



Ensemble mapping as an alternative to baseline seafloor sediment mapping and monitoring

Daphnie S. Galvez^{1,2} · Svenja Papenmeier³ · Lasse Sander¹ · Alexander Bartholomä⁴ · Karen Helen Wiltshire¹

Received: 3 March 2022 / Accepted: 14 June 2022 / Published online: 23 June 2022
© The Author(s) 2022

Abstract

Seafloor sediment mapping traditionally relies on the experience and expertise of practitioners to classify sediment classes based on acoustic backscatter data and ground-truth information. However, recent advancements in automated seafloor mapping present a major development in objective methods that offer practical application for seafloor mapping and monitoring campaigns. In this study, a class-specific approach of ensemble modelling (ensemble mapping) was used to classify the sediment classes of a large-scale seafloor area (1550 km²) in the western Sylt Outer Reef, German North Sea. A pixel-by-pixel comparison of the modelled map and manually digitized map was also conducted to assess the efficiency of the ensemble mapping approach. The resulting seafloor sediment map, with an overall accuracy of 73%, demonstrates five sediment classes that represent most of the seabed of the German North Sea. The manually classified and ensembled maps were 63% identical, but mismatches were observed in the transitional boundaries of soft sediment classes and in stony areas that were not predicted in manual classification. The inconsistencies between the two maps was attributed to the different interpretation of sediment boundaries, the simplification of the sediment classification scheme, and the ability of ensemble mapping to classify more areas than manual classification. This study found that ensemble mapping performs better in characterizing coarse materials and produces maps that are comparable to the maps produced by manual classification, while the production time and degree of subjectivity in the analysis are minimal. Hence, ensemble mapping is a viable alternative to create baseline seafloor sediment maps that can be used for environmental monitoring and resource planning.

Introduction

Detailed and accurate maps of the seafloor environment is one of the pre-requisites of a successful marine spatial planning (Diesing and Stephens 2015) and to designate marine protected areas (Harris and Baker 2012). Seafloor

maps encompass all types of static maps that represent the physical, biological, and geological properties of the seafloor. This can be a map showing the seafloor morphology (Di Martino et al. 2021; Zeiler et al. 2008), benthic habitats (Baker and Harris 2020; Brown et al. 2011), or seafloor sediment distribution (Bockelmann et al. 2018; McBreen et al. 2008). Mapping of the seafloor has been required by the European Union's Marine Strategy Framework Directive (MSFD) (MSFD 2008) to monitor the environmental status of seafloor integrity in European waters (European Commission 2010). As a result, national seafloor mapping programs that are focused on mapping substrate type as a surrogate for biodiversity (Brown et al. 2017; Diesing et al. 2020; Ware and Downie 2020) were established in EU countries (MSFD 2008, Directive 2008/56/EC; (BSH 2016; Downie et al. 2016; Brown et al. 2017).

In Germany for example, the national seabed mapping program SedAWZ ("Full coverage sediment mapping in the German Exclusive Economic Zone") was coordinated by the Federal Maritime Hydrographic Agency (BSH) to provide detailed information on sediment types, hard-substrate

✉ Daphnie S. Galvez
daphnie.galvez@ve.ismar.cnr.it

Alexander Bartholomä
alexander.bartholomae@senckenberg.de

¹ Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Wadden Sea Research Station, Hafenstraße 43, 25992 List, Germany

² Institute of Marine Sciences–National Research Council of Italy (ISMAR-CNR), Arsenale-Tesa 104, Castello 737/F, 30122 Venice, Italy

³ Leibniz Institute for Baltic Sea Research Warnemünde, Seestraße 15, 18119 Rostock, Germany

⁴ Department of Marine Research, Senckenberg Am Meer, Südstrand 40, 26382 Wilhelmshaven, Germany

distribution, and other seabed features on a scale of 1:10,000 for the national implementation of the MSFD (Papenmeier and Hass 2020). The program standardized the seafloor mapping procedure and generated high-resolution seafloor sediment maps of the German North Sea and the Baltic Sea EEZ (geoseaportal.de, last accessed on 01.04.2022) (BSH 2016; Holler et al. 2019; Papenmeier et al. 2019). The published seafloor sediment map of the Sylt Outer Reef (SOR) in the German North Sea was created using the manual classification approach. However, the production of this map requires time and intensive effort from the producer and since the trend nowadays has shifted to the development of automated seafloor mapping methods (Ierodiakonou et al. 2018; Misiuk et al. 2019; Menandro et al. 2020), an upgrade to automated, reproducible, and robust method would be timely.

Seafloor sediment maps are either produced by manual classification, unsupervised classification, or supervised classification algorithms (Lecours 2019). In manual classification, segmentation of the environmental layers (e.g., backscatter data) are performed manually by expert interpretation, but this approach is dwindling because of its subjectivity and inability to repeatedly delineate transitional habitats (Brown et al. 2011; Lecours 2019; Lucieer and Lucieer 2009). In unsupervised classification, environmental data are segmented before assigning a habitat type based on ground-truth observations made by camera or physical sampling (Brown and Collier 2008; Calvert et al. 2015; Kostylev et al. 2001). In contrast, supervised methods use the ground-truth information to form class signatures to be used to identify similar regions in the environmental dataset where no ground-truth data exist to produce a full-coverage habitat map (Calvert et al. 2015). Despite the progress made in recent years, seafloor sediment classification and monitoring using reproducible and automated methods is still in its infancy (Janowski et al. 2020; Zelada Leon et al. 2020). Several studies attempted to address these issues by classifying backscatter mosaics through automated methods using machine-learning and modelling such as pixel-based and object-based analysis (Diesing et al. 2014; Ierodiakonou et al. 2018; Janowski et al., 2020). One of these methods is ensemble modelling or the combination of multiple modelling algorithms to derive an accurate spatial prediction. Ensemble modelling was first introduced to model suitable habitats or species distribution (Araujo and New 2007; Guisan et al. 2017; Araújo et al. 2019), but it has already been applied in the marine environment to map seabed sediments (Diesing and Stephens 2015; Galvez et al. 2021), submarine geomorphology (Pearman et al. 2020), and benthic habitats in remote places (Jerosch et al. 2019). Another approach for sediment mapping is ensemble mapping, which has been suggested to alleviate the difficulties of predicting sediment classes in a complex heterogenous seafloor (Diesing et al. 2020). In this approach, predictions for

each sediment class were generated using single or multiple classification techniques and then the results were combined into a single map by aggregating the modal classes (Diesing et al. 2020; Misiuk et al. 2019).

A class-specific approach of ensemble modelling or “ensemble mapping” was first introduced in our previous study (Galvez et al. 2021) to address the limitation of minimal amount of available ground-truth samples and to reduce the effect of imbalanced ground-truth data in automated seafloor sediment classification. However, the method still needs to be tested in a larger area and to assess its ability to discriminate more sediment classes, in order to evaluate its efficiency for wide-scale seafloor mapping. Therefore, the aim of this study is to test the capacity of ensemble mapping to classify multiple sediment classes of a large seafloor area and to assess if this approach is a good alternative to baseline seafloor sediment mapping.

Physical settings

The study area (1550 km²) is located in the western part of the *Special Area of Conservation* — Sylt Outer Reef (SOR) in the German North Sea, which is approximately 70 km offshore from the island of Sylt, Germany (Fig. 1d). The area belongs to the eastern part of the Paleo-Elbe Valley (PEV) — a valley that was formed by several glacial advances and retreats during the Pleistocene and was submerged during Holocene transgression (Papenmeier and Hass 2020). The modern water depth in the area ranges from 23 to 46 m (Fig. 1b). The slope is increasing towards the northeast, where the water depth also decreases (Fig. 1c).

