



Article Limitations of Predicting Substrate Classes on a Sedimentary Complex but Morphologically Simple Seabed

Markus Diesing ^{1,*}, Peter J. Mitchell ², Eimear O'Keeffe ³, Giacomo O. A. Montereale Gavazzi ⁴ and Tim Le Bas ⁵

- ¹ Geological Survey of Norway, Postal Box 6315 Torgarden, 7491 Trondheim, Norway
- ² Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft NR33 0HT, UK; peter.mitchell@cefas.co.uk
- ³ Marine Institute, Rinville, Oranmore, H91 R673 Galway, Ireland; eimear.okeeffe@marine.ie
- ⁴ Royal Belgian Institute of Natural Sciences, Operational Directorate of Nature, Rue Vautier, 29, 1000 Brussels, Belgium; gmonterealegavazzi@naturalsciences.be
- ⁵ National Oceanography Centre, European Way, Southampton SO14 3ZH, UK; tim.lebas@noc.ac.uk
- * Correspondence: markus.diesing@ngu.no

Received: 28 September 2020; Accepted: 13 October 2020; Published: 16 October 2020



Abstract: The ocean floor, its species and habitats are under pressure from various human activities. Marine spatial planning and nature conservation aim to address these threats but require sufficiently detailed and accurate maps of the distribution of seabed substrates and habitats. Benthic habitat mapping has markedly evolved as a discipline over the last decade, but important challenges remain. To test the adequacy of current data products and classification approaches, we carried out a comparative study based on a common dataset of multibeam echosounder bathymetry and backscatter data, supplemented with groundtruth observations. The task was to predict the spatial distribution of five substrate classes (coarse sediments, mixed sediments, mud, sand, and rock) in a highly heterogeneous area of the south-western continental shelf of the United Kingdom. Five different supervised classification methods were employed, and their accuracy estimated with a set of samples that were withheld. We found that all methods achieved overall accuracies of around 50%. Errors of commission and omission were acceptable for rocky substrates, but high for all sediment types. We predominantly attribute the low map accuracy regardless of mapping approach to inadequacies of the selected classification system, which is required to fit gradually changing substrate types into a rigid scheme, low discriminatory power of the available predictors, and high spatial complexity of the site relative to the positioning accuracy of the groundtruth equipment. Some of these issues might be alleviated by creating an ensemble map that aggregates the individual outputs into one map showing the modal substrate class and its associated confidence or by adopting a quantitative approach that models the spatial distribution of sediment fractions. We conclude that further incremental improvements to the collection, processing and analysis of remote sensing and sample data are required to improve map accuracy. To assess the progress in benthic habitat mapping we propose the creation of benchmark datasets.

Keywords: image-analysis; segmentation; classification accuracy; multibeam echosounder; substrate; bathymetry; backscatter; ensemble; map accuracy; supervised classification

1. Introduction

The majority of anthropogenic pressures on the seabed are concentrated near land within the continental shelf [1–3]. The increasing use of the marine environment and overlapping nature of these

pressures has necessitated the development of marine spatial plans under national or international legislation (e.g., the UK Marine and Coastal Access Act 2009 and the European Union (EU) Maritime Spatial Planning Directive 2014/89/EU). These plans identify suitable zones for different uses, while also providing for environmental protection through the exclusion of all or specific activities through Marine Protected Areas (MPAs). MPAs are one of the main tools with which EU nations plan to meet their obligations under the Habitats Directive (EU Council Directive 92/43/EEC) and regional seas agreements through the Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR Convention). Effective implementation of marine spatial management is underpinned by a detailed understanding of the type, extent, and ecological characteristics of marine habitats to minimise disturbance and to ensure all types of communities are sufficiently protected. This is particularly the case when protecting seabed communities, which are impacted by activities such as bottom-contact fishing, aggregates extraction and offshore infrastructure development. Therefore, to effectively manage the seabed, accurate, high confidence benthic habitat maps are required.

Benthic habitat maps are a simplification of reality and are intended to delineate boundaries around units with similar properties, such as physical characteristics or biological communities. Within Europe this has predominantly focused on substrate type, defined according to the European Nature Information System (EUNIS) hierarchical classification scheme, as a surrogate for biodiversity [4]. These maps form a key step in the delineation and designation of MPAs and subsequently provide baseline information for ongoing monitoring within the MPAs. In the United Kingdom (UK), the MPA network comprises 357 MPAs, the majority of which have been designated under domestic and EU legislation to manage and conserve seabed habitats. Many of these have required the creation of broad or high-resolution seabed maps to support designation and ongoing monitoring of MPAs. Other nations have addressed this need for extensive seabed mapping using a more systematic approach. In Ireland the Integrated Mapping for the Sustainable Development of Ireland's Marine Resource (INFOMAR) programme and its predecessor the Irish National Seabed Survey, have created comprehensive and accessible marine datasets of the physical, chemical and biological features of their seabed in order to underpin a National Marine Planning Framework. Similarly, over the last 15 years the Marine areal database for Norwegian waters (MAREANO) has set about mapping Norway's extensive marine waters and creating detailed seabed geology and biotope maps. This forms a component of Norway's ecosystem-based management approach.

The hardware technology for mapping the seabed has matured by now. Standard mapping is achieved using multibeam echosounder (MBES) or sidescan sonar systems, producing grids of bathymetry and backscatter data with 100% coverage. When complemented with seabed samples, either photographic or physical, these acoustic data products form the basis for the creation of detailed seabed maps, providing an understanding of the actual seafloor substrate and environment. Other data sources may be available too, such as outputs from oceanographic and hydrodynamic models [5,6]. While there has been significant uptake of these standard mapping techniques in the context of seabed mapping, new data acquisition techniques are being developed. The most promising developments include multi-frequency MBESs, which collect data at several frequencies simultaneously and very-high resolution synthetic aperture sonars, which yield extremely detailed sonar imagery of the seabed.

Having high-quality acoustic data products, however, still requires meaningful data analysis. The published literature on this topic has meanwhile grown markedly. Studies on acoustic seabed mapping and classification date back to at least 1993 [7] and a special issue on advances in seafloor mapping using sidescan sonar and MBES bathymetry data was published as early as 1996 [8]. The influential study by Kostylev et al. [9] highlighted the value of MBES data to map seabed habitats. While these authors used expert interpretation to delineate habitats, Dartnell and Gardner [10] presented an early attempt to spatially predict seafloor facies from MBES bathymetry and backscatter using a supervised classification approach. Interest in seabed classification based on acoustic data and seabed samples or observations has increased ever since, witnessed by several special issues dealing with developments in the application of MBES backscatter [11], the geological and biological mapping

and characterisation of benthic marine environments [12], mapping of seabed habitats for better marine management [13], seafloor backscatter data from swath mapping echosounders [14], marine geomorphometry [15] and geological seafloor mapping (https://www.mdpi.com/journal/geosciences/special_issues/Geological_Seafloor_Mapping). Review papers were presented dealing with acoustic seabed classification [16], benthic habitat mapping [17], spatial scale and geographic context in benthic habitat mapping [18], marine geomorphometry [19] and the transfer of knowledge from terrestrial mapping to image-based seabed classification [20]. A sizeable amount of published studies compared predominantly supervised classification approaches in their ability to map seabed substrates and habitats [21–28]. Such studies typically use a measure of map accuracy as a metric to compare the different methods in their ability to derive reliable seabed maps.

The proliferation of seabed mapping studies has shown that MBES data, when combined with groundtruthing samples, can produce accurate and valuable data products [29–32]. However, the diversity of options available means that users often may choose different approaches to the same problem. Such choices are frequently based on aspects like availability of specific software and familiarity of the user with certain types of analytical approaches. Given the breadth of options that are available (e.g., rule-based classification, geostatistics, machine learning, object-based image analysis) it is unlikely that one user will be able to master them all. Knowledge on which approach is most successful under certain circumstances is also still limited. Comparison studies like those mentioned above are therefore still needed. In an endeavour to shed some light on the aforementioned questions we conducted a seabed classification comparison exercise. We invited researchers from the three European seabed mapping programmes MAREANO, INFOMAR and Marine Environmental Mapping Programme (MAREMAP), together with the Royal Belgian Institute of Natural Sciences, to apply their preferred seabed mapping methods to a common dataset of acoustic remote sensing and physical and optical groundtruthing data. To test the limits of the applied methods a highly heterogenous and therefore challenging site to map was selected.

Aim and Objectives

Five approaches to creating a substrate map from a common dataset were undertaken to explore the issues and complexities associated with habitat mapping. The main objectives of the study were as follows:

- (i) examine how each approach succeeded in interpreting the acoustic data;
- (ii) quantitatively measure the accuracy of the outputs from each method of interpretation;
- (iii) explore agreement between the maps and the possibility of creating an ensemble map;
- (iv) discuss the limitations of such approaches in terms of technical issues, prescribed classifications and stakeholder expectations;
- (v) suggest possible solutions to the issues highlighted.

2. Materials and Methods

2.1. Study Site and Data

The area selected for the common dataset was the East of Haig Fras Marine Conservation Zone (MCZ) which covers an area of approximately 400 km² (20 km by 20 km) and is situated within the Celtic Sea on a plateau of the UK continental shelf. Made up of a complex mosaic of rock and sediment substrates, this site has been the focus of four different surveys as part of the UK MPA Programme. The intent of these surveys was to provide evidence for the site designation and subsequent monitoring of seabed habitat condition. An initial acoustic survey in 2012 provided 100% coverage MBES data [33]. Data were acquired using a Kongsberg Simrad EM710 and processed to International Hydrographic Organisation (IHO) Order 1 specifications. Bathymetry data were processed using Caris HIPS and SIPS v7.1 and gridded at 2 m by 2 m resolution. Backscatter were processed in FMGT v7.8.2 using

standard settings and full swath and having removed cross track lines. Backscatter data were exported to the same spatial grid as the bathymetry. Groundtruthing of the site was then undertaken in 2012 and 2013 to support the creation of habitat maps [34] and the designation of the MCZ. A subsequent groundtruthing survey was undertaken in 2015 as the first timepoint in what will be a long-term monitoring programme of benthic faunal communities within the MCZ [35].

For the purposes of this mapping exercise, groundtruth samples collected during all three groundtruthing surveys have been used. For observations of sediment type, only samples collected using a 0.1 m² Mini Hamon grab were retained for analysis and a minimum distance of 300 m between stations was applied to remove resampled stations. From each grab a 0.5 L subsample was used for Particle Size Analysis (PSA). Sedimentary substrates (coarse sediment (CS), sand and muddy sand (Sa), mud and sandy mud (Mu) and mixed sediments (Mx)) were classified according to Long [36] as shown in Figure 1. GPS positions for the sediment grabs were recorded at the instant the grab contacted the seabed. The position was calculated by applying a known offset based on the distance between the side gantry and the antenna location.



Figure 1. Ternary diagram (not to scale) showing four sedimentary substrate classes (after [36]). Composition of samples from the study site are presented in top left panel of Figure 3.

Observations of hard substrate were derived from seabed imagery using a combination of towed camera sledge (2012) and drop camera (2013 and 2015) transects. In 2012, a random subset of the sediment sample locations was targeted for seabed imagery. The 2013 and 2015 surveys then specifically targeted hard substrate based on acoustic data and existing maps [34]. Originally, transects were conducted as 10-minute tows at approximately 0.5 knots (0.26 m s⁻¹) with a still image taken every 30 s. For the mapping exercise only still images where the analyst observed the presence of boulders or bedrock were retained as hard substrate observations. A subset of these were then randomly sampled to produce a single observation per transect. Positional data for seabed imagery were recorded using High Precision Acoustic Positioning (HiPAP), model 500 to measure the offset of the frame relative to the vessel.

In total the groundtruth dataset included 384 samples consisting of five substrate types. A random stratified sampling approach (based on substrate class) was used to split the data into training and testing samples based on a 7:3 split (70% training/30% testing) (Table 1). Only the training samples (totalling 268) were provided to the individual participants for the mapping exercise. Testing samples were used to estimate the accuracy of the individual maps. Table 1 gives an overview of the samples, including the symbols used throughout this study and their relation to EUNIS and the

Marine Habitat Classification for Britain and Ireland (previously Marine Nature Conservation Review, MNCR).Modelled current speed data were also available for the MCZ [35]. Using the hydrodynamic software Telemac2D v7.1, an unstructured triangular mesh was used to model current speeds at various depths within the model domain. The model was developed for the south-west UK continental shelf based on a 3 km node spacing with increased resolution (25 m node spacing) within the study area. By modelling the data across a 30-day spring-neap cycle the mean and maximum current velocity were extracted for the area of interest. This 25 m node spacing was then interpolated to the same spatial grid as the bathymetry.

Table 1. Samples available for the comparison study, classified into five substrate classes, their symbol used throughout the text, their equivalents in the European Nature Information System (EUNIS) and Marine Nature Conservation Review (MNCR) systems and the number of samples in the training and testing datasets.

| Class | Symbol | EUNIS | MNCR | Train | Test | Sum |
|------------------------------|--------|-------|--------|-------|------|-----|
| Bedrock, Boulders or Cobbles | R | A4 | CR | 62 | 27 | 89 |
| Coarse Sediment | CS | A5.1 | SS.SCS | 31 | 14 | 45 |
| Sand and Muddy Sand | Sa | A5.2 | SS.SSa | 74 | 32 | 106 |
| Mud and Sandy Mud | Mu | A5.3 | SS.SMu | 41 | 18 | 59 |
| Mixed Sediments | Mx | A5.4 | SS.SMx | 60 | 25 | 85 |
| Sum | | | | 268 | 116 | 384 |

All participants were provided with the training groundtruth observations and four predictor variables (MBES bathymetry, MBES backscatter, mean current speed at the seabed and maximum current speed at the seabed) with which to derive a seabed substrate map (Figure 2).

2.2. Methods

Unlike studies that tested the performance of different types of spatial prediction methods or machine learning algorithms e.g., [25,28], the choice of the appropriate method was left to the scientists who were involved in this study. Participants were also free to choose which predictor variables to include in the analysis. The five methods used either pixels or image objects/superpixels as the unit of analysis [37]. All classification methods were supervised, i.e., model building was guided by observed classes and their relation to predictor variables. A central assumption of all supervised classification methods is that all substrate types present in the area to be mapped have been observed or sampled.

2.2.1. Method A

The random forest algorithm was used for classification. The algorithm was chosen due to reported high predictive accuracy in studies focusing on the comparison of supervised classifications of MBES data [23,38,39] and has generally proven highly successful in data mining applications [40]. The main underlying assumption of this method is that the predictive power of multiple decision trees is higher than that of a single tree. Bootstrapped samples from the training data are used to construct the individual trees in the forest introducing the first element of randomness. In turn, a random subset of the predictor features is used at the node splits throughout the construction of the model. The result is the construction of unique trees (whose correlation amongst them decreases considerably by this approach). Decisions about the class allocation (labelling) are made on the basis of majority votes of individual trees. Following a feature selection procedure based on the random forest algorithm [41], the random forest model was run growing 500 trees and leaving the parameters as default ($m_{try} = 2$).

To reduce user identified artefacts in the bathymetry a pre-processing step was applied to the bathymetry layer using a majority filter and a neighbourhood of 27 pixels. From this smoothed bathymetry layer, a range of derivatives were generated for inclusion in the model. Feature selection identified five predictor variables that were used for subsequent model training, namely: backscatter,

smoothed bathymetry, Moran's autocorrelation, slope, and bathymetric position index. The routine was implemented in R [42] and models applied using the RandomForest package [43].



Figure 2. Predictor variables, which were provided to the participants of the comparison exercise: (**A**) Multibeam echosounder (MBES) bathymetry relative to Chart Datum and the location of training and testing samples. (**B**) MBES backscatter strength and classified samples. (**C**) Mean current velocity at the seabed from the 30-day model run. (**D**) Maximum current velocity at the seabed from the 30-day model run. (**E**) Location of East of Haig Fras Marine Conservation Zone (MCZ).

2.2.2. Method B

A composite image was generated using 3 variables: backscatter, bathymetry, and maximum current speed. This composite image was classified using the multivariate classification toolset in Esri ArcGIS v10.3 following the methods of Calvert et al. [44] and Dolan et al. [45].

The backscatter data were segmented to generate superpixels for classification. The segmented image was filtered to remove segments generated due to nadir effects. The resulting gaps were filled using neighbourhood analysis of the surrounding cells. This augmented, segmented image was input as the backscatter variable into the composite image. The bathymetry and maximum current speed data did not require any pre-processing and could therefore be input directly as bands 2 and 3 into the composite image. Groundtruth data locations were used to extract statistical descriptions of each substrate class for each band. This statistical information was stored in a signature file and used to classify the entire area using the Maximum Likelihood Classification tool.

2.2.3. Method C

A supervised classification based on substrate observations and a large array of predictor variables was carried out. Subsequently, image pixel-level results were generalised by applying a multiresolution segmentation and averaging over the resulting image objects. Apart from the segmentation process, which was carried out with eCognition software, all processes were run in R statistical software.

Image pre-processing included de-speckling of the backscatter data with a 3 pixel by 3 pixel Lee filter (Speckle Function in ArcGIS Image Analysis). The following derived variables were calculated from the primary variables bathymetry and de-speckled backscatter: backscatter roughness, curvature, eastness, northness, roughness, slope topographic position index (tpi) and vector ruggedness measure (vrm). The calculation was carried out at multiple scales with neighbourhood sizes of 3, 5, 7 ... 51. Subsequently, multi-scale means, and standard deviations were calculated across scales. Together with mean and maximum current speed, 272 potential predictor variables were submitted to a two-step predictor variable selection process [46]: the Boruta algorithm [41] was used to identify important variables. Correlated variables were subsequently removed. The final selection included backscatter (17), roughness (21) backscatter roughness (5), tpi (51) and vrm (3).

A random forest classification model was trained on the training samples and the selected predictor variables. The number of trees in the forest was set to $n_{tree} = 500$, while the number of predictor variables to consider at any given split (m_{try}) was tuned with a grid search algorithm. A value of $m_{try} = 2$ was subsequently used and class probabilities were predicted at the pixel level.

A multi-resolution image segmentation was carried out on the backscatter (17) layer with a scale parameter of 10, shape = 0.1 and compactness = 0.5. The segmentation result was then used as a means of generalising the noisy pixel-level predictions by calculating mean class probabilities per image object and, based on this, the modal substrate class.

2.2.4. Method D

Seabed maps were generated using the Remote Sensing Object Based Image Analysis (RSOBIA) segmentation toolbar [47] in ArcMap, following the approach applied in [48]. The two main independent datasets were considered to be the bathymetry and backscatter mosaic. Both were of comparable acquired resolution and therefore could be taken to be of good coverage of the survey area. The modelled current velocity data were not used.

The RSOBIA toolbar is designed to be quick and user-friendly. The "MBES segmentation" tool requires the bathymetry and backscatter files and the user to define the number of different clusters to be identified and the minimum size of the clusters [48]. The values of these two user defined parameters are based on whether the focus is to produce a high-resolution interpretation of small patches of the survey area or a more regional interpretation of large patches of the survey area. The default values of 20 clusters and 10,000 pixels minimum size were initially used and an

interpretation created. A second run of the program (with a smaller 200 pixels minimum polygon size) produced a much higher resolution interpretation. The RSOBIA tool calculates the slope and local roughness from the bathymetry data and combining the backscatter data with these two derivatives uses a k-means clustering algorithm [49] to find data clusters (classes) which are then put through an iterative elimination process to provide geographically distinct areas. More clusters than final data classes were given so that similar clusters can be amalgamated by the user. The fine scale interpretation was selected as the final segmentation.

The segmentation was then compared with the sample grab data. The segmentation class numbers were correlated with sample type and used to spread classification descriptions to areas without actual samples. In some cases, two differing sample types were found to be in a single segmentation class, at which point expert judgement was given.

2.2.5. Method E

Six derivatives of bathymetry were created using the Terrain Attribute Selection for Spatial Ecology (TASSE) tool [50] based on a circular radius of 25 cells. Some noise was present in the bathymetry data, so this radius was selected to characterise the local depth variation rather than the noise. Backscatter, bathymetry, the six derivatives of bathymetry (bathymetry mean, eastness, northness, Relative Deviation from Mean Value (RDMV), standard deviation and slope) and maximum currents at the seabed were imported into eCognition 9.3 for segmentation using the multiresolution segmentation algorithm. Various scale parameters and input variables were trialled with the final selection based on expert judgement. In this case, oversegmentation was considered preferable to undersegmentation so the intention was to ensure that boundaries visible to the eye were also present in the segmentation. The final segmentation used backscatter, bathymetry, and the standard deviation of bathymetry layers (with weightings of 2:1:1, respectively) and a scale parameter of 30 (default shape and compactness). Objects where a sample was present were classified according to that sample class and the mean of each predictor variable was then calculated for all objects along with geometric variables such as size and shape.

Variable selection was determined using Principle Component Analysis. Six variables were to be included in the model as these explained 80.1% of the variance and had eigenvalues >0.99. For each component, the most highly correlated variable was included. This resulted in bathymetry, backscatter, eastness, northness, RDMV and slope being retained for modelling. The most probable seabed substrate type was modelled using the methodology described in Mitchell et al. [51]. A bootstrap aggregation modelling approach was used to derive multiple random forest [52] models from the training data by repeatedly randomly subsetting the samples with replacement. Models were then used to predict across the study site with the final prediction derived from a plurality vote. Individual random forest models were generated under default settings based on a 70% split of the training data, and the plurality vote was based on 25 subsets of the training data.

2.2.6. Measuring Map Accuracy

The assessment of map accuracy was based on the testing samples, which were not used in the predictions. Predictions made with the methods described above were extracted for every location in the testing dataset and confusion matrices were built. Overall accuracy was used to evaluate the global accuracy of the predictions, while error of omission and error of commission were selected as class-specific metrics of accuracy. The overall accuracy gives the percentage of cases correctly allocated and is calculated by dividing the total number of correct allocations by the total number of samples [53]. The error of omission is the number of incorrectly classified samples of one class divided by the total number of reference samples of that class. The error of commission is the number of incorrectly classified samples that were classified as that class [54]. The overall accuracy, its 95% confidence intervals and a one-sided test to evaluate whether the overall accuracy was significantly higher than the no information rate (NIR) were

calculated by applying the confusionMatrix function of the caret package [55]. The confidence interval is estimated using a binomial test. The NIR is taken to be the proportion of the most frequent class. Errors of omission and commission are not provided by the function but can be calculated from the confusion matrix.

2.2.7. Ensemble Map and Map Agreement

All predictions were rasterised (if necessary) at a resolution of 10 m and aligned. Class-specific agreement between the map outputs was determined by counting the number of instances a class was predicted for every pixel. The respective pixel was assigned the class that was most frequently predicted (modal class) to derive an ensemble map. The agreement between the map outputs of the modal class served as an indicator of confidence in the ensemble predictions.

3. Results

3.1. Data Exploration

Figure 3 summarises the sediment samples (n = 295) and their relationship with backscatter strength. In the ternary diagram, samples plot in the right-hand half indicating a predominance of the sand and, to a lesser extent, gravel fraction. Backscatter strength increases with gravel content and decreases with sand and mud content. In the ternary diagram, backscatter does not exhibit a clear trend in relation to the three grain-size fractions.



Figure 3. (**A**) Ternary mud–sand–gravel diagram showing the sediment samples colour-coded according to Long (2006). (**B**) Same as previous but coloured according to associated backscatter strength (bs) in dB. (**C**) Gravel content versus backscatter strength. (**D**) Sand content versus backscatter strength. (**E**) Mud content versus backscatter strength.

The boxplots of backscatter and bathymetry by substrate class reveal a large overlap in the ranges of those predictor variables (Figure 4). This is particularly obvious for the backscatter response of CS

and Mx, which are near-identical, and to a lesser degree Mu and Sa. Bathymetry does not appear to separate between substrate types apart from Mu. Mean and maximum current speed appear to have virtually no discriminatory power.



Figure 4. Boxplots of the predictor variables (**A**) bathymetry, (**B**) backscatter, (**C**) mean and (**D**) maximum current speed in relation to the five substrate classes.

3.2. Map Accuracy

All maps have overall accuracies significantly higher than the NIR (Figure 5). Overall accuracies vary from 46% to 59%. R is generally predicted with acceptable accuracy (mean omission error \approx 15%, mean commission error \approx 25%), but prediction errors are high for all other substrates (Figure 6). Accordingly, the overall accuracies are low (mean of 50%). The ensemble map does not perform better than the individual methods A to E. Contingency tables relating to the five produced maps are available as Supplementary Tables S1–S5. The data behind Figure 6 is available as Supplementary Tables S6.



Figure 5. Overall map accuracy by method A to E and the ensemble. Error bars give the 95% confidence limits. The dashed horizontal line indicates the no information rate (NIR).



Figure 6. Mean errors of (**A**) omission and (**B**) commission for the five substrate types, averaged across methods A to E. Error bars give one standard deviation.

3.3. Map Agreement

Individual map outputs of the five methods are shown in the Supplementary Figures S1–S5. The agreement between methods for the five substrate classes is shown in Figure 7. These maps might be interpreted as showing the probability of a specific substrate class to occur in a specific location (pixel). There was no class predicted by the majority of maps for only 15.4% of the study area; however, only 15.8% of the area was predicted as the same class in all maps (Table 2). The ensemble map (Figure 8) based on the predictions of Methods A to E shows the modal class, i.e., the substrate class that was most frequently predicted and the agreement between the predictions. The contingency table of the ensemble map (Table 3) shows that R is generally predicted with high accuracy, e.g., 25 of the 27 reference samples labelled as R were also predicted as R. Sedimentary substrates did, however, show a high degree of confusion and low accuracy. This is especially true in the case of CS, which was most frequently predicted as Mx. No class (i.e., map agreement < 3) was most frequently predicted in areas of Mx, followed by Sa. Based on the test set, we also assessed prediction accuracy for individual agreement classes (Table 4). This highlights that map accuracy is higher when all five predictions agree.



Figure 7. Agreement between methods for the five substrate classes: coarse sediment (CS), mud and sandy mud (Mu), mixed sediments (Mx), bedrock, boulders or cobbles (R) and sand and muddy sand (Sa).

Table 2. Total predicted area, as a percentage of the mapped area, for each class in the ensemble map. Derived from right panel in Figure 8. Where the five methods failed to predict a majority class this was classified as 'No class'.

| No. of Maps with Agreement | CS | Mu | Mx | R | Sa | No Class | Sum |
|-------------------------------|-----|------|------|------|------|-------------|-------|
| 1–2 | - | - | - | - | - | 15.4 | 15.4 |
| 3 | 2.6 | 5.6 | 11.3 | 4.5 | 12.9 | - | 36.9 |
| 4 | 0.5 | 4.5 | 7.3 | 3.9 | 15.7 | - | 31.9 |
| 5 | 0.0 | 1.5 | 2.8 | 3.8 | 7.7 | - | 15.8 |
| Sum | 3.1 | 11.6 | 21.4 | 12.2 | 36.3 | 15.4 | 100.0 |

Table 3. Contingency table of the ensemble map. Where the five methods failed to predict a majority class this was classified as no class (NC).

| | | Reference | | | | | |
|------------|-----|-----------|----|----|----|----|-----|
| | | CS | Mu | Mx | R | Sa | Sum |
| Prediction | NC | 1 | 1 | 7 | 2 | 5 | 16 |
| | CS | 1 | 0 | 1 | 0 | 1 | 3 |
| | Mu | 1 | 8 | 1 | 0 | 4 | 14 |
| | Mx | 8 | 1 | 8 | 0 | 4 | 21 |
| | R | 1 | 0 | 2 | 25 | 2 | 30 |
| | