3.10: *Sabellaria* reef-presence map obtained from TNO AI classifier (left). Probability *Sabellaria* presence map generated by predictive modelling using Random Forest from Deltares (right). Results for Area G are displayed.

4 Biodiversity analysis

4.1 Introduction

In this chapter, the results of the boxcore samples which are analysed by Waardenburg Ecology en Eurofins Aquasense are described and compared with current knowledge of the Brown Bank and other areas in the Dutch North Sea.

4.2 Methods

The collected samples were collected and analysed according to Rijkswaterstaat Protocols (RWSV) A2.107 and A2.120, however due to their large size some steps of the protocol were adapted. The adaptations involved increased size fractionation, crushing of the tubes, subsampling of the finest fraction (one sample), and identification of e.g. Bryozoa and Hydrozoa to species level (i.e. where possible). In the following paragraphs the used method is described in detail. A visual overview of the processing steps is given in Figure 4.1.

Upon arrival of the samples in the laboratory, an entry check of the samples was made for completeness (i.e. jar undamaged and sample data label correct) and whether the samples were correctly fixed. The entire sample was transferred to a 500 µm sieve. The samples were rinsed in the sieve with tap water. One or more sieves with a larger mesh size were placed on the finer sieves (Figure 4.1). The mesh size of the finest sieve was 500 µm. Furthermore, sieves with mesh sizes of 10 mm, 4 mm, 2 mm and 1 mm were used in the analyses. *Sabellaria* tubes were removed from the fractions and carefully crushed to obtain all organisms in the tubes. The crushed tubes were placed back in the sieve and processed further. In one case (sample F1-B2), the finest fraction was subdivided into subsamples. Since almost all organisms were already removed from the sample (>90%), it was decided to take a sub-sample based on volume and to deviate from the protocol. Of the finest fraction, 18.75% was analysed. The other samples were not divided into subsamples and were sorted in their entirety. The rinsed sample was sorted into plastic screening containers with tap water under the binocular. All organisms and fragments of organisms were removed from the samples and sorted by species group (Polychaeta, Crustacea, Mollusca, Echinodermata and other). The organisms were preserved in 70% ethanol and kept for identification. The residual material was stored in a sample jar in 4-6% formaldehyde.

All organisms were identified to species level or lowest taxonomic level possible (in accordance with Analysis instruction A2.107, version 8). Bryozoa and Hydrozoa, were also identified to species level. Due to the recent discovery of *Aglaophamus agilis* outside the Cleaver Bank, it was decided to report all juvenile Nephtyidae at family level. This deviated from previous Waardenburg Ecology and Eurofins Aquasense reports, which reported juvenile shimmy worms (*Nephtys*) at the genus level. Because several new and unique species for the Netherlands were found in these samples, the species list is not yet definitive. Some species were sent to experts abroad and have not yet been given a final name. They have currently been identified at a higher taxonomic level.

The naming was noted in accordance with the most current TWN (Taxa Waterbeheer Nederland) list (22-07-2024). If necessary, methylene blue combined with chrystal violet or methyl green was used for the identification. These dyes enhance certain difficult-to-see features. During the identifications, the heads of the organisms were counted per species. In the case of Polychaeta, for example, many individuals were often damaged and incomplete. The headless parts were collected and pooled with the complete individuals of the same taxon for determination of ash-free dry weight.

The biomass was determined for each taxon in each sample, expressed in ash-free dry weight (AFDW). The direct ashing method was chosen to determine the biomass, in accordance with the protocol. At least two specimens of all species were kept aside for the reference collections. Several specimens of some rare and new species were also kept separate, so that they could be sent to experts, for example. In the absence of one biomass value, the value -9999 was entered.

The macrofauna analyses were carried out in Waardenburg Ecology's and Eurofins Aquasense ISO 17025 accredited lab in accordance with the following protocols:
- *A2.107 Waterbodem marien – Uitzoeken en determineren van macrozoöbenthos* (version 8, September 2021);
- *A2.120 Bepaling biomassa macrozoo¨benthos* (version 4, September 2021).

To compare results with previous samples, the MWTL (Monitoring Waterstaatkundige Toestand des Lands) dataset of the Brown Bank and Cleaver Bank were used from 2021. Every three years, boxcore samples are collected on the Dutch continental shelf and analysed according to the very same RWS protocols used in this study. These lab analyses were executed by the same analysts executing this study, thus results are comparable. In this study, only the dataset of 2021 of the Brown Bank and Cleaver Bank was used as comparison. For details on the standard monitoring, see van Son et al. (2022) [1].

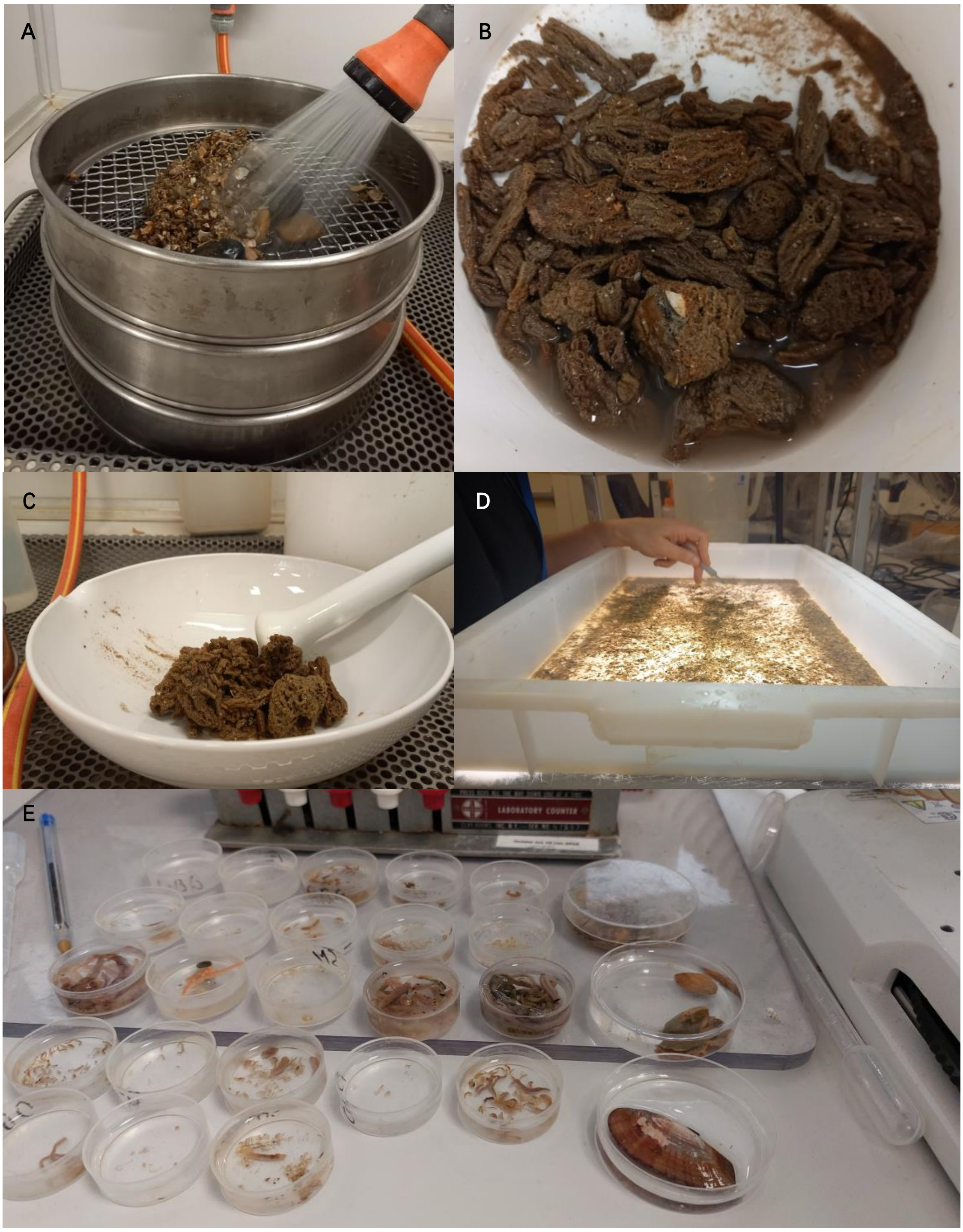


Figure 4.1: Processing steps in the lab. A: samples were rinsed and sieved in the lab. B: clumps of *Sabellaria* were separated from the rest of the sample. C: *Sabellaria* clumps are carefully crushed to obtain all organisms in the tubes. D: all organisms were carefully picked from the size-fractionated samples. E: an impression of the number of organisms in a sample.

4.3 Results

4.3.1 *Sabellaria* survey samples

During the two-week sampling campaign (August 21 to September 1, 2023), a total of nine boxcore samples (0,078 m^2) were collected in which *Sabellaria* tubes were visibly present in the field. Samples that did not contain *Sabellaria* tubes in the field were rejected and excluded from further analysis. Boxcore samples were taken in two search areas, B and F. In area B, samples were taken at four locations: B1, B3, B4 and B9. At location B1 and B4, a duplicate sample was taken (<1 meter). In area F, three samples were taken. These three samples were taken within a radius of 50 meters.

In total, 243 taxa were found in the nine samples (Table 4.1). The highest number of taxa found in one sample was 112 (F1-B2)) and the lowest number was 49 (B1-B1). It was striking that some species exhibited very high abundance in the samples. Sample F1-B3 contained the most individuals; in total, more than 50.000 per square meter. The lowest number of individuals was observed in sample B1-B1, concurring with the previously noted low number of taxa in this sample; just over 3.000 individuals per square meter. In the samples, a considerable number of rare species for the Netherlands were found. Some were previously only known from another area (such as the Cleaver Bank) and at least seven species are new to the Netherlands.

The measured biomass values were lower than the actual collected biomass (Table 4.1). This is because quite a lot of individuals were kept separately, and some biomass values did not meet the quality requirements. The described values, however, provide a general impression in which samples the most biomass was observed; samples F1-B2 and F1-B3. The biomass values were mainly determined by the presence or absence of burrowing megafauna and to a lesser extent by shellfish.

Table 4.1: Key figures *Sabellaria* survey 2023 (n indicates number of individuals).

Location	Number of taxa	Individuals (n/m^2)	Biomass (mg/m^2)
B1-B1	49	3154	.409.5
B1-B2	60	8854	7493.0
B3-B1	55	9490	3931.2
B4-B11	93	19179	7898.3
B4-B12	72	10382	6866.9
B9-B7	85	16624	9754.8
F1-B1	74	12675	8254.8
F1-B2	112	44197	23466.7
F1-B3	87	50188	16348.4
Sum	243		
Median	74	12675	7898.3
Max	112	50188	35840.4
Min	49	3154	3931.2

4.3.2 *Sabellaria*-reef types

The nine samples consisted of three different types of *Sabellaria spinulosa* "reef":

A. some loose *Sabellaria spinulosa* clumps (locations B1-B1, B1-B2, and B3-B1);
B. reef-structures with elevated *Sabellaria spinulosa* (B4-B11 and B4-B12); and
C. reef-structures of *Sabellaria spinulosa* in the sediment (B9-B7, F1-B1, F1-B2 and F1-B3).

Type A – Loose clumps of *Sabellaria spinulosa*
Three samples (B1-B1, B1-B2, and B3-B1) contained few small *Sabellaria spinulosa* clumps, almost all of which laid loose on the sand (Figure 4.2). The top sediment layer of these three samples was just below 10 cm thick and consisted of medium sand (Wentworth scale). Below the top sediment layer was a silt-rich sediment (50% silt, 50% sand).

Compared to the other identified *Sabellaria spinulosa* reef types, fewer species and individuals were found in type A. The samples contained little or no living *Sabellaria spinulosa* but did contain some clumps with empty tubes (Figure 4.2). Relatively abundant animals were found in these tubes, mostly Sipuncula (peanut worms, deposit feeders) and Nemertea (ribbon worms, predators).

Figure 4.2: Type A sample (location B1-B2), containing some empty Sabellaria tubes, a sandy top layer and a silty sublayer.

Phoronida (horseshoe worms) were abundant (558 n/m^2) in all three samples. These small worm-shaped organisms with a horseshoe-shaped tentacle crown live in a sandy tube in the sediment and sometimes occur in very high densities in the Dutch North Sea. The remaining fauna was characterised by species that are either characteristic of the Brown Bank and

Zeeland coast, or by species that occur on the Cleaver Bank or Oestergronden. They occurred in low densities in the samples.

The clumps found in the samples, assigned to Type A, does not form a *Sabellaria spinulosa* reef according to OSPAR [9] and Gubbay (2007) [19]. The *Sabellaria* is not elevated, occurs in low densities and the biodiversity is relatively low (compared to the other described types). It should be noted that in this region, *Sabellaria spinulosa* reefs probably could potentially develop that do meet the definition of a *Sabellaria spinulosa* reef. A hypothesis is that the clumps are broken parts of a former reef, which have been destroyed by natural or human action.

Type B – Elevated *Sabellaria spinulosa*
In some areas, elevated *Sabellaria spinulosa* (>5 cm) was found using the ROV. Area B4 was located between two mega ripples, east of the deepest part of the trough. Here, the highest density (>50% coverage) and highest reefs (>10 cm) were found. Two samples were taken next to each other (B4-B11 and B4-B12). As with type A, the top layer consisted of sand with underneath a layer with approximately 50% sand and 50% silt/clay.

Striking in the type B samples is that the video showed relatively abundant larger species, such as *Cancer pagurus* (edible crab) and *Necora puber* (velvet crab), which are associated with a *Sabellaria spinulosa* reef. These observations were notable, as most tubes did not contain any *Sabellaria spinulosa*. The larger tubes that were clearly visible on the video images were almost all empty. On the outside of these tubes, young *Sabellaria spinulosa* worms had settled. These were mostly alive. In both samples, approximately one third of the tube was embedded in the sediment and the majority (>66%) rose up from the seabed (Figure 4.3). Both samples contained a typical *Sabellaria spinulosa* reef community supplemented with species known from a comparable sediment type. *Sabellaria spinulosa* (small specimens) were by far the most abundant (6.160 n/m^2). As is often found in a *Sabellaria spinulosa* reef, many crustaceans were present [19]. Small and juvenile crabs, such as *Pisidia longicornis* (1.718 n/m^2) were frequently abundant. These crabs were often found in or between the empty tubes. Another example of a numerous crustacean is the (sub) surface deposit feeder *Abludomelita obtusata* (987 n/m^2). An amphipod that is regularly found in the Dutch North Sea, but not in such high densities. Like type A, Phoronida (horseshoe worms) were also quite abundant in these samples (981 n/m^2).

The type B samples aligns closest with the described definitions for a *Sabellaria spinulosa* reef. Depending on how certain factors are weighted, the type B boxcore samples result in a reefiness score of medium to high.

This reef type was mapped by the SSS and classified as elevated reef. In the video classification this reef type is classified as elevated *Sabellaria*.

Figure 4.3: Type B sample (location B4-B11), containing elevated *Sabellaria*, a sandy top layer and a silty sublayer.

Type C –*Sabellaria* crusts

Sabellaria spinulosa reefs situated completely in the sediment, with tubes that were tens of centimetres long and protruded a maximum of a few centimetres above the seafloor, were defined as type C reefs. In area B, one sample of this type was taken (B9-B7, Figure 4.4) and in area F, three samples (F1-B1, F1-B2 and F1-B3). This type of reef occurs in the flat parts of the troughs and are hardly visible on the SSS and MBES (Figure 4.5) and could not clearly be mapped with the employed acoustic sytems. They are also very poorly visible on the video images, because they barely protrude above the sea floor (Figure 4.6). The video classification indicates that this reef type (classified as flat *Sabellaria*) might be the most common on the Brown Bank (see Table 3.2).

Sample B9-B7 was located in an area with approximately 15% silt/clay, 80% (mostly fine) sand and 5% gravel. Below this top layer (5-15 cm), a sediment layer containing on average 20-40% silt/clay and hardly any coarse sand and gravel was found.

Samples from area F consisted of ~80% (fine) sand and about 15% silt/clay (Figure 4.4). Below this top layer (~7 cm), a layer containing more gravel (20%) instead of sand (~7cm) was seen. The deepest layer seemed to contain marine clay. About 75% was silt/clay, 20% sand and 5% gravel. Where *Sabellaria spinulosa* was not present, the sediment composition was similar to the layer present below the *Sabellaria*; about 10-15% silt/clay with 65-70% sand and 15-20% gravel.

Figure 4.4: Type C sample (B9-B7) with a lot of *Sabellaria spinulosa*.

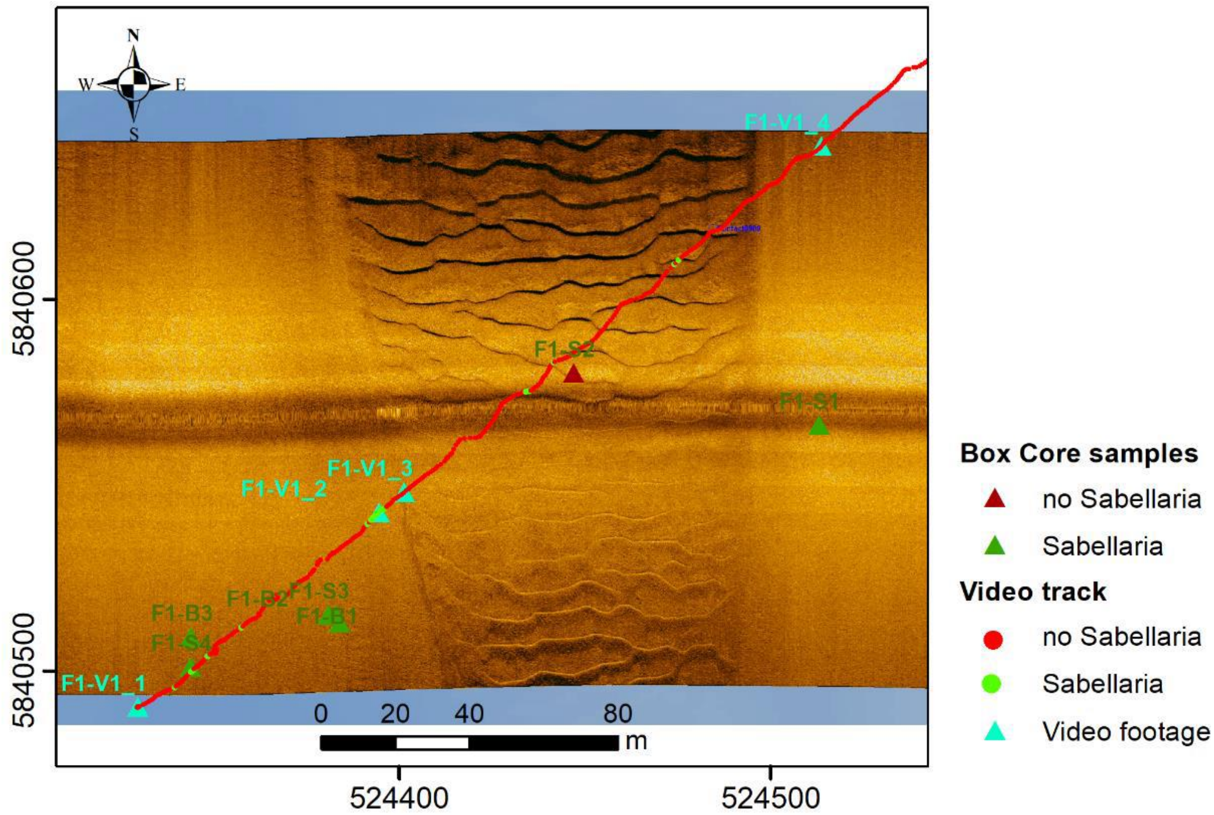


Figure 4.5: Side scan sonar image of area F, where *Sabellaria spinulosa* was found at the location marked with green dots (video track classification) and triangles (boxcore samples).

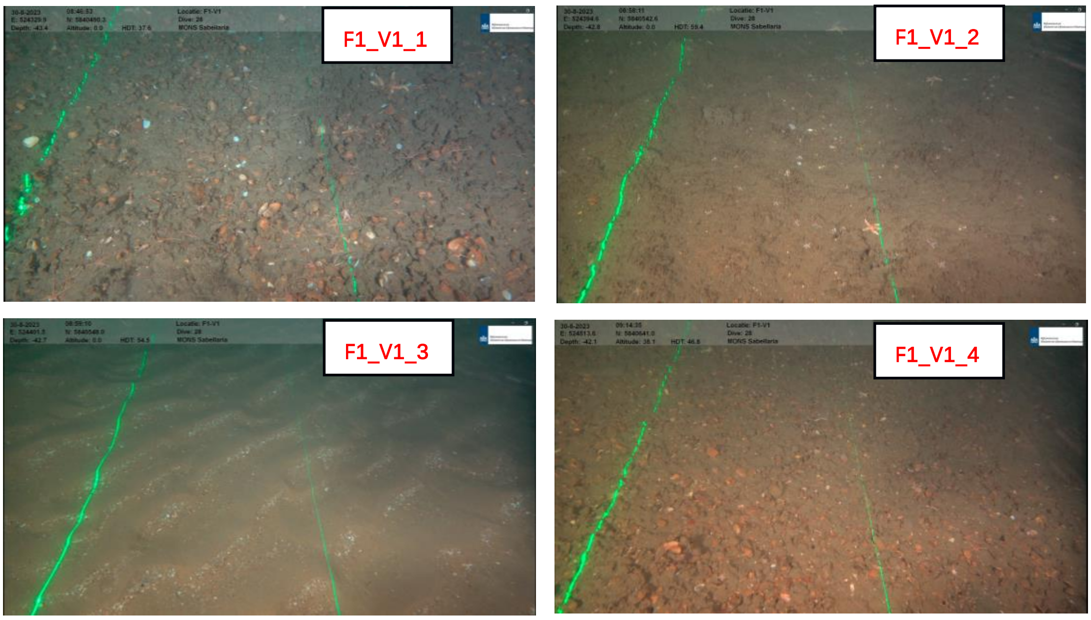


Figure 4.6: ROV video footage. Locations of this footage are shown by light blue dots in Figure 4.5. (Left top) F1-V1_1 showing shell bed with *Sabellaria spinulosa*, (right top) F1-V1_2 showing fine-grained sediment with *Sabellaria spinulosa*, (left bottom) F1-V1_3 showing fined-grained sediment and (right bottom) F1-V1_4 showing shell and gravel bed.

Slightly more species were found in the type C samples than in the type B samples. More striking was the number of new species for the Netherlands and that relatively many of the species found were previously only known from the Cleaver Bank or Oestergronden. Both areas contain a lot of fine sediment (silt/clay) and on the Cleaver Bank there is also relatively much coarser sediment, such as gravel, stones and boulders. By far the most common species in the samples was *Sabellaria spinulosa* (average: 15,353 n/m^2). The highest density of *Sabellaria spinulosa* occured in the type C samples. The community appeared to consist of a few larger specimens and many small individuals of *Sabellaria spinulosa*. In addition to *Sabellaria spinulosa*, another new species for the Netherlands was very abundant in the samples, *Aphelochaeta sp.* (6,612 n/m^2). Both species of worms live in the tubes of *Sabellaria spinulosa*. These worms were found in the tubes that were silted up. Another worm that was quite abundant was *Aonides oxycephala* (1840 n/m^2). Until now, this worm was only known from the Cleaver Bank in much lower densities (maximum 67 n/m^2).

According to the most common definitions of a *Sabellaria* reef (such as OSPAR), the *Sabellaria spinulosa* should protrude above the sediment. Something that hardly happened within the Type C samples. During an inter-agency workshop [19], it was indicated that it is not clear whether this form of reef should be considered as a stage in reef development or whether they are the result of growth in different environmental conditions. All *Sabellaria* types found in this research seem to have originated on either a shell, a pebble or a stone. In this respect there is no differences between the types. However, the type of habitat is different between types A, B and C. The type A and B reefs are mostly located between sandy megaripples, while the type C reefs were found in a more silt/clay sediment.

4.3.3 *Comparison between Sabellaria* samples from the Brown Bank and Cleaver Bank

Brown Bank

The *Sabellaria* samples, collected in this study on the Brown Bank, clearly differ from samples previously collected on the Brown Bank during the MWTL campaign 2021. The median species richness was four times higher (Figure 4.7, Table 4.2). Species richness in the *Sabellaria* samples (n=9) ranged between 49 and 112 unique taxa, whereas the species richness in the Brown Bank samples from 2021 (n=6) ranged between 13 and 29 unique taxa. Median macrofauna abundance in *Sabellaria* samples was 12 times higher (Figure 4.8, Table 4.3). Abundances in *Sabellaria* samples varied between 3154 and 50188 individuals per m^2, whereas in the Brown Bank samples it varied between 474 and 1821 individuals per m^2.

Cleaver Bank

In the Netherlands, the Cleaver Bank is the most species-rich area in terms of macrofauna [26]. Due to the presence of many different (micro) habitats, a lot off species can occur. The median species richness of the *Sabellaria* samples (n=74) was 1.5 times higher than the 2021 MWTL Cleaver Bank samples (n=49), while the total sampling area is 4.74 times lower (Figure 4.7, Table 4.2). Also, the median abundance was higher in the *Sabellaria* samples (n=12.675) by a factor of 4 compared to the Cleaver Bank samples (n=3.133).

These large differences between *Sabellaria* and Brown Bank samples are partly due to the difference in sampling area, the different sampling strategies and the presence of *Sabellaria* tubes. The total sampling area was 1.5 times larger in the *Sabellaria* survey. Sand banks, sand waves, mega ripples and small ripples are found on the Brown Bank. The Brown Bank samples of MWTL are generally taken on top of the sand waves, and sediment substrate therefore usually consists of medium to coarse sand. The macrobenthos community in these samples consists of a relatively low abundance of organisms that are typically found in sand wave habitats. In contrast, the *Sabellaria* samples were collected on flat seabeds and in the troughs of megaripples, and sediment composition consisted of fine sand or silty sand with some gravel, stones or shells. As a result, in addition to a number of typical fine to medium sand species, many species were also observed in the *Sabellaria* samples that were not yet known from this region or the Netherlands.

The samples were selected on board, based upon the visual presence of *Sabellaria* tubes in the boxcore samples. The higher habitat variability and the presence of *Sabellaria* clumps in these samples therefore also partly explains the large differences observed in species richness and abundance between the *Sabellaria* and Brown Bank samples..

The found species community in the *Sabellaria* samples partly correspond with Cleaber Bank samples. About 50% of the found species (n=120) were also found in the 2021 Cleaver Bank samples. It can be suspected that these species were found in the *Sabellaria* samples, because of suitable environmental conditions, partly due to the soil composition and partly due to the microhabitats created by the *Sabellaria* tubes.

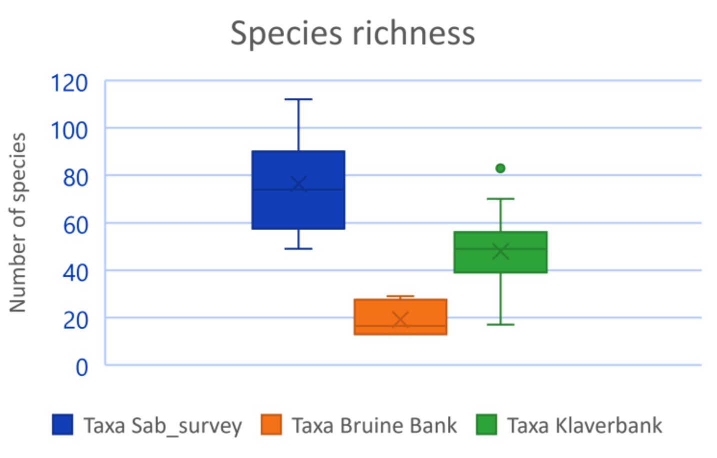


Figure 4.7: Species richness of *Sabellaria* samples (this study), Brown Bank (Bruine Bank) samples (MWTL study) and Cleaver Bank (Klaverbank) samples [1]. The box indicates the 25th and 75th percentile, the horizontal stripe within each box the median and the whiskers the maximum and minimum value of all values that are not considered as outliers (dots).

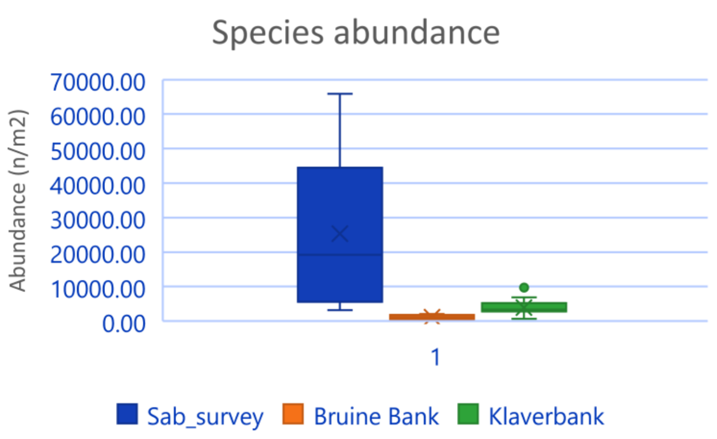


Figure 4.8: Abundance of *Sabellaria* samples (this study), Brown Bank samples (MWTL study) and Cleaver Bank samples [1]. The box indicates the 25th and 75th percentile, the horizontal stripe within each box the median and the whiskers the maximum and minimum value of all values that are not considered as outliers (dots).

Table 4.2: Species richness *Sabellaria* survey (this study), Brown Bank (MWTL, 2021) and Cleaver Bank (MWTL, 2021).

Area	Sabellaria survey	Brown Bank	Cleaver Bank
Number of samples	9	6	37
Total sampling area (m²)	0.702	0.468	3.33
Sum, number of taxa	243	62	326
Median number of taxa	74	17	49
Max number of taxa	112	29	83
Min number of taxa	49	13	17

Table 4.3: Species abundance *Sabellaria* survey (this study), Brown Bank (MWTL, 2021) and Cleaver Bank (MWTL, 2021)

Area	Sabellaria survey	Brown Bank	Cleaver Bank
Number of samples	9	6	37
Total sampling area (m^2)	0.702	0.468	3.33
Median n/m^2	19179	1.103	3389
Max n/m^2	65897	2.090	9711
Min n/m^2	3154	474	689

4.3.4 Rare species

A total of 243 different taxa were found in the 9 *Sabellaria* samples. This is an extremely high number compared to all other macrofauna projects in the Dutch North Sea. In addition, many observations have been made that are remarkable.

***Sabellaria spinulosa* reef species**
Some of the newly found species seem to be directly associated with the *Sabellaria spinulosa* reef. Not surprisingly, *Sabellaria spinulosa* was the most common species. More remarkable was that beside *Sabellaria spinulosa*, another worm from the same family was found: *Lygdamis muratus* (Figure 4.9). A large species that was previously unknown in the Netherlands, but had previously been found in England [27] and Belgium [28]. The tubes consist of much coarser material than *Sabellaria spp.* and always occurs in low densities. The species was found in samples B9-B7 and F1-B2. Both samples belong to a type C reef.

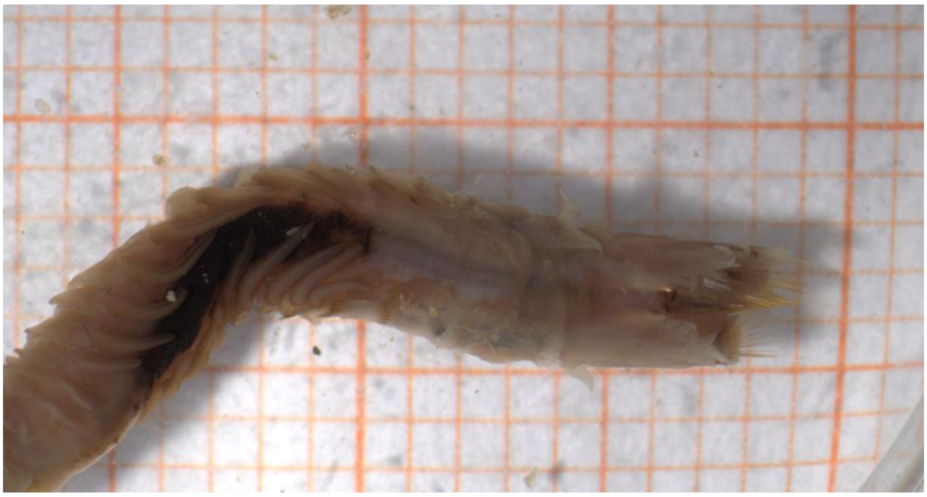


Figure 4.9: Lygdamis muratus. A new worm found in the Dutch part of the North Sea from the family Sabellariidae.

An unknown species of *Aphelochaeta* was found at all sampled locations. From one specimen in B1-B2 to hundreds of specimens in the type C samples. This worm had not been found before in the Netherlands and is possibly new to science. The worm lives in the *Sabellaria* tubes that are silted up. The worm closely resembles *Protocirrineris stormae*, a new species from the Delta, but has only one pair of gills.

Another example is *Syllides bansei* (Figure 4.10, left). This small worm was only found in the type B and C samples with relatively high numbers of *Sabellaria spinulosa*. Perhaps the species are dependent on each other. *Syllides bansei* was originally described in Florida (USA) and has been confirmed by an expert.

Another species that occurred regular in all the different *Sabellaria* types and is new for the Netherlands was the worm *Phyllodoce longipes* (Figure 4.10, right). All species of *Phyllodoce* known from the Netherlands were found in the *Sabellaria* samples. Possibly all species are predators and/or scavengers. Little is known about the factors determining which species occurs were. *Phyllodoce longipes* was for the first time described in Chile. However, it is question whether *Syllides bansei* and *Phyllodoce longipes* are exotic species. There is much discussion about their taxonomy, and it is expect that they do not yet have their own scientific name.

Figure 4.10: (left) Chaetae of *Syllides bansei*, found among others in sample B4-B11 and B4-B12. (right) *Phyllodoce longipes* from sample B4-B11. A new species for the Netherlands with its characteristic brown band under its head and the strongly asymmetrical prechaetal lobe (not visible in the figure).

Species only known from the Cleaver Bank

With the exception of B1-B2, the worm *Aonides oxycephala* was found in all samples. In the type C samples in fairly large numbers (average 1874 n/m^2). This relatively large worm determined an important part of the worms (Polychaeta) biomass. So far this species has only been found on the Cleaver Bank with a maximum density of 67 n/m^2.

The worm *Sphaerodorum gracilis* (Figure 4.11) is known at two locations on the Cleaver Bank. This species with a possible slightly northern distribution was first described in Norway. In reality, it concerns a cryptic species where multiple species hide under the same name. In the *Sabellaria* samples this species was only found in the three F samples (type C) with an average density of 22 specimens per square meter.

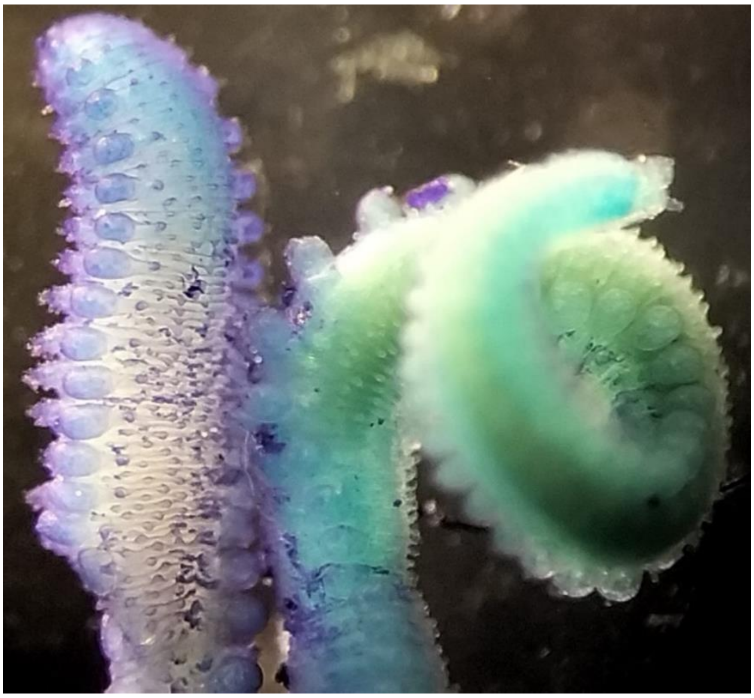


Figure 4.11: *Sphaerodorum gracilis*, a worm with conspicuous papillae on its body, stained with methyl blue.

Rare species for the Southern North Sea

Some of the species found occur mainly in the south of Europe. An example of such a species is the shell *Diplodonta rotundata* (Figure 4.12, left). So far only known from the Brown Bank region and wind farm Princes Amalia Windpark (PAWP) in the Netherlands. A total of seven specimens were found divided over sample B1-B2, F1-B1 and F1-B2. It is striking that the species had not yet been found in the MWTL sample points before 2002 and is always found in low numbers (a total of 9 specimens over all years).

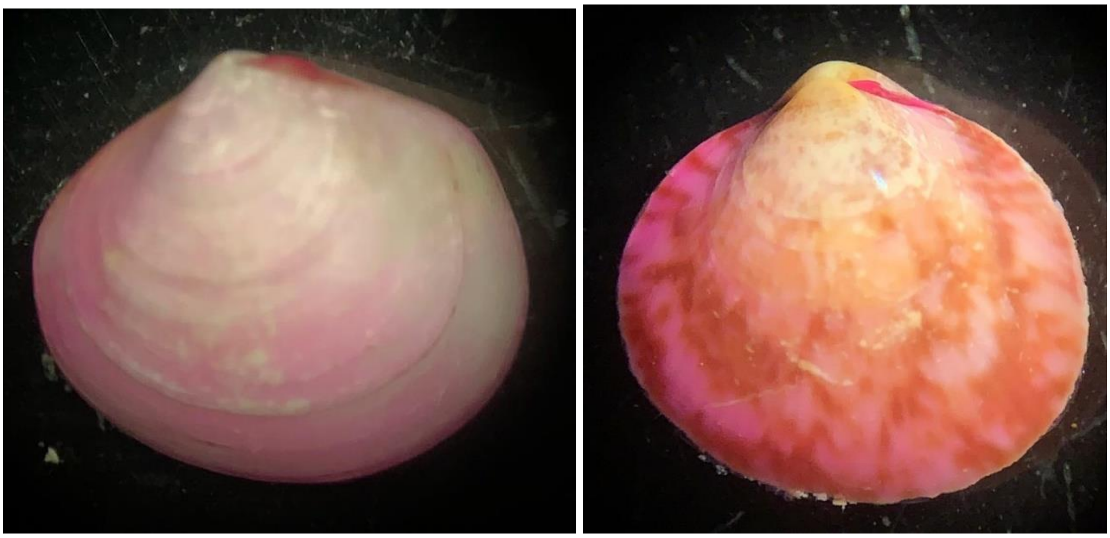


Figure 4.12: (left) *Diplodonta rotundata*, a shellfish that is slowly being observed more in the Dutch North Sea, stained with rose bengal. (right) *Glycymeris glycymeris*, a species that is very rare in the Netherlands, stained with rose bengal.

In sample F1-B2 one specimen of *Glycymeris glycymeris* (Figure 4.12, right) was found. This eye-catching coloured shell is a typical NE Atlantic species, but very rare in the Netherlands. Only a few observations are known, none of which come from the MWTL survey. The species is extremely variable in shape and colour. This can make it difficult to identify the species correctly [29].

Large species

Most macrofauna sampling on the North Sea takes place in spring (March) and not in late September. This means that the chance of finding certain species is different. Also the extent to which a certain species has grown can vary. Possibly because of this or because of the habitat is different some species were much larger than they had been found so far in the Netherlands. Mainly in the family Spionidae. Two examples are the worms *Spio symphyta* and *Dipolydora sp. B*. Small specimens of *Spio symphyta* are often found, hampering identification. Important identification characteristics, such as the number of teeth on the chaetae, change as they grow larger (1 tooth eventually disappears). As a result, the current identifications of *Spio symphyta* and *Spio decoratus* was difficult.

Dipolydora sp. B (Figure 4.13) is probably found in low numbers with some regularity in the macrofauna samples around the Brown Bank. This species is usually determined by genus and in a few cases as *Dipolydora saintjosephi*. With small specimens one classifies it as *Dipolydora saintjosephi,* while some important characteristics, such as the number of chaetae, change as it grows larger. The exact name, distribution and its role in the ecosystem, is currently unknown. Several specimens have been sent to an expert for further analysis.

Figure 4.13: *Dipolydora sp. B*, stained with rose Bengal.

5 Discussion

5.1 Mapping *Sabellaria*

The development of the classification model and its performance assessment described in this report is as good as can be done using the limited available training and test data. Even though special attention was paid to the selection of a well-balanced training and test dataset, the dependency on ROV video data for reliable ground truthing does not allow the manual selection of high confidence training data.

The classification model was able to map elevated reefs (type B) but did not enable a distinction between the coverage density (low < 20%, medium 20 to 50%, high >50%). Furthermore, the SSS data was not sensitive enough to be able to detect flat *Sabellaria* and consequently the classification model considered flat *Sabellaria (type C)* as no reef. The MBES backscatter data was not able detect either the flat or the elevated *Sabellaria*. Solely to some extent the elevated *Sabellaria* was visible in the MBES bathymetry but in this project no methodology was developed enabling the effective usage of the MBES bathymetry to automatically detect the elevated *Sabellaria*. A combination of using the SSS and MBES data simultaneous could not be established because of the poor positioning of the SSS data (primarily during the first week of the trial due to a missing USBL sensor on the tow fish) .

The calculated reef coverage for areas A-B-F-G (3.2-4.6%) are influenced by choice of area boundary. Depending on survey location within the swale between the tidal ridges, larger or smaller areas of high or low reef percentage would be surveyed which influences the overall reef coverage in percent. Therefore, the interpretation of the estimated reef coverage should be made under the light of the surveyed area. For example, the top of the tidal ridges were not surveyed for *Sabellaria* because previous studies have only found *Sabellaria* in the swale ([15] [16]). Two lines were surveyed to verify existing bathymetric maps, but not to demonstrate the presence or absence of *Sabellaria*. Surveying these areas and including the data in the analysis would most likely decrease the estimated reef coverage in percent.

The developed SSS processing workflow is adaptable to other environmental areas and potentially applicable to classify other reef types, habitats or objects (if SSS has sufficient sensitivity). Clearly, training and testing new classifiers with sample datasets from such new environments and SSS configurations are indispensable for validated predictions. This was also apparent in this project, where the majority of the training and testing samples were located in area B leading to a higher classification accuracy compared to the other areas.

Furthermore the acoustic shadow of the megaripples are more present on the north side of the sonar track and become larger towards the outer range. This is caused by the orientation of the sonar w.r.t. the megaripples. The acoustic shadow appears in the deepest part of the troughs of the megaripples, where the video recordings have indicated the presence of elevated *Sabellaria*. Since the acoustic shadows are the blind zones of the sonar, these areas are not captured by the classification and consequently in the computation of the reef coverage. A higher flying height of the tow fish would reduce the acoustic shadow but also decrease the contrast of the SSS. Sailing perpendicular to the megaripple direction (i.e. north-south orientated track lines) would reduce the blind zone but results in lower quality acoustic images of the reefs [4].

It was also observed that elevated reefs could not be successfully detected close to the nadir. This is most likely due to the dependency on a sufficiently high enough incident angle causing the blotchy pattern of high and low reflectivity indicative for the presence of *Sabellaria*. The insensitivity of the SSS data to detect *Sabellaria* in the nadir region was taken into account in the post-processing of the classified tracks. In future measurements campaigns, this would ideally already be considered in the data acquisition and pre-processing steps by planning narrower SSS lines .

The manual winch operating mode and the missingUSBL locator during the first week in combination with a sailing direction perpendicular to the main tidal current resulted in a SSS data uncertainty up to ~30m. Because of this uncertainty and the low sensitivity of the SSS at short ground ranges (this data was removed from the final map production), the final *Sabellaria* presence maps has only a resolution of 100x100m in order to merge separate SSS transects into a full-coverage map. The coarse map resolution limits the possibility to compare the identified *Sabellaria* hotspots with other data layers (e.g. morphological details such as the slope derived from the MBES bathymetry). Since the actual resolution of the SSS images is already very high with 10x10cm, improvements on the positioning are more important than on the acoustic sensor.

The second week SSS dataset, which was acquired using a USBL system, has a significantly higher positioning accuracy with an estimated maximum uncertainty of 10 m. The positioning accuracy was assessed manually by comparing the positioning offset between the same feature visible in the MBES and SSS data. The MBES data has an accuracy of a few centimetres and was considered as the ground truth. The high positioning error occurring in the first week resulted from the estimation of the tow fish position based on the measured cable length. Firstly, the cable bends while towing the tow fish and therefore only a rough estimate of the actual distance between vessel and tow fish can be provided. In addition, currents, in particular when sailing perpendicular to the current, lead to an offset towards starboard or port side which cannot be corrected for. With a USBL locator (second week), the position of the tow fish (w.r.t. the vessel) can be measured with an accuracy of 1 to 2 meters under good conditions. Beside the uncertainty in the positioning of the tow fish, several additional factors (e.g. heading sensor uncertainty, sound speed variation, deviation from the flat seabed) contribute to the total positioning uncertainty of the georeferenced sonar image. Since all factors vary while sailing, the positioning uncertainty varies though out the datasets.

5.2 Habitat analysis

The majority of the data comparison cannot prove a statistical relationship between *Sabellaria spinulosa* and a preference towards a certain sediment type. The boxcore samples were not taken randomly, but either aimed at finding *Sabellaria* or to ground truth the MBES/SSS data. In order to demonstrate whether the presence of *Sabellaria* does influence the distribution of silt, clay and sand in the top and bottom layer, more samples should be taken proportionally per strata. *Sabellaria* which form crusts (flat) or are elevated were found in a variety of sediment types. It is hypothesized that *Sabellaria* forms crusts due to the sediment type and this should be seen as a separate form of a *Sabellaria* reef.

Only the comparison between the MBES-based sediment map and the video recordings reveals a tendency of the elevated *Sabellaria* towards finer sand. However, this sediment map is only derived from MBES backscatter using the relationship between backscatter and sediment composition described in literature. This relationship is based on modelling and experiments showing that in general backscatter increases with increasing coarseness and hardness of the sediment ([24] [30] [25]) as for example from mud to sand and gravel or from fine to medium and coarse sand. Therefore, the uncertainty is higher compared to a analysis of the sediment samples and drawing only conlsuion on this observation should be considered with caution.

The statistical analysis of the data cannot prove a preference of *Sabellaria* towards certain morphological features. However, the visual inspection of the video recordings and SSS results has shown that the elevated *Sabellaria* predominantly occurs in the troughs towards the gentle slopes (north east of the megaripples, see Figure 2.5). Most likely the statistical analysis for the elevated *Sabellaria* is hampered by the low positioning accuracy of the SSS data. Also, it was investigated if the BPI (Bathymetric Positioning Index, for explanation see Section 3.2) metric could be used as an indicator for the troughs and therefore be used for the identification of suitable elevated *Sabellaria* hotspots. The results however showed that the deployed method (using the median of the two different BPIs) did not give satisfactory performance. Here it should be noted that the success of the BPI application depends on the degree of correspondence between the lateral scales of the BPIs and of the bedforms. This method requires further development such as testing different BPI settings and better positioning of the SSS data to enable its use.

The higher navigational accuracy of the second week SSS data (due to the presence of the USBL) would allow for a more elaborate comparison against the MBES backscatter, sediment composition and morphology. Due to budget constraints, this detailed analysis could not be realized within the current project. It may however be worthwhile to further investigate the correlation between the MBES derived parameters and SSS based reef classification. To further improve the navigational estimates it is also recommended to (i) use the smoothed tow fish heading estimates and (ii) correct for the bathymetry in the ground projection. For comparison at the resolution of the megaripples (~meters) it may also be needed to apply feature matching for individual transects between the MBES and SSS to correct for remaining navigation and sound propagation errors. For the final maps, it needs to be noted that reducing the grid size with the current data set will induce significant data gaps.

It should be noted that the results of the habitat analysis are only valid within the certainty of the underlying data and the environmental area under investigation. Uncertainties are related to the positioning uncertainty of all measurement devices but in particular the SSS, interpretation of the ROV video footage, spatially limited representation of the seabed by a boxcorer, ambiguity of the acoustic imaging (only an indirect measure) and its classification. Extrapolation of the results to different areas should be considered with caution.

The Brown Bank is an intensively fished area [15] and elevated *Sabellaria Spinulosa* is assumed to be vulnerable for physical disturbances ([3] [31]). Therefore, the identification of relatively large areas populated with elevated *Sabellaria Spinolusa* is remarkable (see study by van der Rijden et al. [15]). Similar to this project, they also reported elevated *Sabellaria Spinulosa* in the troughs of the megaripples. They hypothesized that the troughs provide shelter for *Sabellaria* from the fishing gear, as it jumps from megaripple crest to crest, which could enable the formation and persistence of these reefs despite being under fishing pressure. It should be noted, that the relation between bottom trawling and the presence of *Sabellaria Spinulosa* was not a central topic of this project. Nevertheless the SSS and MBES backscatter images, which are very suitable sensors to identify trawl marks ([32] [33]), were manually scanned for signs of bottom trawling. Only very few trawl marks were observed in the surveyed area. Bruns et al. [32] estimated the persistence of trawl marks to

2-7 days in the German Exclusive Economic Zone (North Sea) and up to 5 months in the Dogger bank (North Sea). Since the Brown Bank is reported as an intensively fished area, the persistence of trawling marks is expected to be low.

6 Conclusions

6.1 Mapping *Sabellaria*

The multi-scale and multi-sensor approach for detecting and mapping *Sabellaria spinulosa* has been successful. The mulitbeam echosounder (MBES) was crucial to provide a detailed morphological overview over the study area for selecting locations for ground truthing the seabed during the trial. Furthermore, it was used to manually geolocate the poorly georeferenced SSS data in order to select training and testing samples for the AI classifier. In post-processing, the MBES bathymetry and backscatter data were employed for the habitat analysis. The side-scan sonar (SSS) has been proven as the sensor of choice for detecting elevated *Sabellaria spinulosa* reefs during the trial and as the input for the classification method to automatically map elevated *Sabellaria*. The ROV video recordings were crucial for ground truthing the undisturbed seabed on a slightly larger scale and with better lateral coverage than the boxcore samples. They were used to validate the acoustic patterns visible in the SSS images during the trial and were successfully employed to train and test the classification model in post-processing. The boxcore samples provided a detailed information about the sediment grain size and allowed a biodiversity species analysis of the *Sabellaria* reefs.

The chosen classification model, called support vector machine (SVM), was trained and tested on data from area A, B, F and G to detect elevated *Sabellaria*. In the video recordings elevated *Sabellaria* was identified as individual clumps with an elevation of more than 5 cm up to a high seabed density of more than 50% with an elevation of 20 cm. The precision (true positive prediction rate) and sensitivity (true positive rate) of the *Sabellaria* reef detection algorithm has been assessed to 0.518 (51.8%) and 0.491 (49.1%), respectively. The predicted percentages of *Sabellaria* coverage are reduced by the precision of 0.518 to correct for this uncertainty. This approach resulted in reef coverage percentages estimated at 4.6% (Area A), 4.1% (Area B), 3.5% (Area F) and 3.2% (Area G). In addition to the automized validation of the classification results, a manual inspection of the *Sabellaria* areas showed that the classification model generally predicted reefs at logical locations with SSS patterns similar to the locations where *Sabellaria* was found during the monitoring campaign (typical blotchy patterns dissimilar from ones caused by megaripples). Results for area E were not included in this report. The SSS data in Area E, located in the south-west of the IJmuiden Ver area, was acquired with a different track orientation. The limited number of training samples in this area did not allow a proper training of the classifier towards the different track orientation.

The reef-presence maps in area A, B, F and G show large-scale spatial patterns of *Sabellaria spinulosa* reefs, indicating a preference of elevated *Sabellaria spinulosa* for settlement to the east of the deepest part of the swale between the tidal ridges (i.e. valley between the sand bank). A fine-scale visual inspection of the classification maps, confirmed by the ROV video recordings, indicates the preference of elevated *Sabellaria* for settlement in the troughs towards the stoss side (i.e. side with the more gentle slope) of the finer-scale megaripples as well. A preference of *Sabellaria* for certain sediment compositions could not been proven via the various acquired datasets and analysis methodologies.

6.2 Biodiversity

The video recordings have revealed that*Sabellaria spinulosa* forms reef with an elevations of up to 20 cm. Crusts of Sabellaria (also described as flat *Sabellaria*) are difficult to observe on the video recordings, MBES and SSS. It is expected from the boxcorer samples and the video recordings that this type of reef is more common than the elevated Sabellaria in the Brown Bank region.

Many more species have been observed in very high abundances in the Brown Bank than previously known. Many species are new or rare in the Netherlands or were previously only known from the Cleaver Bank and/or Oestergronden. Approximately 80% of the taxa which were found in the *Sabellaria* samples were not found before in the Brown Bank area. On average, the samples were also richer in species than the Cleaver Bank samples taken for the MWTL(Monitoring van de Waterstaatkundige Toestand des Lands) programm.

A total of three reef types were observed in the box corer samples; some loose *Sabellaria spinulosa* clumps (type A), reef-structures with elevated *Sabellaria spinulosa* (type B) and reef-structures of flat *Sabellaria spinulosa* in the sediment (type C). Type A samples consisted of loose clumps of *Sabellaria* tubes. Most tubes did not contain *Sabellaria*. Nevertheless, these samples were more species-rich than the MWTL samples from the Brown Bank taken in 2021. Elevated *Sabellaria* is designated as type B. These samples contained many species associated with a *Sabellaria* reef. The reefiness score was medium to high. Most species were found in the Type C samples. Many of the species were previously rare or unknown in the Netherlands. It is hypothesized that this type of reef is created due to the type of sediment and should be considered as a separate type of *Sabellaria* reef. These results show that *Sabellaria* reefs sustain a unique ecosystem with high environmental value.

6.2.1 Lessons learned and recommendations

A USBL sensor is crucial for accurate positioning of the SSS image and should be attached to the tow fish at every survey. In addition, the smoothed fish heading should used instead of the ship heading for improved ground projection of the sonar data. An accurate positioning accelerates and simplifies the identification of suitable ground truth locations for the presence of reefs on board of the vessel. During the classification procedure more accurate positioning simplifies the selection of training and testing samples and could replace the hybrid approach of manually selecting the sample dataset. Furthermore, it would allow to create a final *Sabellaria* presence map with a higher resolution. This would be beneficially for the habitat analysis. Finally, it would improve the use of the SSS data in combination with the MBES data as input layers in the classifier, thereby most likely improving the classification accuracy.

The along-track resolution of the SSS data (~ 30 to 45 cm) was relatively poor, in particular compared with the relatively high across-track (~3 cm) resolution. A higher across-track resolution would allow to acoustically ensonify the reefs, which consisted mostly of individual patches, more densely. This would lead to a sharper image with reef patterns being easier to distinguish for the classifier. A higher along-track resolution could be achieved with a slower vessel speed. In this trial the vessel speed of the ARCA varied between 4.5 and 7 kn depending on the orientation of the vessel towards current and waves. However, achieving a lower speed with the ARCA in open water and still keeping a straight track line might be challenging.

With the current data acquisition setup, a balance will need to be made between acquiring more sonar (to improve coverage) or ROV Video data (to improve confidence in final maps). The SSS processing on board of the vessel was key to the successful detection of the *Sabellaria* reefs. For future campaigns it may also be considered to train and test the AI model on board of the vessel with the incoming ground truth data to further optimize (i) the search for reefs and (ii) the acquisition of high quality training and testing data.

For future measurement campaigns also the deployment of an Autonomous Underwater Vehicle (AUV) able of acquiring both high-resolution SSS and ground truth video data may provide value. Adding further autonomous capabilities to the AUV may even enable the AUV to search for and identify *Sabellaria* reefs in areas with a complex morphology.

Furthermore, it is advisable to compare the sediment data from the boxcorer with the presence/absence of the different Sabellaria reef types. In combination with the SSS and MBES it would facilitate to estimate where the different reef types may occur in the area. To determine the effect of the different *Sabellaria* reefs on the sediment composition, a different sampling strategy is required. For this, it is advisable to take a proportional amount of samples per strata.

References

[1] L.M. Van Son, B. Dzon and L. Leeuwis, “Macrozooebentos monitoring in de Rijkswateren met Boxcorer, MWTL 2021 - Noordzee. RWS: BM 22.32,” Eurofins Aquasense, Amsterdam, 2022.

[2] J. Asjes, H. Merkus, O. Bos, J. Steenbergen, S. Stuijfzand, I. van Splunder, T. Kooten, S. Rivero and G. A. J. Vis, “Monitoring en Onderzoek Naturversterking en Soortenbescherming (MONS),” 2021.

[3] OSPAR_Commision, “Background document on Sabellaria spinulosa reefs. Biodiversity Series (614/2013),” OSPAR, 2013.

[4] T. Gaida, B. Binnerts, C. van der Stappen and J. Cuperus, “Sabellaria monitoring trial: Summary report (id56-sabellaria-monitoring-trial-summary-report),” TNO, 2023.

[5] “Data managment factsheet: Monitoring Onderzoek Natuurversterking Soortenbescherming,” Witteveen & Bos, 2024.

[6] JNCC, “UK Biodiversity Action Plan; Priority Habitat Descriptions,” BRIG (ed. Ant Maddock), 2008.

[7] A. Jackson and K. Hiscock, “Sabellaria Spinulosa Ross worm. In Tyler-Walters H. and Hiscock K. Marine Life Information: Biology and Sensitivity Key Information,” 2008. [Online]. Available: https://www.marlin.ac.uk/species/detail/1133. [Accessed 2 7 2024].

[8] L.A. Jones, K. Hiscock and D.W. Connor, “Marine habitat reviews. A summary of ecological requirements and sensitivity characteristics for the conservation and management of marine SACs,” Petersborough, Joint Nature Conservation Committee. (UK Marine SACs Project report), 2000.

[9] OSPAR, “OSPAR List of Threatened and/or Declining Species and Habitats. Reference Number: 2008-6,” 2008.

[10] “Min I&W, Min LNV. Marine Strategy (part 1). Update of current environmental status, good environmental status, environmental targets and indicators. 2018-2024 (Issue part 1).,” 2018.

[11] P.M.J. Herman and F.F. van Rees, “Mapping Reef forming North Sea Species,” 2022.

[12] O. G. Bos, J.W.P. Coolen and J.T. van der Wal, “Biogene riffen in de Noordzee. Actuele en potentiële verspreiding van rifvormende schelpdieren en wormen (C058/19),” Wageningen Marine Research, 2019.

[13] H. Sas, L. van Duren, P. Herman, T. van der Have, P. Kamermans, O. Bos, E. Kingma, T. Bouma and E. Kardinaal, “Reef-building species and biogenic reef enhancement in the Dutch North Sea. Background documents,” 2023.

[14] ECOWENDE, “ECOWENDE,” [Online]. Available: https://ecowende.nl/en/) . [Accessed July 2024].

[15] K.J. van der Reijden, L. Koop, S. O'Flynn, S. Garcia, O. Bos, C. van Sluis, D.J. Maaholm, H.P.M.J., D.G. Simons, H. Olff, T. Ysbaert, M. Snellen, L.I. Govers, A.D. Rijnsdorp and R. Aguilar, “Discovery of Sabellaria spinulosa reefs in an intensively fished area of the Dutch Contintental Shelf,” *Journal of Sea Research,* vol. 144, pp. 85-94, 2019.

[16] K.J. van der Reijden, L. Koop, S. Mestdagh, M. Snellen, P.M.J. Herman, H. Olf and L.L. Govers, “Conservation Implications of Sabellaria spinulosa Reef Patches in a Dynamic Sandy-Bottom Environment,” *Frontiers in Marine Science,* vol. 8, 2021.

[17] E.G.R. Bakker, J. de Jong, R. Middelveld and D.B. Kruijt, “Macrozoöbenthos bemonstering Noordzee met de Bodemschaaf. Rapportage 2021 en 2022. Rapport22-0306,” 2022.

[18] L. van Walraven, R. Jak, D. van Oevelen, M. Keur, M. Polling, N. Peck and J. Hoekendijk, “Noordzeeloket,” 2023. [Online]. Available: https://www.noordzeeloket.nl/publish/pages/228846/2-zooplankton-wozep-mons-kennisdag.pdf.

[19] S. Gubbay, “Defining and managing Sabellaria spinulosa reefs: Report of an inter-agency workshop 1-2 May, 2007.,” JNCC Report No. 405., Peterborough, Joint Nature Conservation Committtee (JNCC)., 2007.

[20] V. Hendrick and R. Foster-Smith, “ Sabellaria spinulosa reef: a scoring system for evaluating 'reefiness' in the context of the Habitats Directive,” *Journal of the Marine Biological Association of the United Kingdom,* vol. 86, no. 4, pp. 665-677, 2006.

[21] F. Bianconi, R. Harvey and W. Wang, “A texture analysis approach to identifying Sabellaria Spinulosa colonies in sidescan sonar imagery,” in *Irish Machine Vision and Image Processing Conference*, 2011.

[22] S. Solikin, H.M. Manik, S. Pujiyati and Susilohadi, “Support Vector Machine Classification Method for Predicting Jakarta Bay Bottom Sediment Type using Multibeam Echosounder Data,” *Science & Technology,* vol. 28, no. 2, pp. 477-491, 2020.

[23] “GIT,” [Online]. Available: https://wozep.nl/git/MONS/04_benthos_and_habitats/-/tree/main/sabellaria_reefs.

[24] APL-UW, “High-Frequency Ocean Environmental Acoustic Models Handbook,” Applied Physics Laboratory, Seattle, Washington, 1994.

[25] X. Lurton, An Introduction to Underwater Acoustics, Springer-Verlag Berlin Heidelberg, 2010.

[26] “Aquadesk,” [Online]. Available: https://live.aquadesk.nl/viewer/1. [Accessed 1 11 2024].

[27] E.J. Allen, “Pallasia murata n.sp.: a new British Sabellarian,” *Journal of the Marine Biological Association of the United Kingdom,* vol. 7, no. 2, pp. 299-304, 1904.

[28] ILVO, “Macrobenthos monitoring in function of aggregate extraction activities in the Belgian part of the North Sea,” 2016. [Online]. Available: https://www.vliz.be/en/imis?dasid=4451&doiid=199. [Accessed 2024 11 1].

[29] F. Nolf and F. Swinnen, “The Glycymerididae (Mollusca: Bivalvia) of the NE Atlantic and the Mediterranean Sea,” *Neptunea,* vol. 12, no. 3, pp. 1-35, 2012.

[30] X. Lamarche and G. Lurton, “Backscatter measurements by seafloor-mapping sonars. Guidelines and Recommendations,” 2015.

[31] N. Gibb, H. Tillin, B. Pearce and H. Tyler-Walters, “Assessing the sensitivity of Sabellaria spinulosa to pressures associated with marine activities.,” JNCC, 2014.

[32] I. Bruns, P. Holler, R.M. Capperucci, S. Papenmeier and A. Bartholomä, “Identifying Trawl Marks in North Sea Sediments,” *Geosciences,* vol. 10, no. 11, 2020.

[33] T.C. Gaida, “Onderzoek naar zeebodem in de Noordzee,” De Rijke Noordzee, 2020. [Online]. Available: https://www.derijkenoordzee.nl/nieuws/onderzoek-naar-zeebodem-in-de-noordzee.

Appendix A
Feature layers

Figure A.1 to Figure A.4 show the 20 features layers used in the simulation run 2.

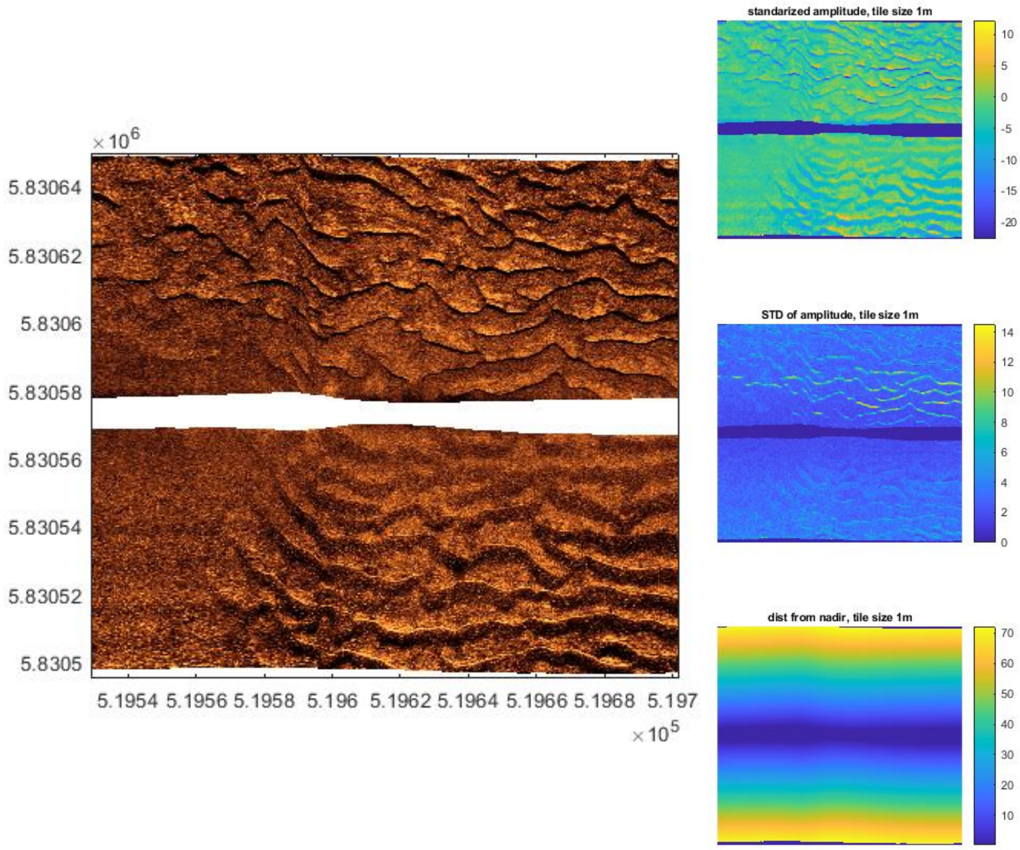


Figure A.1: (left) SSS image of track Bsub_W_2230831185800.sdf. (right) Basic statistic feature layers for a tile size of 1 x 1 m.

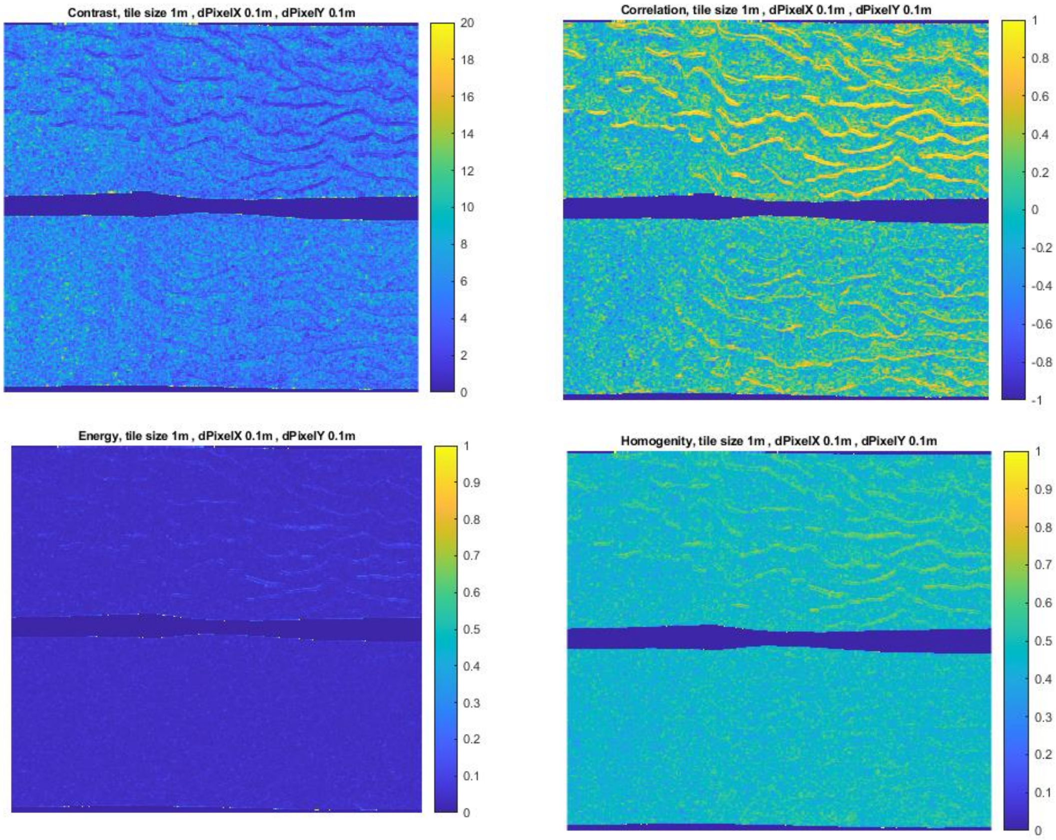


Figure A.2: GLCM matrix computed from a tile size (tile) of 1 x 1 m and pixel pairs of 0.1 m.

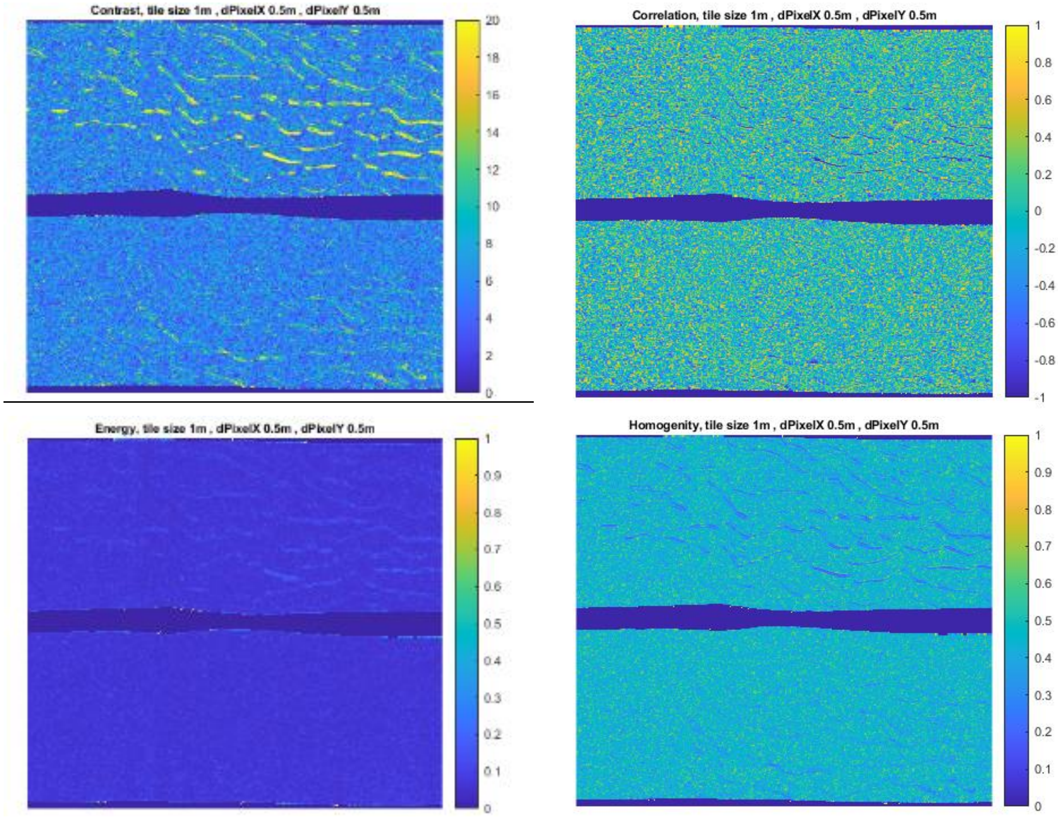


Figure A.3: GLCM matrix computed from a tile size (tile) of 1 x 1 m and pixel pairs of 0.5 m.

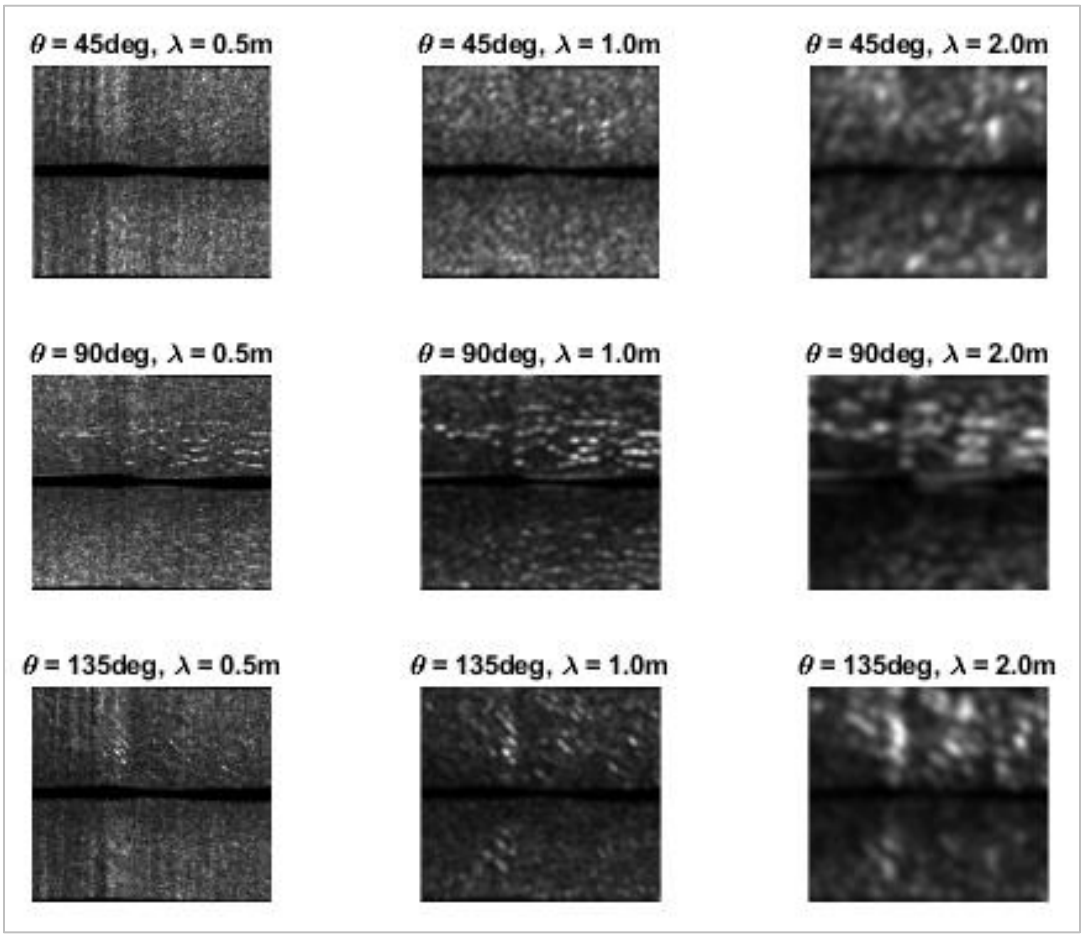


Figure A.4: Gabor filter showing different direction angles θ and wavelengths λ.

Classifier performance analysis

Variety of classifiers and feature layer combinations test

The performance of 5 classification methods for 24 feature layer combinations was tested. This resulted in 254simulation runs per classification method. The feature layers used in the 24 simulation runs are shown in Table A.1. The Precision and Sensitivity per class for all runs are shown in Figure A.5. The scores were obtained from the validation using the test sample dataset from area A, B, F and G.

In general, the SVM and neural network perform best, in particular towards the detection of the reef class evaluated by the precision being considered as the most important score (Figure A.5). Considering that a high precision and sensitivity is also desirable for the no reef and acoustic shadow classed, runs 1, 2, 4, 15, 16, 21, 22 and 23 perform almost equally well. For all runs these feature combinations are relatively similar. Considering run 2 as the default run, described in Section 2.2.1 and used for the results presented in this report, the other runs differ in the tile size for the basic statistic and GLCM features, number of wavelengths, value of wavelengths and orientations in the Gabor filter and the use of the ground range.

Table A.1: Description of features layers used in the simulation runs. *Amp* is the median SSS amplitude computed over the given tile size, *STD* is the standard deviation of the SSS amplitude computed over the given tile size and GR is the ground range or also called distance from nadir. *All* in the GLCM column means that all four GLCM layers are computed for the given tile size and pixel pairs. *All* in the Gabor column means that all layers for the given wavelengths and angles are computed while PC means that the principal components were computed from the given wavelengths and angles.

Run	Basic statistic	GLCM	Tile size (m)	Pixel pair (m)	Gabor	Wavelength (m)	Angles (deg)	PC	Feature count
1	Amp, STD, GR	all	1, 2.5	0.1, 0.5	all	0.5, 1, 2	45, 90, 135	-	31
2	Amp, STD, GR	all	1	0.1, 0.5	all	0.5, 1, 2	45, 90, 135	-	20
3	Amp, STD, GR	all	2.5		all	0.5, 1, 2	45, 90, 135	-	20
4	Amp, STD, GR	all	1		all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	35
5	Amp	-	1	-	-	-	-	-	1
6	Amp, GR	-	1	-	-	-	-	-	2
7	Amp, STD, GR	-	1	-	-	-	-	-	3
8	-	all	1	0.1, 0.5	-	-	-	-	8
9	GR	all	1	0.1, 0.5	-	-	-	-	9
10	-	-	-	-	all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	24
11	-	-	-	-	all	0.8, 1, 1.5, 2	45, 60, 90, 135	-	16
12	-	-	-	-	all	0.8, 1, 1.5, 2	45, 90, 135	-	12
13	-	-	-	-	PC 1 & 2	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	2
14	-	-	-	-	PC 1, 2 & 3	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	3
15	Amp, STD	all	1	0.1, 0.5	all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	34
16	Amp, STD	all	1	0.1, 0.5	all	0.5, 1, 2	45, 90, 135	-	19
17	Amp, STD	all	1	0.1, 0.5	PC 1 & 2	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	-	12
18	Amp, STD, GR	all	1, 2.5	0.1, 0.5	all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	yes	1

Run	Basic statistic	GLCM	Tile size (m)	Pixel pair (m)	Gabor	Wavelength (m)	Angles (deg)	PC	Feature count
19	Amp, STD, GR	all	1, 2.5	0.1, 0.5	all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	yes	3
20	Amp, STD, GR	all	1, 2.5	0.1, 0.5	all	0.5, 0.8, 1, 1.5, 2, 4	45, 60, 90, 135	yes	5
21	Amp, STD, GR	all	1	0.1, 0.5	all	1	45, 90, 135	-	14
22	Amp, STD, GR	all	1	0.1	all	1	45, 90, 135	-	10
23	Amp, STD, GR	all	1	0.1, 0.5	all	0.8	45, 90, 135	-	14
24	Amp, STD, GR	all	1	0.1	all	1	90	-	8

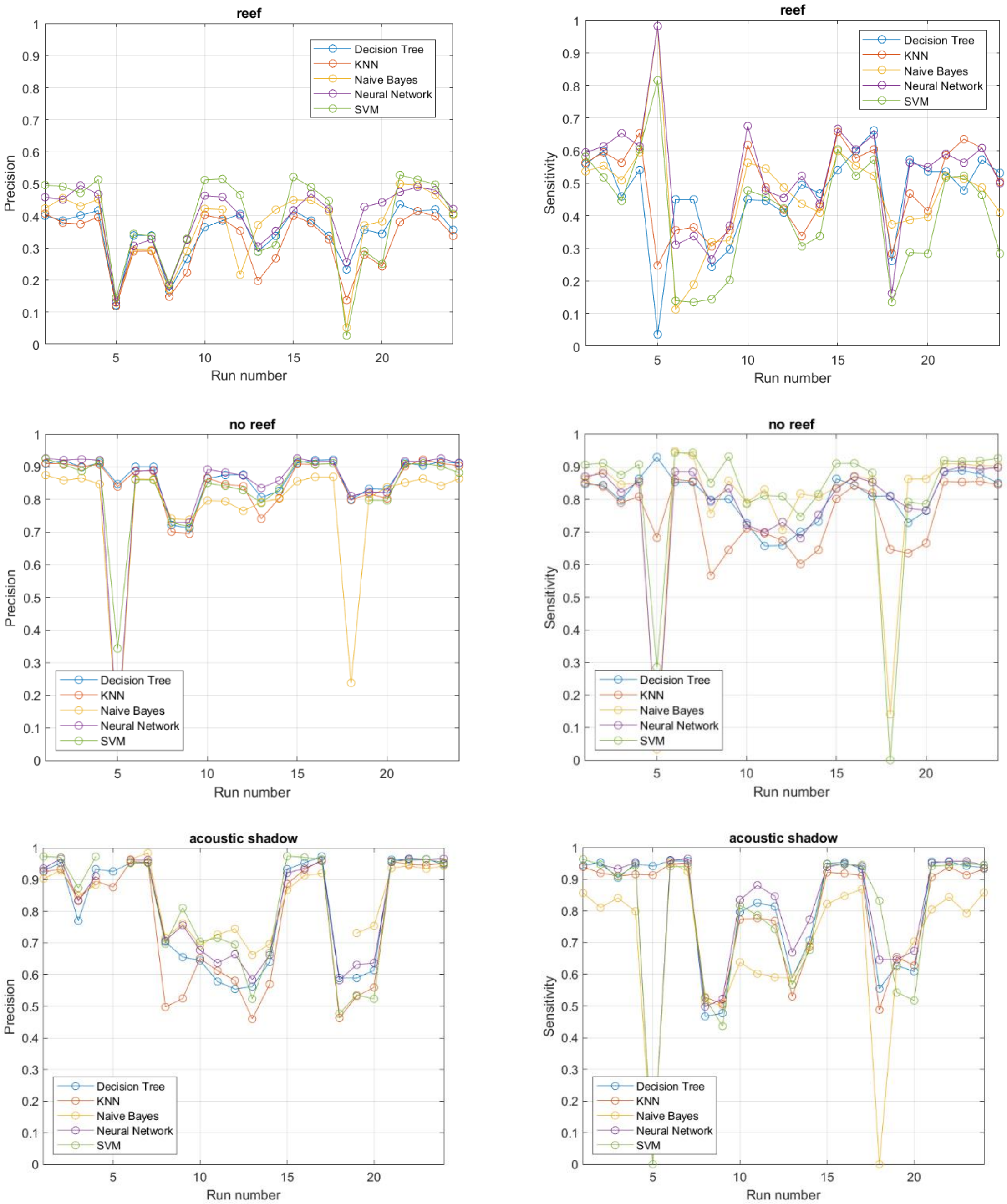


Figure A.5: Performance analysis of all classifiers and a variety of feature layer combinations. 24 different feature layer combinations were tested. Here, the test was carried out with Matlab default settings, i.e. for the SVM a linear Kernel function was used. (left) Precision and (right) Sensitivity of reef class obtained from the computation of the confusion matrix. Run 2 of SVM was selected because of (1) high Precision and Sensitivity of the reef class while also having high values for the other classes.

<u>Classifier evaluation using different classifier setting test</u>

Figure A.6 shows the confusion matrix obtained for the SVM classifier using different Kernel functions (i.e. linear, gaussian, polynomial 3th degree and 4th degree). The highest precision score for the reef class is obtained using the gaussian kernel function with 52.7% followed by the linear Kernel function with 49.1% and polynomial functions with ~45%. The sensitivity score is also highest for the gaussian function with 62.6% compared to 51.8% using the linear Kernel function. However, the option to select the gaussian function was discovered at a late stage of the project and the classification results were already obtained using the linear function. This result can be considered in the discussion that optimizing classifier settings can achieve an improvement in the classification accuracy.

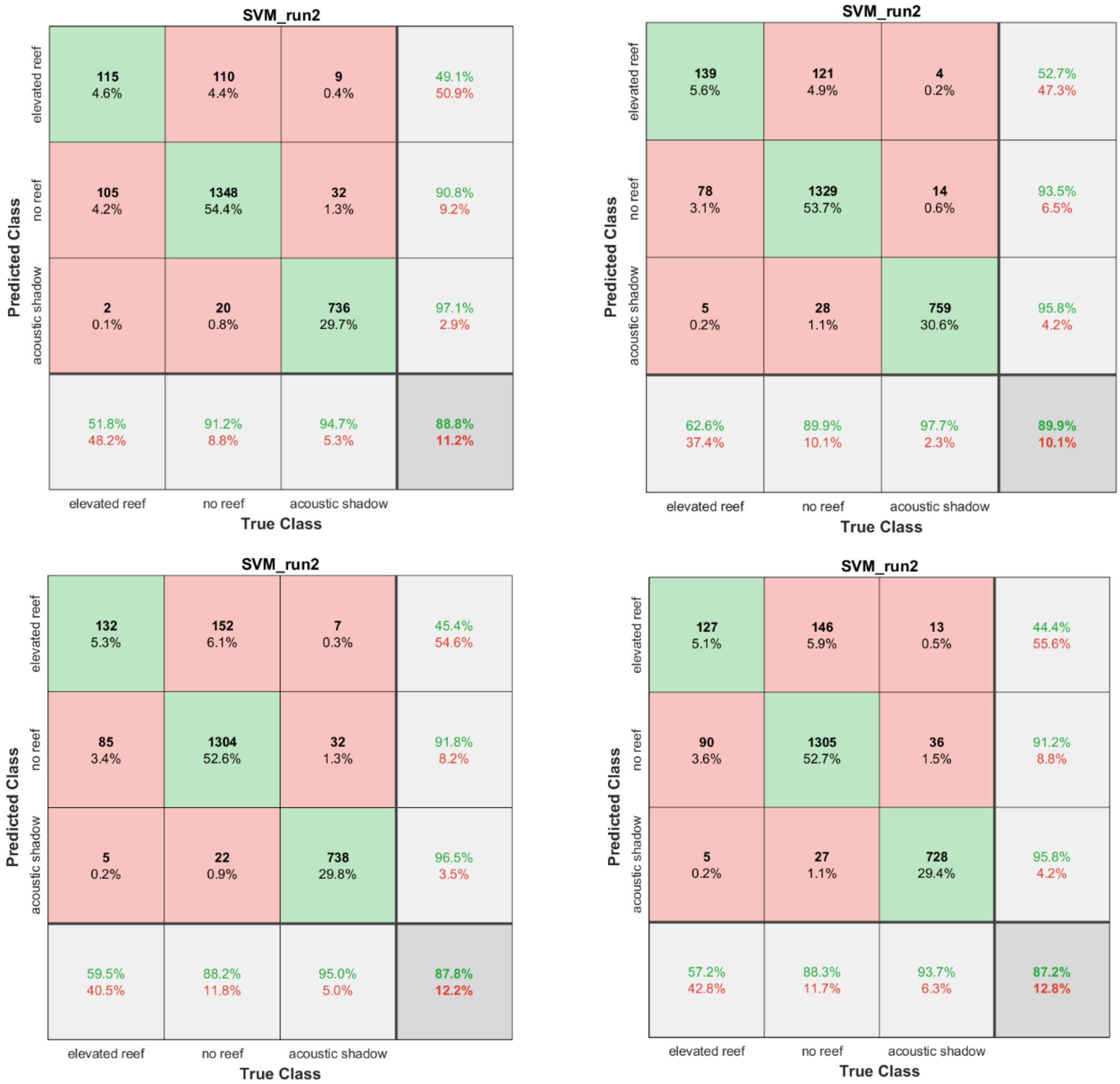


Figure A.6: Confusion matrix obtained from SVM classifiers using different Kernel functions: (top left) linear, (top right) gaussian, (bottom left) polynomial 3rd degree and (bottom right) polynomial 4th degree. Model are trained and validated on samples of area A, B, G and F.

Classifier evaluation using testing sample dataset from different areas

The confusion matrix for the trained linear SVM model can be obtained for testing samples collected in different single areas and subset combinations of areas. This allows a test of the performance of the classifier in different areas. Figure A.7 shows that the performance of the classifier towards the detection of reefs is highest in Area B with a precision of 80.4%. If testing samples from area B, F and G are included into the validation, the precision droops to 64.6 % and if area A, B, F and G are included, as shown in Section 2.2.4, the precision even drops to 49.1%. Since the majority of the training samples is located Area B, it shows that classifier performance reflects the availability of sufficient training samples and ideally the training samples are equally distributed over the study area.

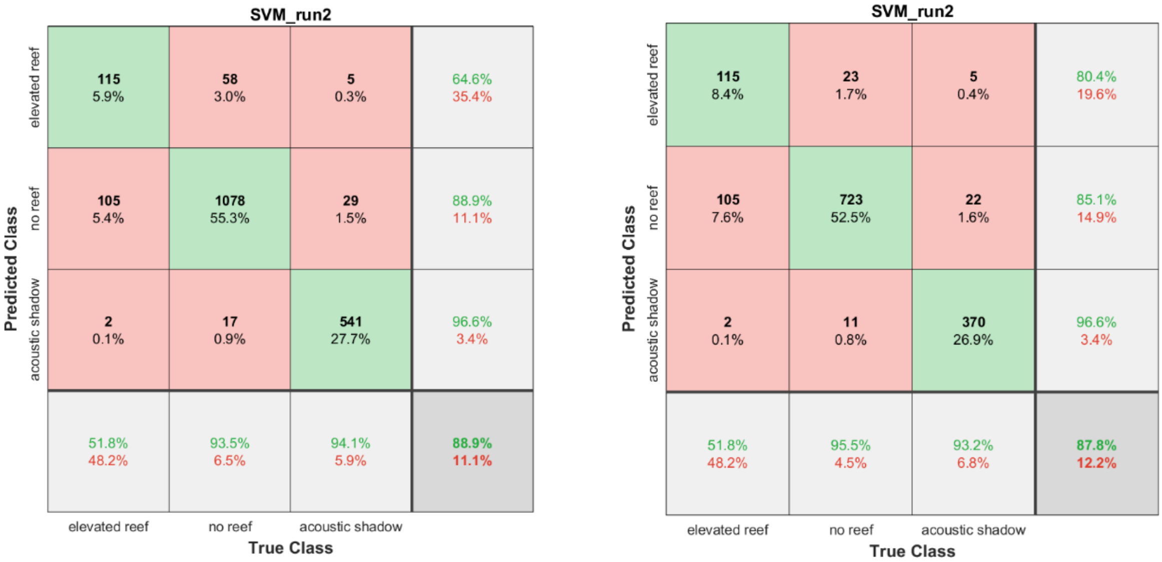


Figure A.7: Confusion matrix of trained linear SVM model obtained from testing samples from area (left) B, F and G and (right) only B.

Sensitivity analysis

A sensitivity analysis was carried out for the effect of the ground range on the classification results. Figure A.8 shows that the reef class is not equally distributed over the ground range indicating that the ground range has an effect on the reef classification. In particular, for a ground range below 30 m significantly less reef was detected. Figure A.9 shows that much less training samples at the shorter ground ranges were available. Therefore, the classifier is less well-trained for the detection of reefs. One reason why lower numbers of training and testing samples were selected at a shorter ground range is that the typical blotchy patterns indicating the reefs were not clearly visible and therefore it was more difficult to pick suitable samples. Most likely the low incident angle at this region does not allow to visualize the small-scale relief of the reefs. Since the number of testing samples for the reef class are also lower for the shorter ranges, the contribution to the confusion matrix is also lower. Based on these observations, a more representative estimation of the reef coverage is obtained by using only the classification results for a ground range from 30 to 75 m for the generation of the final reef-presence map.

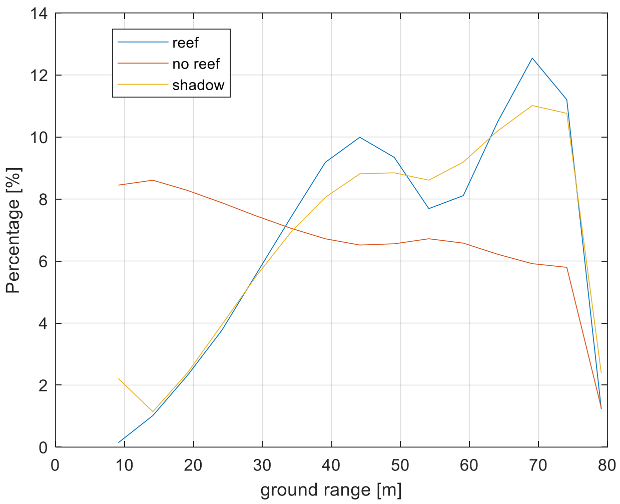


Figure A.8: Sensitivity analysis on the effect of the ground range on the classification results. The percentual distribution of a class along the ground range is shown.

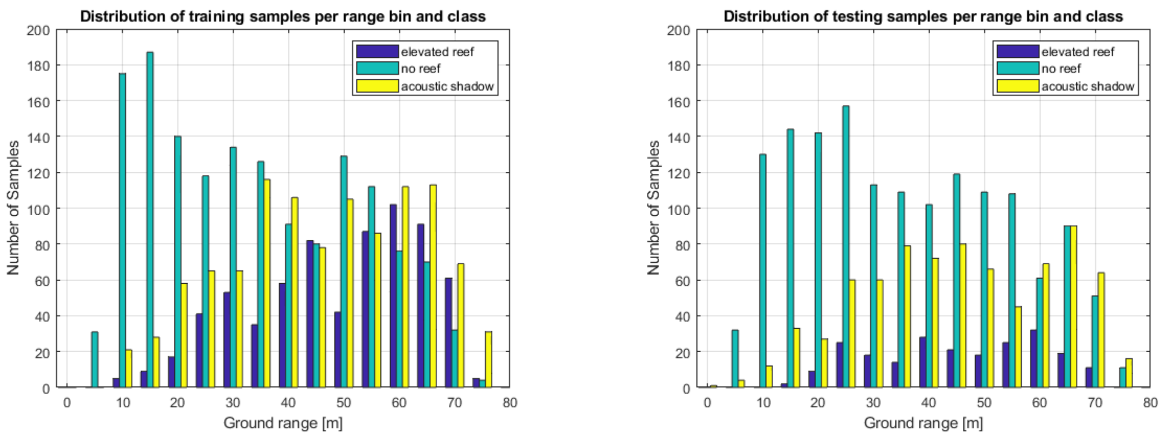


Figure A.9: Distribution of (1) training and (2) testing samples over the ground range.

Appendix B
Training and validation data

Figure B.1 shows subsets of SSS images and corresponding classification results obtained from the trained linear SVM classifier. In the SSS images the testing samples, picked by the TNO expert via the hybrid approach using expert knowledge and the video recordings, are plotted on top. In the classified images the positive (green) or false (red) detection of the given samples are plotted on top. In these examples, all testing samples of the acoustic shadow and no reef class are correctly predicted, as it was expected from the performance analysis considering the full testing dataset in Figure A.5. Approximately half of the elevated reef class samples are correctly predicted.

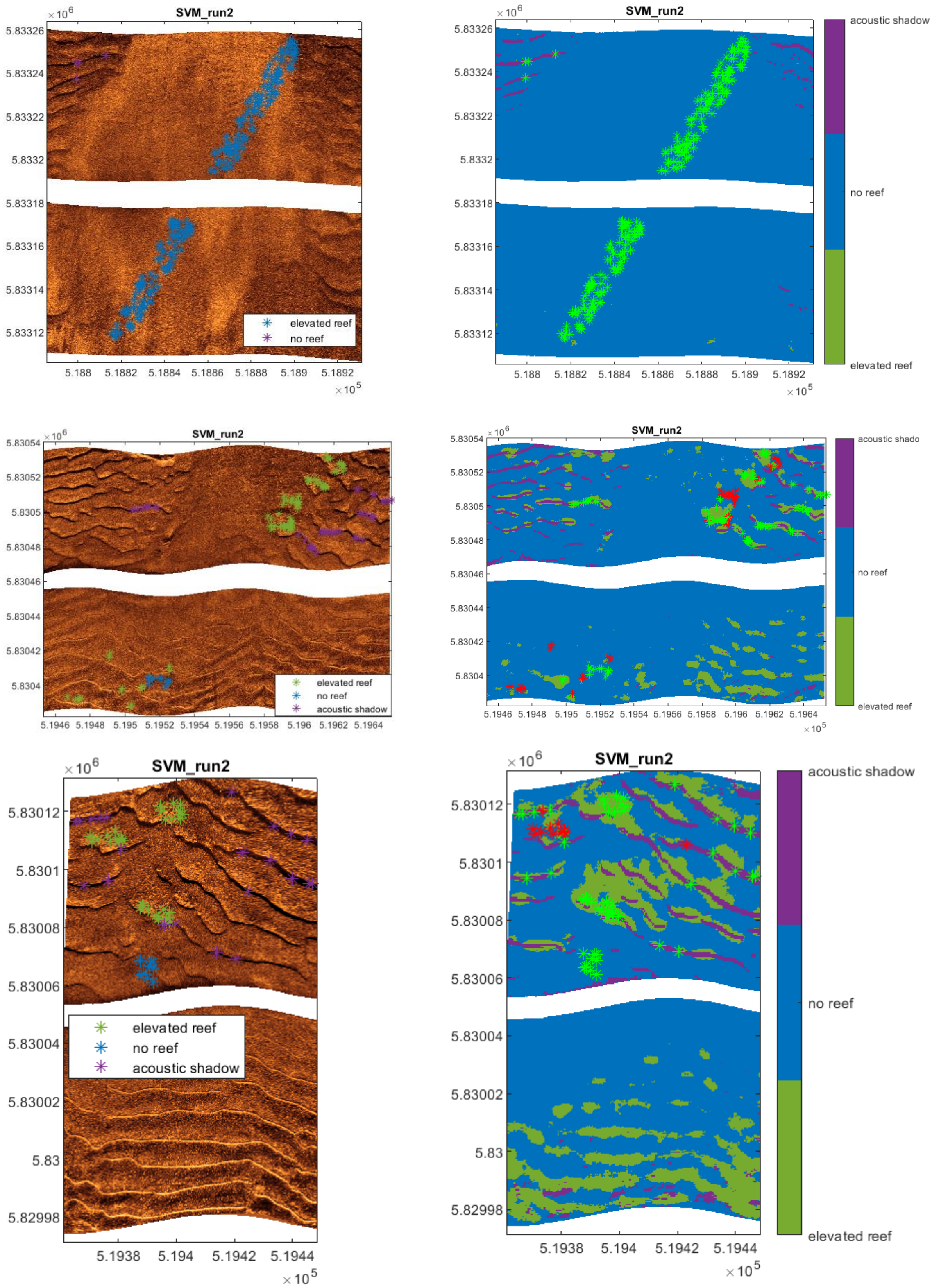


Figure B.1: Example lines for validation of trained linear SVM model using testing sample dataset. (left) SSS image with testing samples for validation indicated by asterisks. (right) Classification results with indication of positive (green asterisks) and negative (red asterisks) detections.

Appendix C
MBES-based maps

An unsupervised k-means clustering was applied to the acquired and processed MBES backscatter maps [4], for clustering into three acoustic classes. The MBES backscatter values were averaged over a grid cell of 5 x 5 m before being put into the k-means clustering. This grid cell size should be sufficiently large to account for the positioning uncertainty of the boxcorer and providing a robust backscatter value insensitive to intrinsic noise and small-scale seabed variation. An increasing class number correspond to increasing backscatter values (Figure C.1). The acoustic classes were correlated with the grain size analysis of the top layer of the boxcore samples. Some correlation of high gravel content with high acoustic class number exist but there is no correlation of the acoustic class with the mud content (Figure c.2). For acoustic class 1 no representative correlation can be obtained because it is underrepresented with only one sample. For next campaigns, the boxcore samples should be more equally-well distributed over the range of backscatter values (or acoustic classes). Since the number of samples is not well distributed over the classes and the correlation is ambiguous, a sediment map solely from the boxcore samples cannot be established. However, using the physical relationship that coarser and/or harder sediment generally corresponds to higher backscatter and finer and/or softer sediments correspond to lower backscatter values ([24], [25]), the clustered maps can be considered as a coarseness/roughness sediment map where the degree of coarseness and roughness increases from class 1 to 3. Visually, the extensive areas classified as acoustic class 3 (yellow - coarse/hard substrates) correspond with a flat bathymetry and also often with an extensive shell bed identified in the MBES data and video recordings, respectively.

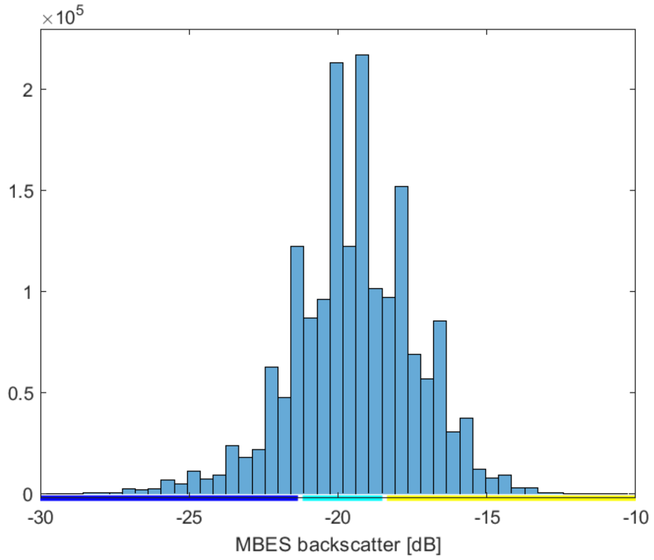


Figure C.1: Histogram of processed MBES backscatter value from area A, B, F and G. The blue, cyan, and yellow horizontal bars indicate values corresponding to acoustic class 1, 2 and 3, respectively. The boundaries are a result of the k-means clustering.

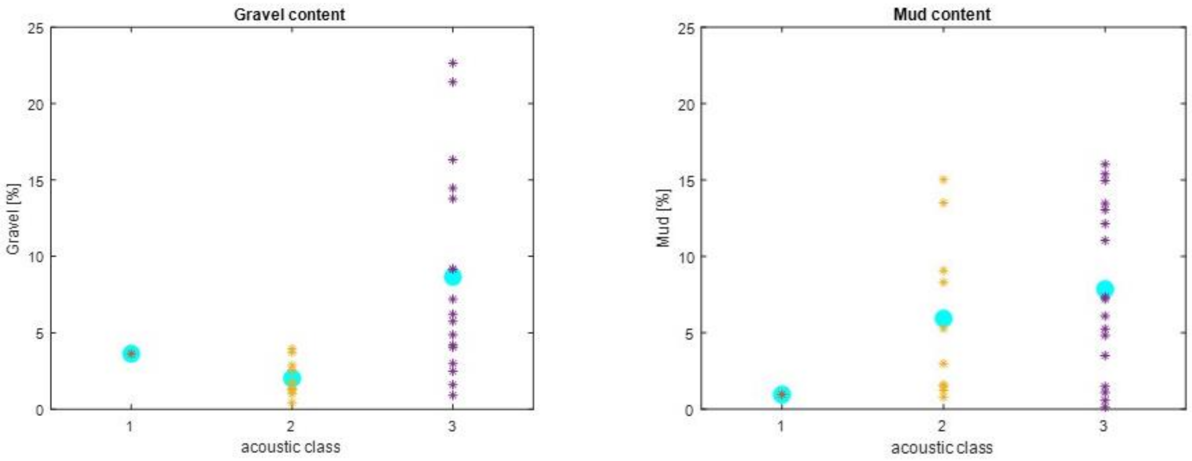


Figure C.2: Correlation between acoustic class and the (left) gravel content and (right) mud content obtained from the top layer of the boxcore samples. The acoustic class were obtained from the k-means clustering of the MBES backscatter data. Asterisks indicate values of individual samples and light blue dot the median of all samples.

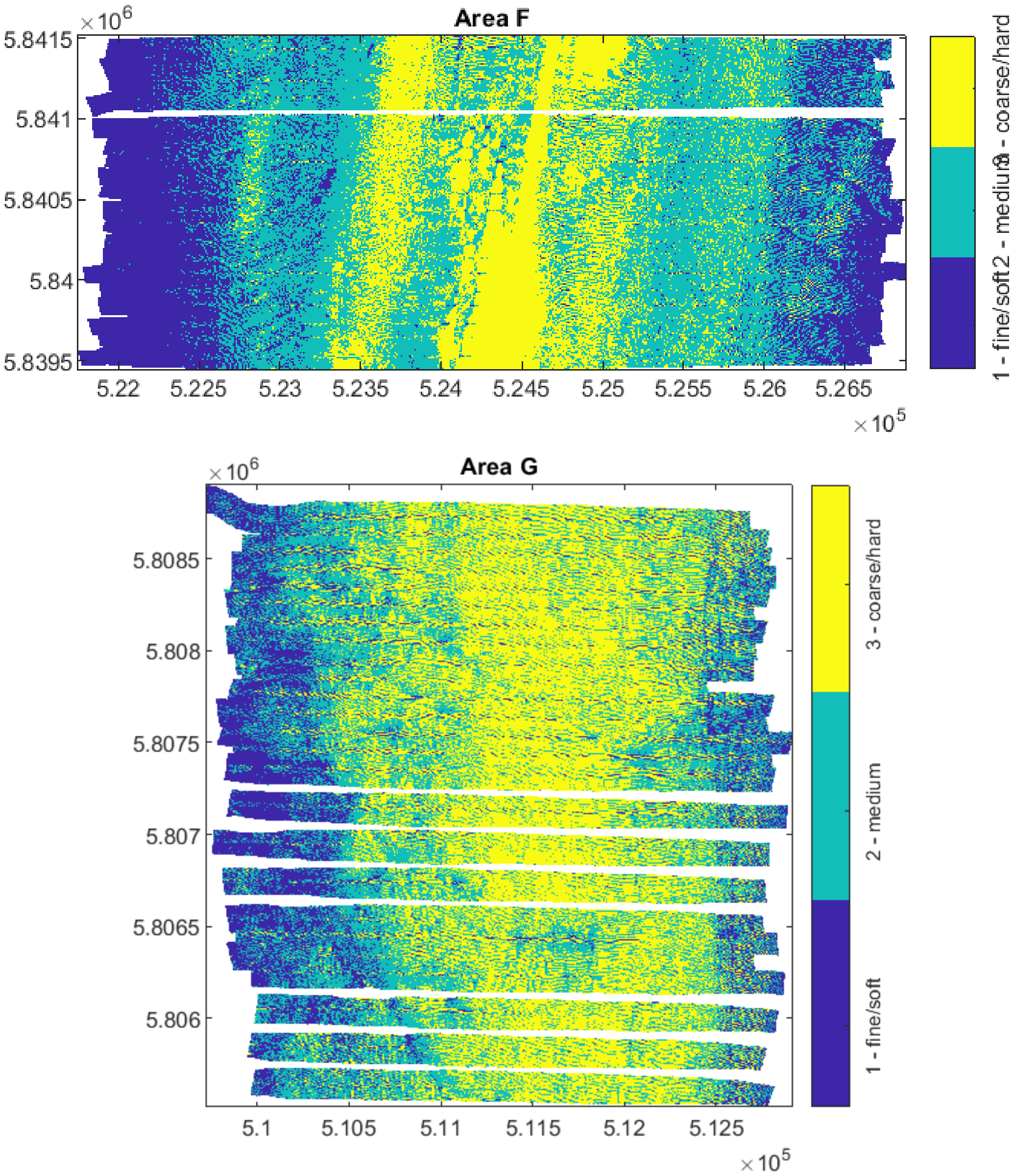


Figure C.3: Sediment coarseness/roughness map of (top) Area F and (bottom) Area G. Class 1 corresponds fine/soft sediment, Class 2 to medium coarse/hard sediment and Class 3 to coarse/hard sediment. Data gaps are related to missing MBES tracks.

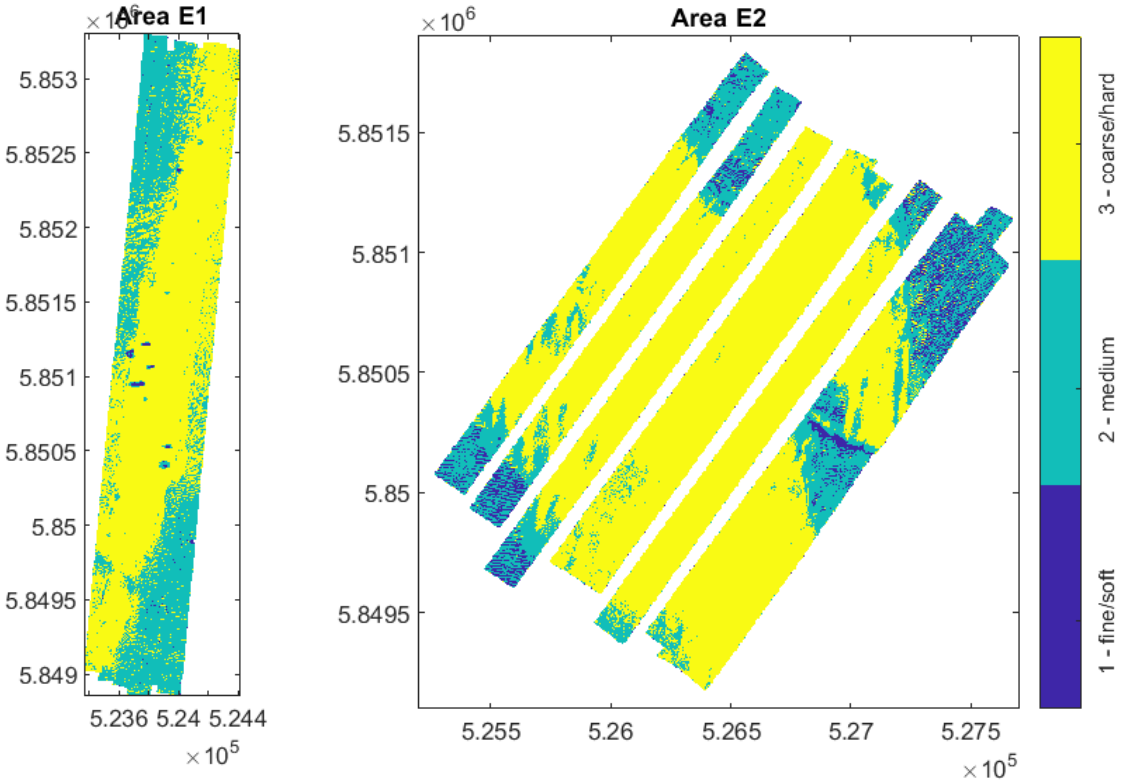


Figure C.4: Sediment coarseness/roughness map of (left) Area E1 and (right) Area E2. Class 1 corresponds fine/soft sediment, Class 2 to medium coarse/hard sediment and Class 3 to coarse/hard sediment.

Appendix D
MBES bathymetry maps

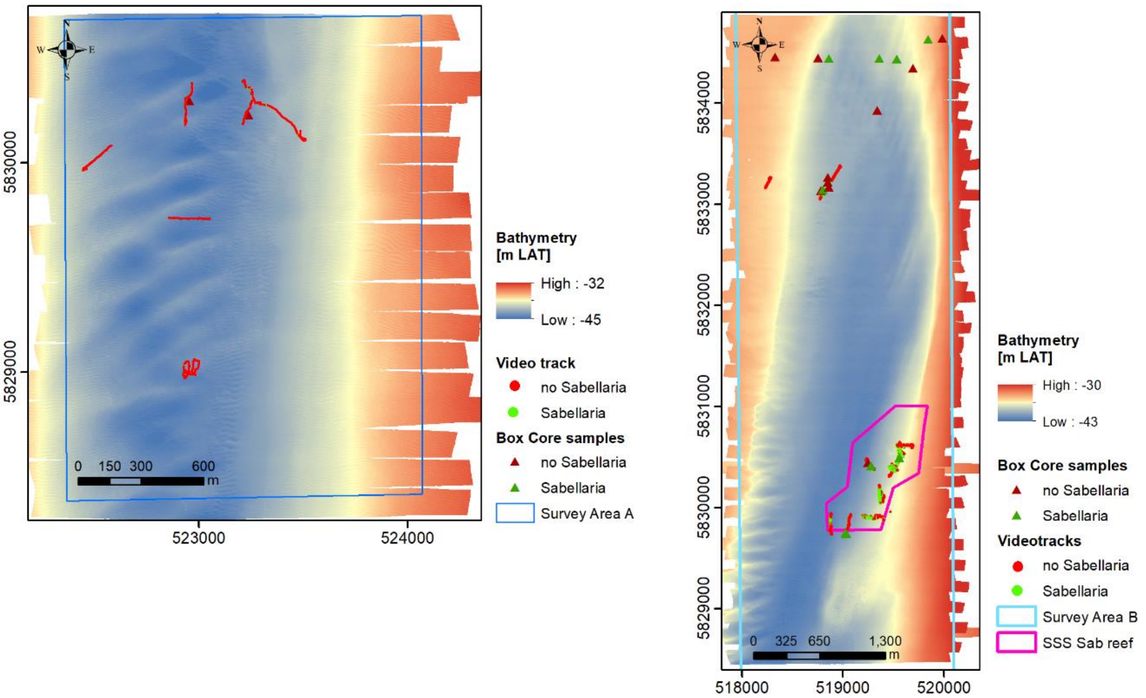
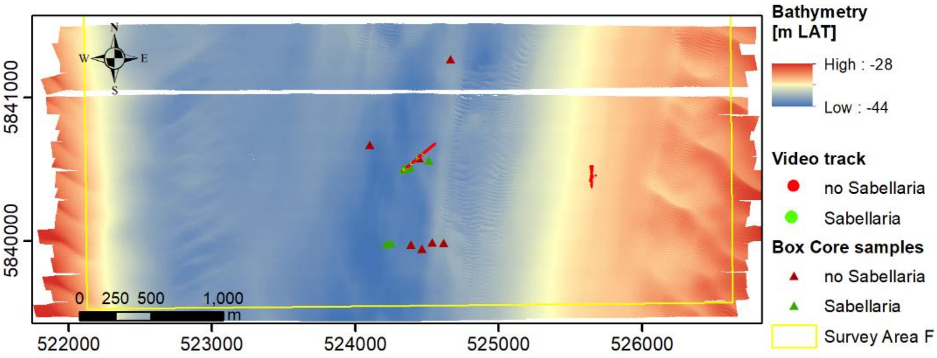


Figure: D.1: MBES bathymetry of Area (left) A and (right) B. In both maps, boxcore locations as well as video tracks are classified to show the absence and presence of *Sabellaria*. The pink polygon shows the manually mapped area where a blotchy pattern in SSS imagery indicates the presence of elevated *Sabellaria* reefs.



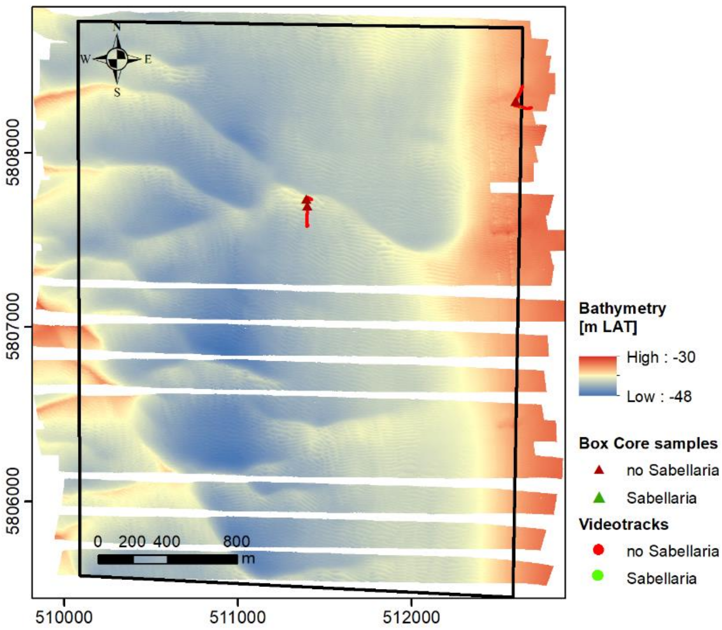


Figure: D.2: MBES bathymetry of Area (top) F and (bottom) G. In both maps, boxcore locations as well as video tracks are classified to show the absence and presence of *Sabellaria*.

Defence, Safety & Security

Oude Waalsdorperweg 63
2597 AK Den Haag
www.tno.nl

innovation
for life