Seafloor substrates in the western SOR are characterized by heterogenous seafloor patterns on the east and homogenous seafloor patterns on the west. The heterogenous area is composed of patches of coarse-grained materials, which are classified as lag deposits or as sorted bedforms, and surrounded by finer materials (Diesing et al. 2006; Mielck et al. 2015; Galvez et al. 2020; Papenmeier and Hass 2020). The homogenous area is characterized by finer materials, which is composed of Holocene fine and medium marine sands. These fine materials were reworked during the postglacial transgression and are defined as mobile sediments (Figge 1980; Zeiler et al. 2000; Diesing and Schwarzer 2006; Kösters and Winter 2014).

The seafloor features in the SOR reflect the glacial origin of the subsurface (till) deposits and represent the highly dynamic oceanographic processes on the water column and on the seafloor (Diesing et al. 2006; Heinrich et al. 2017; Feldens et al. 2018; Bartholomä et al. 2020; Galvez et al. 2020). The southern German North Sea is exposed to strong winds from the west, which results in a residual cyclonic circulation that is strongly affecting the sediment transport

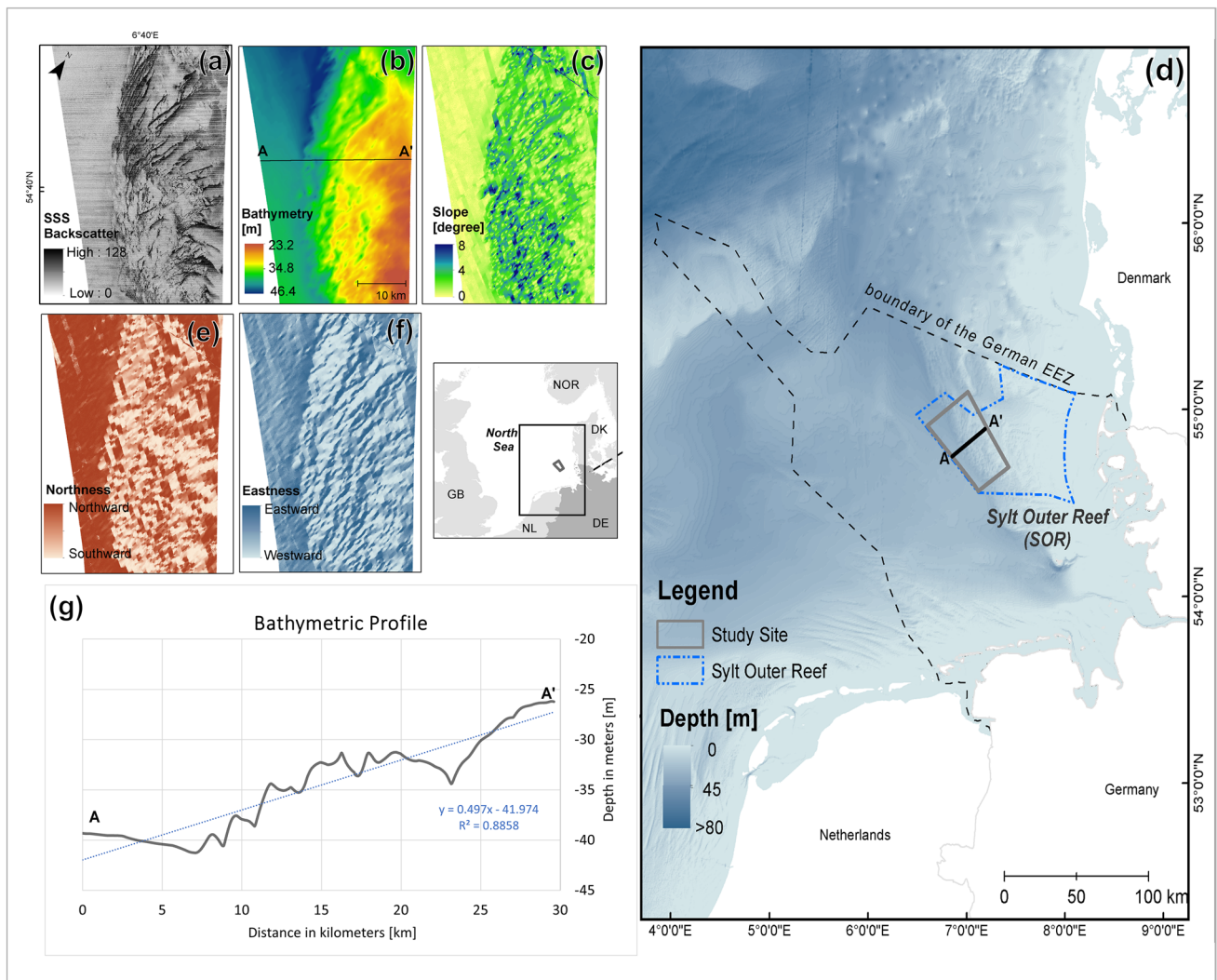


Fig. 1 Location of the study area and its morphological features: SSS backscatter intensity, bathymetry, slope, and aspects

(Staneva et al. 2009; Port et al. 2011; Kösters and Winter 2014; Callies et al. 2017).

Methodology

Data collection and processing

The side-scan mosaic used in this study was collected in April–May 2013, while the ground-truth data were obtained from multiple expeditions from 2013 to 2019 in the SOR with the German research vessels “Heincke” and “Senckenberg”.

Side-scan data were collected with an Edgetech 4200 MP side-scan sonar (Edgetech, West Wareham, MA, USA) at a frequency of 300 and 600 kHz. The range was set at 230 m to achieve a 10% swath overlap and spatial grid resolution of

1 m. The sonar was towed at a speed of 5 km behind the vessel and was kept at 5–10 m above the seafloor. The SSS raw data were processed in SonarWiz (Chesapeake Technology, California, CA, USA) for slant range correction, speed, lay-back, and gain normalization. The nadir line was cut out to 5 m in both port and starboard direction to reduce the noise in the final mosaic. To enhance the quality and spatial accuracy of the SSS data, the mosaics underwent quality control following the procedures of Galvez et al. (2020). Multibeam echosounder (MBES) data were simultaneously collected with a hull-mounted Kongsberg EM710 system (Kongsberg Maritime AS, Kongsberg, Norway). The MBES has two positioning units. The primary positioning system is from Trimble SP461 DGPS (0.5–3 m accuracy), while the secondary unit is DEBEG/Leica GPS (5–15 m accuracy). The very shallow mode with a frequency range of 65–106 kHz and pulse length of 0.2 ms, which is ideal for the < 100-m-depth

range (Kongsberg Maritime 2018), was used in our surveys. The default maximum reliable swath width was 90°. Post-processing of MBES data was conducted in QPS Qimera v2.0.1 software (Quality Positioning Services BV, Zeist, the Netherlands) to correct the raw MBES data from tidal effects and reject invalid soundings.

Ground-truth information was collected by both underwater video and sediment grain-size sampling in 429 different locations, following a standardized sampling protocol (BSH 2016). Sediment samples were collected using a Van Veen Grab (HELCOM standard), while underwater videos were obtained using a Kongsberg OE14-366 Zoom Camera (Kongsberg Maritime AS, Kongsberg, Norway) and a GoPro 3 + Black Edition (GoPro, Inc., San Mateo, CA, USA). The cameras were mounted on a robust metal frame with a laser scale, and deployed underwater as close as possible to the seafloor for at least 5 min while the ship was drifting at a speed of less than 1 kn. The on-board control unit of the camera system was connected to the GPS network of the research vessel for positioning. Ground-truth data from underwater videos were only used in the model to predict the location of lag sediments, because grab samples of lag sediment (e.g. boulders, rocks) were not collected. Furthermore, additional ground-truth information on hard substrates (lag sediment) that was obtained from parametric sediment echosounder using automated stone detection was used in the model (Papenmeier and Hass 2018).

Sediment samples were processed in the home laboratory to remove the carbonate and organic matter using chemical treatment according to the procedures described in Hass et al. (2010). Subsequently, the samples were analysed using a CILASS 1180L laser particle sizer (LPS, range: 0–04–2500 µm). Particles larger than 2000 µm were dry-sieved. Grainsize statistics were calculated in GRADISTAT v8.0© (Blott and Pye 2001). Underwater videos were screened for image quality to omit blurred footage, and the remaining videos were then converted into individual images at 2-s intervals using the scene video filter of VLC media player (VideoLan project, version 3.2.1.0).

All ground-truth samples were categorized according to Folk and Ward's (Folk and Ward 1957) and BSH sediment classification (BSH 2016). The level A and C of the BSH sediment classification scheme was used to classify the ground-truth samples as lag sediment, coarse sediment, medium sand, fine sand, and fine sediment (Supplementary Table 1). The two levels were chosen to harmonize the sediment classes that were used for both ensemble mapping and manual classification.

Manual classification

The SSS backscatter mosaic of the western SOR was manually digitized by an expert using a digitizer tablet

and ArcGIS v.10.4, following the mapping procedures in the “Guideline for Seafloor Mapping in German Marine Waters” (BSH 2016). The backscatter mosaic was characterized based on all ground-truth information from sediment samples and underwater videos and was classified according to all the levels (level A to C) of the BSH sediment classification scheme (Supplementary Table 1). However, for this study, the original manually classified map was reclassified into a simpler classification scheme (only levels A and C) to enable map comparison analysis. The original manually classified sediment distribution map is presented in Supplementary Material Fig. 1. The accuracy of the manually classified map was not calculated because all of the ground-truth data was used in the production of the map; hence, no validation samples remained available for the accuracy analysis.

Ensemble mapping

The ensemble mapping presented in this study was described in detail in our previous publication (Galvez et al. 2021). To cite, ensemble mapping is done by predicting each sediment classes one by one (class-specific) through ensemble modelling, and then each prediction was assembled into one map according to their probability score. The workflow for ensemble mapping presented in Fig. 2 was adapted from the seafloor sediment mapping workflow in the BSH's guideline for seafloor mapping (BSH 2016). The ensemble mapping process is divided into four steps: data preparation, data selection, ensemble modelling, and ensemble mapping.

Data preparation

The ground-truth samples were used as the response variable in the ensemble model, while the textural features and seafloor properties were used as the predictor variables. For this purpose, ground-truth samples and acoustic data were processed beforehand to be compatible for the models.

The location of the sediment and video samples were converted into points and binary format. For example, locations in which coarse sediment was observed were assigned 1, while areas where there was no coarse sediment were assigned 0. Pseudo-absences were also generated for each sediment class in three iterations using a random strategy to prevent sampling bias and to reduce the effect of spatial autocorrelation in the model (Barbet-Massin et al. 2012; Guisan et al. 2017; Galvez et al. 2021).

The original grid resolution (1 × 1 m) of the backscatter mosaic was rescaled to 10 m to match the resolution of the bathymetric data from the BSH (geoseaportal.de, last accessed on 01.04.2022). The rescaling was conducted using the “nearest neighbor” method of the “Resample” tool of ArcGIS 10.7.1. After rescaling, the SSS backscatter

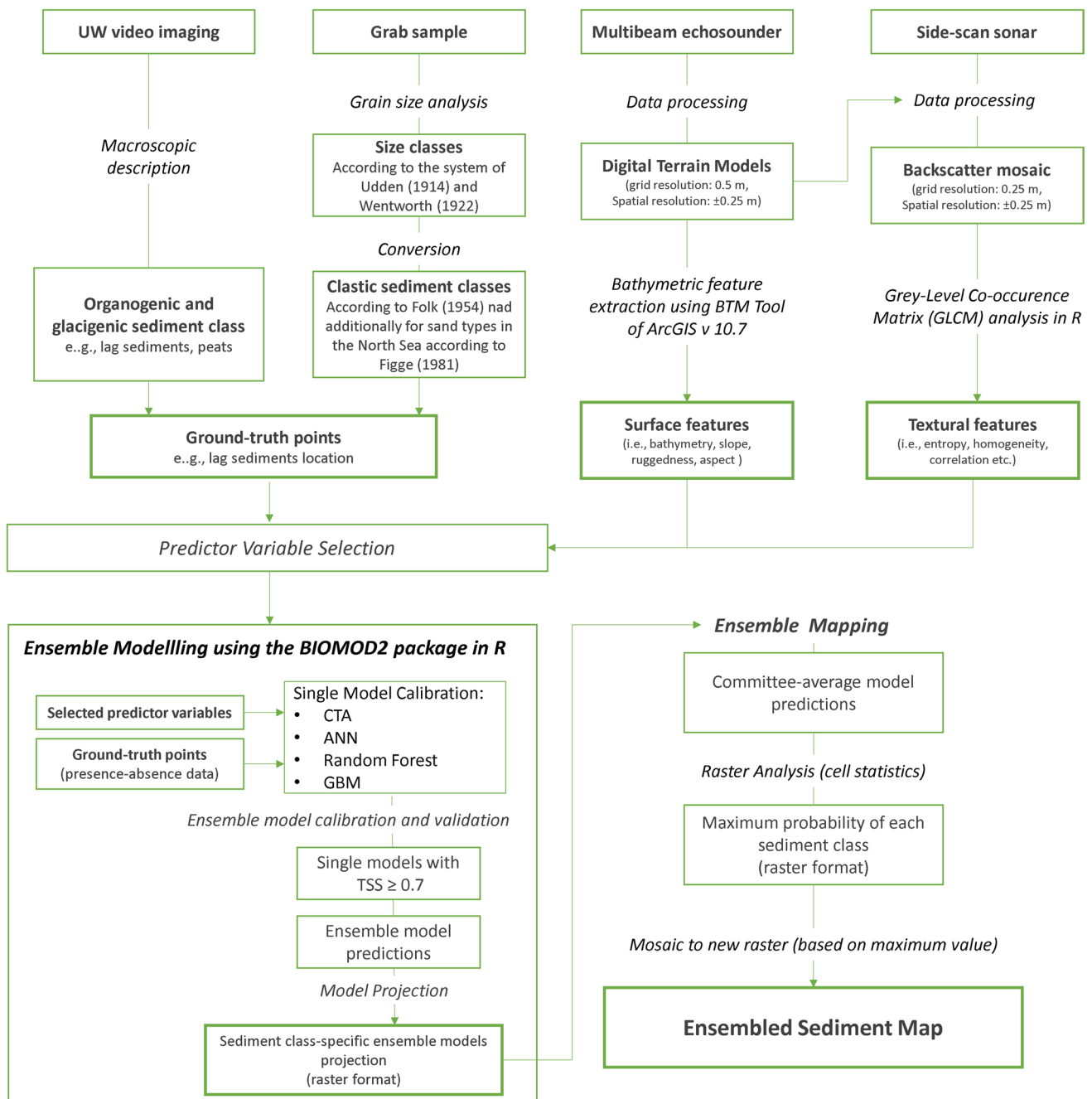


Fig. 2 The seafloor sediment mapping workflow using the ensemble mapping approach which was adapted from the BSH Guideline (BSH, 2016)

mosaic was analysed using the grey-level co-occurrence matrix package in R (GLCM v.1.6.5.) to extract textural features (i.e. homogeneity, correlation, entropy etc.) (Haralick et al. 1973). The bathymetric data was used to calculate seafloor properties (i.e. slope, aspect, and ruggedness) using the Benthic Terrain Modeler v3.0 Toolbox of ArcGIS 10.7.1 (Shaun Walbridge et al. 2018). Information on the textural features, seafloor properties, and

ground-truth samples that were used in the ensemble models is summarized in Supplementary Information 2.

Data selection

Proper selection of appropriate predictor variables is important in a statistical model to avoid model overfitting and multi-collinearity (Breiner et al. 2015; Naimi and Araújo

2016). In this study, predictor variables were selected in an iterative process to ensure that only the most appropriate predictors will be used for the model. Initially, the variance inflation factor (VIF) was used to detect collinearity between predictors and to remove redundant variables. A VIF greater than 10 indicates a collinearity problem (Chatterjee and Hadi 2006). The VIF analysis was performed using the “vifstep” function in the R package “usdm” (Naimi et al. 2014). All predictor variables were analysed in a stepwise procedure, whereas variables with VIF of > 5 were removed. Further feature selection was conducted during model calibration based on the variable importance score of the predictors. The variable importance function in the “BIOMOD2 package” in R (Thuiller et al. 2009; R Core Team 2020) uses a machine-learning approach to randomize one of the variables in each permutation and calculate a correlation score between the standard prediction and the new prediction. The higher the value, the more importance the predictor variable has on the model (Thuiller et al. 2009).

Ensemble modelling

Ensemble modelling was performed using the “BIOMOD2” package within the statistics software R (CRAN) v.4.0.3 (Thuiller et al. 2009; R Core Team 2020). The package was developed for ecologists to predict species distribution, but it can also be used to model any binomial data (Thuiller et al. 2009). Four classifiers were selected within the BIOMOD package for the ensemble modelling: classification tree analysis (CTA), random forest (RF), artificial neural networks (ANN), and generalized boosted models (GBM) (Galvez et al. 2021).

During the ensemble modelling process, the four classifiers trained 70% of the ground-truth data and validated the remaining 30% using a 20-fold cross-validation method. After 240 model runs, the classifiers voted for the best prediction based on the individual model’s performance to generate the final prediction. The model performance was assessed based on their true skill statistics (TSS) value, which is the evaluation of the predictive accuracy of the spatial distribution model (Thuiller et al. 2009). Models with TSS value of ≥ 0.7 indicate very good performance (Thuiller et al. 2009, 2010; Guisan et al. 2017); therefore, only single models with TSS of at least 0.7 were included in the ensemble model. The ensemble model was built based on the committee average of the single models and was used to predict the distribution of the sediment class. Committee average is the sum of the TSS value of the single models divided by the total number of models (Guisan et al. 2017; Pearman et al. 2020). BIOMOD also applies the same weight to all predictions to remove the bias across the selected models when building an ensemble model (Thuiller et al. 2009; Guisan et al. 2017). The uncertainty of the ensemble model

predictions was calculated based on the coefficient of variation of the model predictions. The R script used for ensemble modelling in this study is available at https://github.com/galvezDS/galvezDS_seafloorSed_ensembleModelling.git (accessed on 02 March 2022).

After ensemble modelling, the class-specific sediment distribution maps were filtered automatically in ArcGIS using the “Majority Filter” to remove the noise or artefacts from the ensemble models. The majority filter used a 4×4 grid cell window to determine the most common value (majority) to replace the smaller cells/pixels in the raster image (ESRI 2021).

Ensemble mapping

The last step of ensemble mapping is the aggregation of all ensemble model predictions into one map to generate a seafloor sediment distribution map. Each sediment class that was modelled separately was aggregated into one map in ESRI ArcGIS v 10.7 using the raster analysis functions (see Galvez et al. 2021 for the detailed procedure). Subsequently, the accuracy of the ensembled map was measured using the confusion matrix to calculate its overall accuracy. The “caret” package in R was used to generate a confusion matrix (Kuhn 2008).

Results

Seafloor sediment distribution in the western SOR

Five sediment types were classified by the ensemble model in the western SOR namely: lag sediments (LagSed), coarse sediments (CSed), medium sand (mSa), fine sand (fSa), and fine sediments (FSed) (Fig. 3). Each sediment class was predicted with very good accuracy through ensemble modelling (Fig. 3; Table 1). Medium sand has the highest accuracy while fine sand received the lowest model performance (TSS = 0.70) (Table 1). The class-specific predictions show fine sediment and fine sand as the most dominant classes in the study area. Fine sediment covers 32% of the study area, and fSa occupies 31% (Table 2). In contrast, lag sediment occupies only 7% or 105 km² of the 1550-km² study area.

The ensembled map shows the overall distribution of the five sediment classes in the western Sylt Outer Reef (Fig. 4). The accuracy assessment shows that ground-truth samples were successfully classified in the map with an overall accuracy of 73%. In some areas in the northwest, the model predicted high-backscatter areas as both CSed and LagSed class.

LagSed was observed as areas with the highest backscatter intensity (120–180 grey values) in the SSS mosaic and situated in deeper depths (35–40 m) than its surroundings.

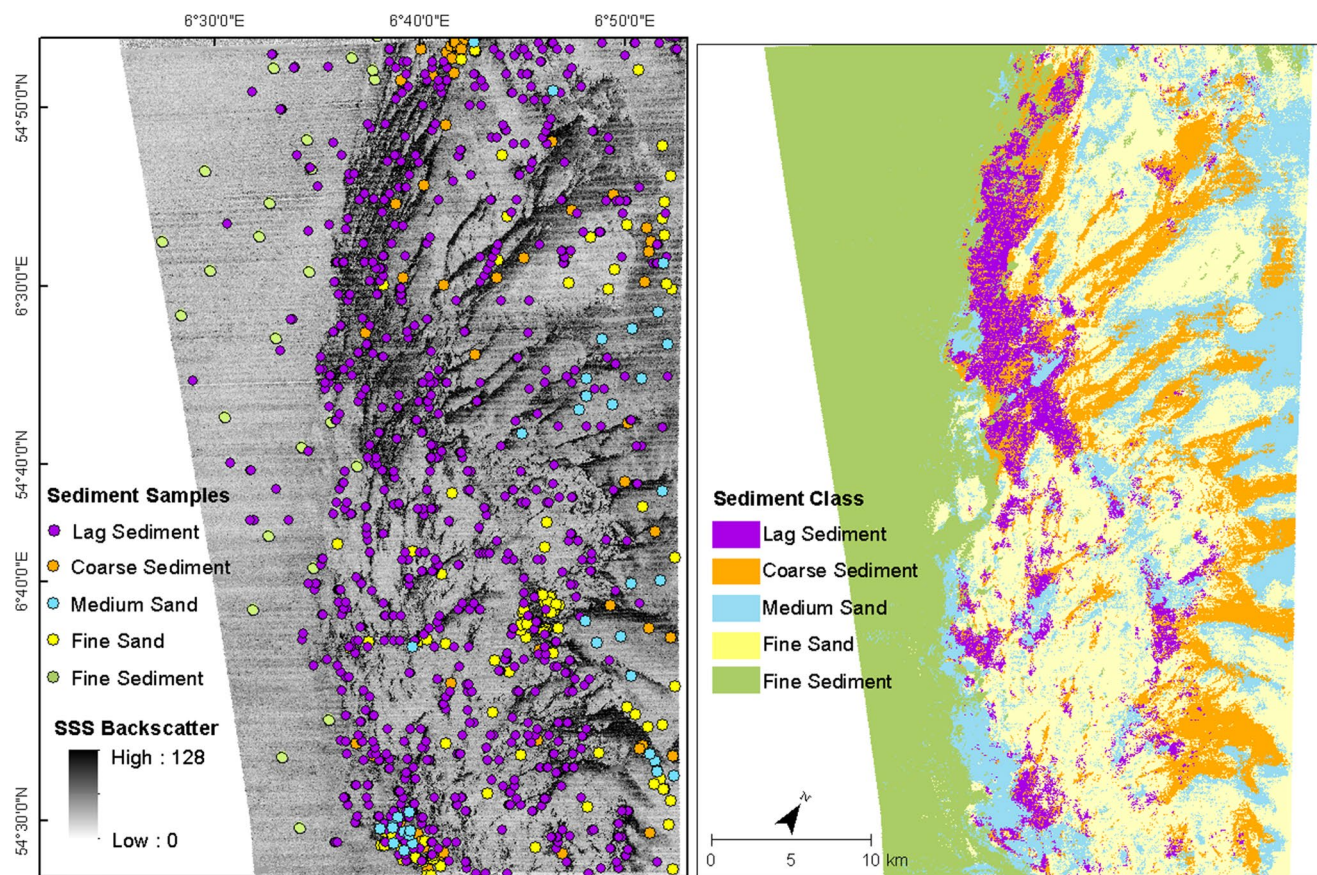


Fig. 3 Mean probability of occurrence and uncertainty of the predictions of each sediment class derived from ensemble models

Table 1 Model performance of the class-specific predictions based on true skill statistics (TSS) and receiving operating characteristic (ROC). TSS value of at least 0.70 indicates very good performance

Ensemble models	TSS	ROC
Lag sediment	0.84	0.97
Coarse sediment	0.80	0.96
Medium sand	0.86	0.97
Fine sand	0.70	0.92
Fine sediment	0.74	0.92

They were mostly found closed to the boundary of the heterogenous and homogenous seafloor and occupies 105 km² (7%) of the study area (Table 1). The Csed class was also reflected with high backscatter intensity (100–200 grey

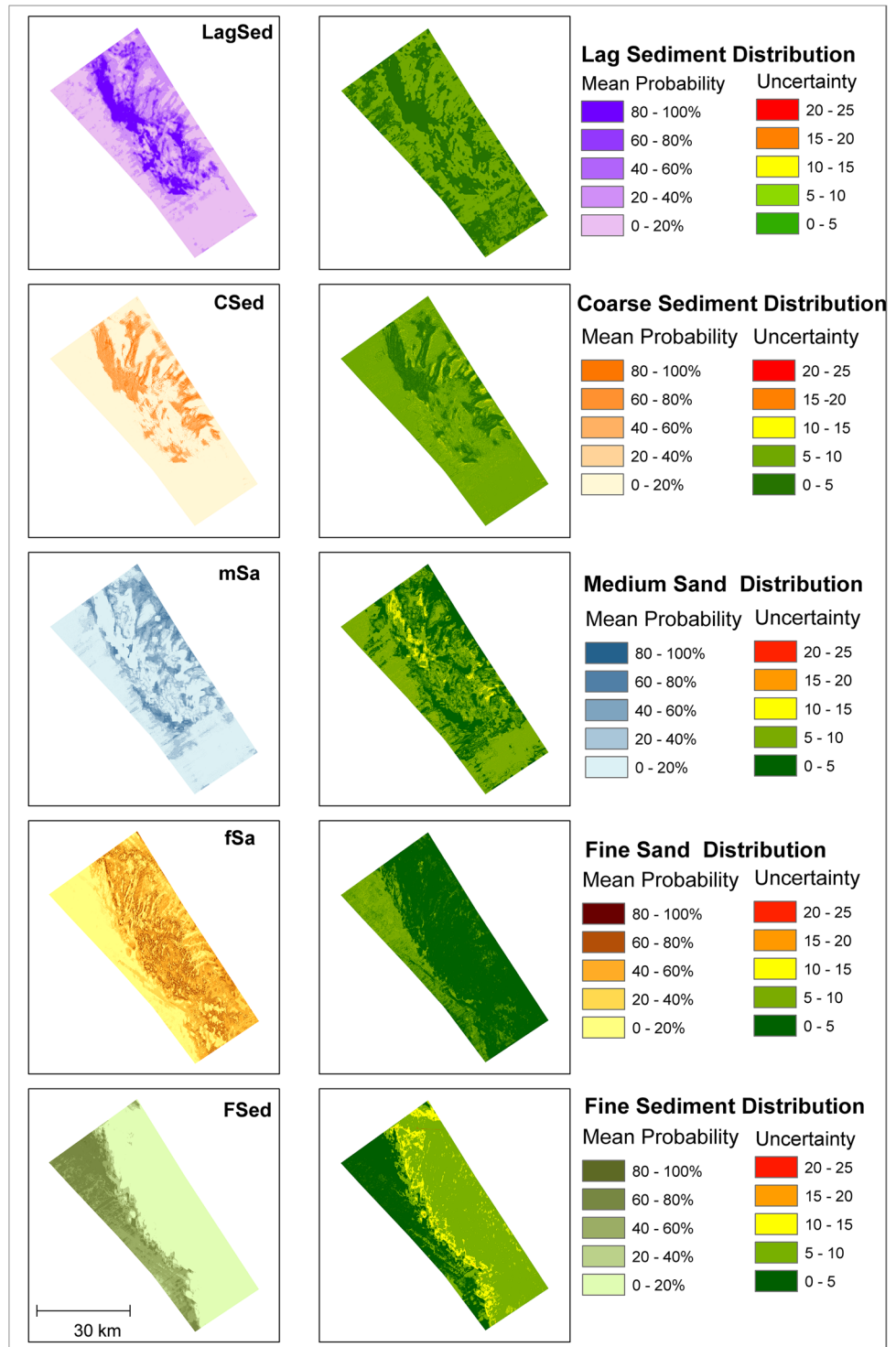
values) and occupies the easternmost part of the study area (184 km²). CSed areas appear as patches of bedform features with northeast-southwest orientation and have a sickle-like shape (Fig. 3). Their boundaries are sharp on one end and fuzzy on the other side (Fig. 3).

Medium sands were generally observed between the coarser sediments (LagSed or CSed) and fine sand classes (Fig. 4) and appeared in the SSS mosaic with medium backscatter intensity (40–50 grey values). They occupied approximately 228 km² (16%) of the total study area (Fig. 3; Table 1). Fine sands dominated the eastern section and were mainly observed in the surroundings of coarser materials (i.e. LagSed and CSed) and in areas with water depths of

Table 2 Total area coverage of the classified sediment types using manual classification and ensemble modelling, with respect to the total study area. The last two columns show the agreement and disagreement of the two maps

	Manual classification	Automated classification	Agree	Disagree
Lag sediment	68 km ² (4%)	105 km ² (7%)	27 km ² (2%)	42 km ² (3%)
Coarse sediment	184 km ² (12%)	225 km ² (15%)	86 km ² (6%)	97 km ² (6%)
Medium sand	228 km ² (15%)	246 km ² (16%)	83 km ² (5%)	145 km ² (9%)
Fine sand	641 km ² (41%)	481 km ² (31%)	356 km ² (23%)	285 km ² (18%)
Fine sediment	429 km ² (28%)	493 km ² (32%)	420 km ² (27%)	8 km ² (1%)
Total	1550 km ²	1550 km ²	63%	37%

Fig. 4 (left) Location of ground-truth samples that were classified into sediment types. (right) Sediment classification map using ensemble mapping



25–30 m and steeper slopes ($2\text{--}3^\circ$). Fine sand areas can be seen in the SSS mosaic as low backscatter areas (20–40 grey values) and dominated the 31% of the study area (Table 2).

Lastly, fine sediment (FSed) occupies the western section, which has lower elevation and flat gradient compared to the eastern section of the western SOR (Figs. 1 and 3) The FSed area was characterized by low backscatter intensity (< 20

grey values) in the SSS mosaic and has deeper water depths that range from 40 to 60 m. Fine sediment occupies most of the study area by approximately 32% (Table 1).

In summary, the heterogeneous eastern seafloor was characterized by patches of coarser materials that were organized into bedform features (Figs. 1a and 4). The bedform features have sickle-like shapes with northeast-southwest orientation

and are composed of either LagSed or CSed classes (Fig. 4). Moreover, these features were found in shallower depths (5–10 m difference from the homogenous seafloor), steeper by up to 3°, and surrounded by sandy materials (mSa and fSa) (Figs. 1c and 4). The sandy materials were found on the eastern side of the bedforms, which is the most elevated and shallowest areas (25 m). On the other hand, the western part of the study area or the homogenous seafloor is dominated by fine sediments, has a relatively flat surface, and is located in deeper water depths (> 40 m) (Figs. 1b and 4).

Map comparison

The ensemble map is 63% identical with the manually classified map (Fig. 5; Table 2). A 37% inconsistency between the two maps was observed along sediment boundaries and in areas that were classified differently through manual classification. For instance, sediment boundaries were depicted as fuzzy boundaries in the ensemble map, but they were delineated as sharp boundaries in the manually classified map. The ensemble map was also able to classify more LagSed, CSed, mSa, and FSed areas than the manually classified map (Fig. 5; Table 2). For example, the ensemble mapping has classified more mSa areas in the southwest and central part of the study area than manual classification (Fig. 5; Table 2).

On the other hand, the manual classification identified more FSed and fSa areas in the map than the ensemble model. The highest disagreement between the two maps was observed in areas of mSa and fSa, while high agreement was found in the location of FSed, CSed, and LagSed classes (Fig. 5; Table 1).

Overall, ensemble mapping was able to identify similar sediment distribution pattern with manual classification, but the time and effort spent were less. The ensemble mapping process of a large area (1550 km²) was conducted in 1–2 days. In contrast, without a structure or guideline to follow at the beginning, the digitalization of the SSS mosaic took 2 years to finished.

Discussion

The goal of this study is to assess the efficiency of ensemble mapping in creating a baseline seafloor sediment map of a large area, and if this approach can substitute manual classification for baseline mapping. The results suggest that the ensemble map is in accordance with the manually classified map by 63%, which means that the ensemble mapping approach produced comparable information with the manually classified map. The method was also able to classify more sediment classes in contrast to our previous study (Galvez et al. 2021) where only two sediment classes

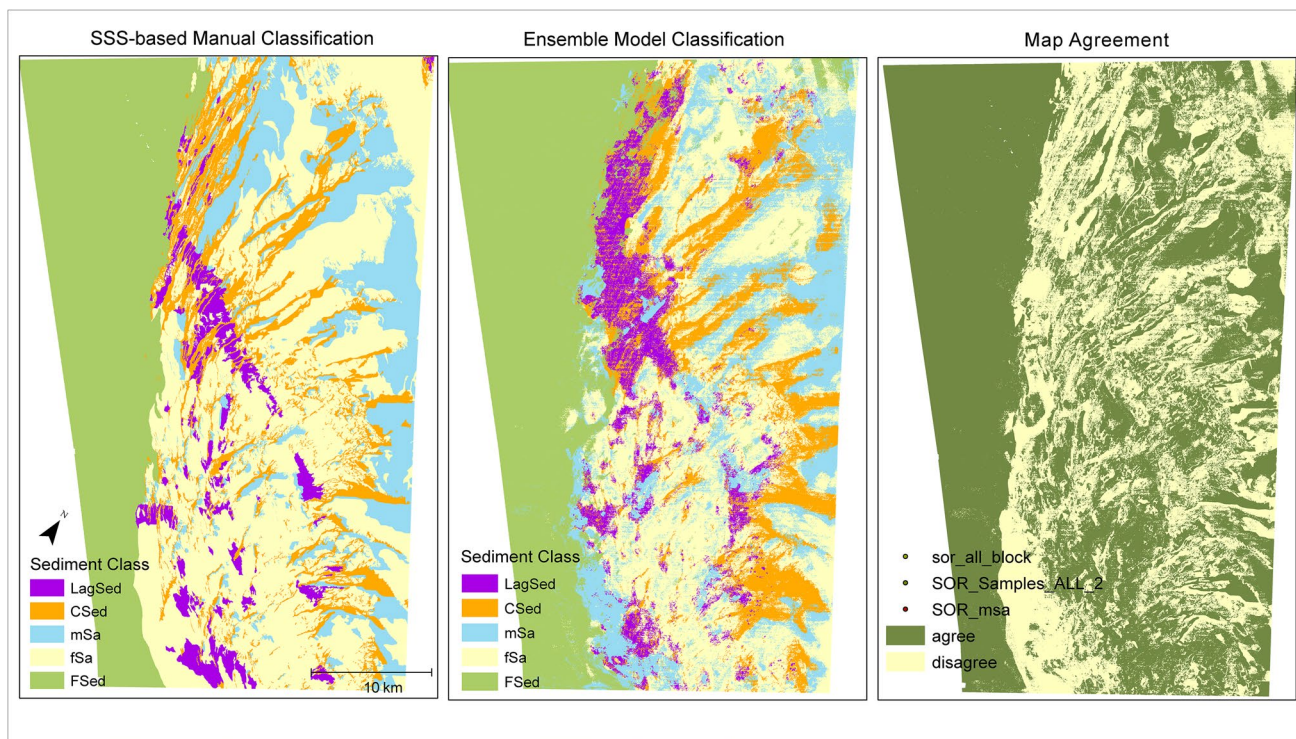


Fig. 5 Comparison of the two maps produced via manual classification (simplified) and ensemble modelling, and their agreement

were classified. Moreover, ensemble mapping performs better in characterizing coarse materials (i.e. LagSed and CSed) because more data was used in the model to predict their distribution, including the map of hard substrates distribution, textural features, and seafloor properties that improved the ensemble model performance. In contrast, only ground-truth data from videos and sediment sampling were used during manual classification to characterize the coarse sediment areas.

Sediment boundaries between soft sediments are naturally transitional and indeterminate and are therefore difficult to interpret subjectively or through manual classification (Lucieer and Lucieer 2009). Ensemble mapping was able to delineate the transitional boundaries of the soft sediment classes better than manual classification. The inconsistencies in the maps highlighted the low capacity of manual classification to distinguish backscatter intensities of transitional areas and soft sediments (i.e. medium, and fine sand). This drawback of manual classification was also identified in previous studies and was addressed by introducing objective classification methods (e.g. Lucieer and Lucieer 2009; Mascioli et al. 2021), as what was conducted in this study.

The ensemble mapping approach also provided class-specific sediment distribution maps with the uncertainty of the predictions (Fig. 3). These maps are useful for planning offshore installations of wind farms, identifying dredging areas, selecting sampling areas for future seafloor mapping efforts, and delineating soft-sediment habitats for environmental management. For example, tubeworms are typically observed in the fine sediment areas of the SOR (Heinrich et al. 2017; Schönke et al. 2017) and these areas were accurately depicted in the output map, which can help environmental managers to locate their habitats.

In some cases, manual classification still has an advantage over automated methods, i.e. it allows unwanted information or artefacts in the backscatter mosaics to be ignored, which can reduce the sources of uncertainties in the final map (Heinrich et al. 2017). More information can also be provided in manually classified maps (e.g. characterization up to sub-classes), because the sediment data used to generate the maps has more information about grain size (i.e. sediment fractions, sorting, skewness, etc.) that can help the mapping expert in deciding the location and type of sediment class during manual classification even with small number of ground-truth data. For example, the original manually classified map in Supplementary Fig. 1 described up to 35 sediment classes, but this was reduced into five sediment classes in the ensemble map. Detailed sediment classification is more challenging in automated classification especially in supervised methods because the ground-truth data were already assigned into class signatures prior to classification. Although a more detailed class signature can be made before modelling, the model requires sufficient and

well-distributed (spatially) ground-truth information to be able to predict accurately.

The disagreement between the two maps can be caused by different factors such as mismatches in scale (e.g. grid resolution), simplification of the sediment classification scheme, and smaller number of ground-truth samples that was used in manual classification. The manually classified map was classified at 1-m grid resolution, but the ensemble map was modelled at 10-m resolution; hence, the mismatches between the two maps can be attributed to the differences in their grid resolution. Simplifying the classification scheme has also resulted to differences between the two maps, which are typical when multiple classes are nested within a broader class (Mitchell et al. 2019). For example, some areas are classified as “fSa + mSa” in the original manually classified map (Supplementary Fig. 1), but these areas were classified only as fine sand in the maps presented in this study (Fig. 5). Lastly, a manually classified map was created with less ground-truth samples than the ensemble map, which may be the cause of poor characterization of hard substrates and medium sand that resulted to inconsistencies in two maps. Manual classification was conducted in 2015, and only the samples available at that time were used for the classification. The ensemble map, on the other hand, used all the ground-truth samples that were collected from 2013 to 2019 to model the sediment distribution. The use of legacy data is not uncommon in the field of seafloor mapping; in fact, it has been utilized to produce a generalized seafloor sediment distribution map of the north-west European continental shelf (Mitchell et al. 2019). However, using ground-truth data from different time periods may incorporate some errors in the map caused by the spatiotemporal variations in sediment conditions. The use of legacy data may be acceptable for various applications such as baseline seafloor mapping that solely aims to inform on the general seafloor sediment distribution but would not necessarily be appropriate for monitoring seafloor sediment changes or to inform certain anthropogenic activities such as aggregate extraction and offshore engineering (Mitchell et al. 2019). Nevertheless, the presented seafloor sediment distribution maps in this study are accompanied by map accuracy information, which can guide the map users to evaluate the usability of the maps for their activities.

In summary, this study supported the applicability of the ensemble mapping approach to map a larger seafloor area and suggests that ensemble mapping is a promising tool to generate a baseline map of the spatial distribution of seafloor sediments in the western Sylt Outer Reef, German North Sea. This study found that ensemble mapping performs better in characterizing coarse materials and produces maps that are comparable to the maps produced by manual classification, while the production time and degree of subjectivity in the analysis are minimal. Though this comparison

was inconclusive as to whether one classification method is better than the other, there were significant differences that determine their suitability for seafloor mapping. For instance, the ensemble mapping approach predicted under-sampled sediment classes accurately and identified transitional areas that are usually difficult to classify manually. This may be an asset in monitoring transitional and sandy areas on the seafloor. In some cases, more information on sediment composition is required to monitor certain benthic habitat communities or seafloor activities (i.e. offshore engineering, dredging); for this purpose, manual classification is more suitable.

Conclusions

In this study, the efficiency of ensemble mapping was proven by comparing its capacity to characterize seafloor sediment distribution with the traditional manual classification method. The advantage of ensemble mapping over manual classification is that it provides comparable information required for mapping sediment distribution, while the production time and expert knowledge needed are minimal. Moreover, the reproducibility of the map is very important to adapt in the continuous environmental monitoring activities. The approach presented here is not only fast and efficient but also allows integration of new input data such as hard substrate maps, hydrodynamic models, and additional ground-truth samples that can be collected in future seafloor mapping campaigns. Alternately, the ensembled map can also be used as data input to model seafloor sediment dynamics and hydrological process or can be integrated into an ecosystem-based management system. In conclusion, ensemble mapping of seafloor sediment distribution is an alternative to manual seafloor classification for the creation of maps that can be the basis or reference point of seafloor monitoring campaigns.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00367-022-00734-x>.

Acknowledgements This research is part of the project AMIN I-III, which is a research and development cooperation between the Alfred-Wegener-Institute, Helmholtz Centre for Polar and Marine Research (AWI), and the Federal Maritime and Hydrographic Agency in Hamburg (BSH). The project is part of the bigger project SedAWZ, coordinated by the BSH and financed by the Federal Agency for Nature Conservation (BfN). We would like to thank the crew of the research vessel ‘Heincke’ and ‘Senckenberg’. This manuscript is dedicated to our beloved colleague H. Christian Hass.

Author contribution Daphnie S. Galvez: conceptualization, methodology, investigation, writing—original draft, visualization; Svenja Papenmeier: conceptualization, investigation, visualization, data collection, writing—review and editing; Lasse Sander: writing—review and editing; Alexander Bartholomae: writing — review and editing;

Karen Helen Wiltshire: writing — review and editing. All authors have read and agreed to the published version of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was funded by Bundesamt für Seeschifffahrt und Hydrographie (grant no. 10038521).

Data availability Hass, H Christian; Papenmeier, Svenja (2019): Sidescan sonar survey (300/600 kHz) during R/V Heincke cruise HE400 in the German Bight with link to raw data files. Alfred Wegener Institute—Wadden Sea Station Sylt, PANGAEA, <https://doi.org/10.1594/PANGAEA.905799>. Other data that were used in this study will be archived in the PANGAEA data bank (www.pangaea.de).

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Araújo MB, Anderson RP, Márcia Barbosa A et al (2019) Standards for distribution models in biodiversity assessments. *Sci Adv* 5:4858. <https://doi.org/10.1126/sciadv.aat4858>
- Araujo M, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22:42–47. <https://doi.org/10.1016/j.tree.2006.09.010>
- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many?: How to use pseudo-absences in niche modelling? *Methods Ecol Evol* 3:327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Bartholomä A, Capperucci RM, Becker L et al (2020) Hydrodynamics and hydroacoustic mapping of a benthic seafloor in a coarse grain habitat of the German Bight. *Geo-Mar Lett* 40:183–195. <https://doi.org/10.1007/s00367-019-00599-7>
- Blott SJ, Pye K (2001) GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surf Process Landforms* 26:1237–1248. <https://doi.org/10.1002/esp.261>
- Breiner FT, Guisan A, Bergamini A, Nobis MP (2015) Overcoming limitations of modelling rare species by using ensembles of small models. *Methods Ecol Evol* 6:1210–1218. <https://doi.org/10.1111/2041-210X.12403>
- Brown LS, Green SL, Stewart HA et al (2017) Semi-automated mapping of rock in the Irish Sea, Minches, western Scotland and Scottish continental shelf. JNCC, Peterborough
- BSH (2016) Guideline for seafloor mapping in German marine waters using high-resolution sonars. BSH No. 7201, p 147

- Callies U, Gaslikova L, Kapitza H, Scharfe M (2017) German Bight residual current variability on a daily basis: principal components of multi-decadal barotropic simulations. *Geo-Mar Lett* 37(2):151–162
- Chatterjee S, Hadi AS (2006) Regression analysis by example. Fourth. John Wiley & Sons Inc, Hoboken, New Jersey
- Diesing M, Stephens D (2015) A multi-model ensemble approach to seabed mapping. *J Sea Res* 100:62–69. <https://doi.org/10.1016/j.seares.2014.10.013>
- Diesing M, Kubicki A, Winter C, Schwarzer K (2006) Decadal scale stability of sorted bedforms, German Bight, southeastern North Sea. *Cont Shelf Res* 26:902–916. <https://doi.org/10.1016/j.csr.2006.02.009>
- Diesing M, Schwarzer K (2006) Erforschung der FFH-Lebensraumtypen Sandbank und Riff in der AWZ der deutschen Nord- und Ostsee (FKZ-Nr. 802 85 270). <https://doi.org/10.13140/2.1.4779.6004>
- Downie AL, Dove D, Westhead K et al (2016) Semi-automated mapping of rock in the North Sea. Joint Nature Conservation Committee, Peterborough
- ESRI (2021) Majority Filter (Spatial Analyst)—ArcMap | Documentation. Environmental Systems Research Institute, Inc. website. <https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/majority-filter.htm>. Accessed 7 Jun 2022
- Feldens P, Schulze I, Papenmeier S et al (2018) Improved interpretation of marine sedimentary environments using multi-frequency multibeam backscatter data. *Geosciences* 8:214. <https://doi.org/10.3390/geosciences8060214>
- Figge K (1980) Das Elbe—Urstromtal im Bereich der Deutschen Bucht (Nordsee). *E&G Quaternary Sci J* 30:203–212. <https://doi.org/10.3285/eg.30.1.16>
- Folk RL, Ward WC (1957) Brazos River bar [Texas]; a study in the significance of grain size parameters. *J Sediment Res* 27:3–26. <https://doi.org/10.1306/74D70646-2B21-11D7-8648000102C1865D>
- Galvez DS, Papenmeier S, Hass HC et al (2020) Detecting shifts of submarine sediment boundaries using side-scan mosaics and GIS analyses. *Mar Geol* 430:106343. <https://doi.org/10.1016/j.margeo.2020.106343>
- Galvez D, Papenmeier S, Sander L et al (2021) Ensemble mapping and change analysis of the seafloor sediment distribution in the Sylt Outer Reef, German North Sea from 2016 to 2018. *Water* 13:2254. <https://doi.org/10.3390/w13162254>
- Guisan A, Thuiller W, Zimmermann NE (2017) Habitat suitability and distribution models: with applications in R. Cambridge University Press, Cambridge
- Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans Syst, Man, Cybern SMC* 3:610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Hass HC, Kuhn G, Monien P, Brumsack H (2010) Climate fluctuations during the past two millennia as recorded in sediments from Maxwell Bay, South Shetland Islands, West Antarctica. *Geol Soc, London, Spec Publ* 344:243–260. <https://doi.org/10.1144/SP344.17>
- Heinrich C, Feldens P, Schwarzer K (2017) Highly dynamic biological seabed alterations revealed by side scan sonar tracking of Lanice conchilega beds offshore the island of Sylt (German Bight). *Geo-Mar Lett* 37:289–303. <https://doi.org/10.1007/s00367-016-0477-z>
- Ierodiaconou D, Schimel ACG, Kennedy D et al (2018) Combining pixel and object based image analysis of ultra-high resolution multibeam bathymetry and backscatter for habitat mapping in shallow marine waters. *Mar Geophys Res* 39:271–288. <https://doi.org/10.1007/s11001-017-9338-z>
- Jerosch K, Scharf FK, Deregibus D et al (2019) Ensemble modeling of Antarctic macroalgal habitats exposed to glacial melt in a Polar fjord. *Front Ecol Evol* 7:207. <https://doi.org/10.3389/fevo.2019.00207>
- Kösters F, Winter C (2014) Exploring German Bight coastal morphodynamics based on modelled bed shear stress. *Geo-Mar Lett* 34:21–36. <https://doi.org/10.1007/s00367-013-0346-y>
- Kuhn M (2008) Building predictive models in R using the caret package. *J Stat Softw* 28:1–26
- Lucieer V, Lucieer A (2009) Fuzzy clustering for seafloor classification. *Mar Geol* 264:230–241. <https://doi.org/10.1016/j.margeo.2009.06.006>
- Mascioli F, Piattelli V, Cerrone F et al (2021) Feasibility of objective seabed mapping techniques in a coastal tidal environment (Wadden Sea, Germany). *Geosciences* 11:49. <https://doi.org/10.3390/geosciences11020049>
- Menandro PS, Bastos AC, Boni G et al (2020) Reef mapping using different seabed automatic classification tools. *Geosciences* 10:72. <https://doi.org/10.3390/geosciences10020072>
- Mielck F, Holler P, Bürk D, Hass HC (2015) Interannual variability of sorted bedforms in the coastal German Bight (SE North Sea). *Cont Shelf Res* 111:31–41. <https://doi.org/10.1016/j.csr.2015.10.016>
- Misiuk B, Diesing M, Aitken A et al (2019) A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using random forest. *Geosciences* 9:254. <https://doi.org/10.3390/geosciences9060254>
- Mitchell PJ, Aldridge J, Diesing M (2019) Legacy data: how decades of seabed sampling can produce robust predictions and versatile products. *Geosciences* 9:182. <https://doi.org/10.3390/geoscience9040182>
- Naimi B, Araújo MB (2016) sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography* 39:368–375. <https://doi.org/10.1111/ecog.01881>
- Naimi B, Hamm NAS, Groen TA et al (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography* 37:191–203. <https://doi.org/10.1111/j.1600-0587.2013.00205.x>
- Papenmeier S, Hass H (2018) Detection of stones in marine habitats combining simultaneous hydroacoustic surveys. *Geosciences* 8:279. <https://doi.org/10.3390/geosciences8080279>
- Papenmeier S, Hass HC (2020) Revisiting the Paleo Elbe Valley: reconstruction of the Holocene, sedimentary development on basis of high-resolution grain size data and shallow seismics. *Geosciences* 10:505. <https://doi.org/10.3390/geosciences10120505>
- Pearman TRR, Robert K, Callaway A et al (2020) Improving the predictive capability of benthic species distribution models by incorporating oceanographic data – towards holistic ecological modelling of a submarine canyon. *Prog Oceanogr* 184:102338. <https://doi.org/10.1016/j.pocean.2020.102338>
- Port A, Gurgel K-W, Staneva J et al (2011) Tidal and wind-driven surface currents in the German Bight: HFR observations versus model simulations. *Ocean Dyn* 61:1567–1585. <https://doi.org/10.1007/s10236-011-0412-9>
- R Core Team (2020) R: a language and environment for statistical computing. Version 4.0.3. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>. Accessed 22 June 2022
- Schönke M, Feldens P, Wilken D et al (2017) Impact of Lanice conchilega on seafloor microtopography off the island of Sylt (German Bight, SE North Sea). *Geo-Mar Lett* 37:305–318. <https://doi.org/10.1007/s00367-016-0491-1>
- Staneva J, Stanev EV, Wolff JO, Badewien TH, Reuter R, Flemming B, Bartholomä A, Bolding K (2009) Hydrodynamics and sediment dynamics in the German Bight. A focus on observations and numerical modelling in the East Frisian Wadden Sea. *Cont Shelf Res* 29(1):302–319
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* 32:369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>

- Thuiller W, Lefebvre B, Araujo M (2010) Presentation manual for BIOMOD. LECA-CNRS, Université Joseph Fourier, Grenoble
- Walbridge S, Slocum N, Pobuda M, Wright D (2018) Unified geomorphological analysis workflows with benthic terrain modeler. *Geosciences* 8:94. <https://doi.org/10.3390/geosciences8030094>
- Zeiler M, Schulz-Ohlberg J, Figge K (2000) Mobile sand deposits and shoreface sediment dynamics in the inner German Bight (North Sea). *Mar Geol* 170:363–380. [https://doi.org/10.1016/S0025-3227\(00\)00089-X](https://doi.org/10.1016/S0025-3227(00)00089-X)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.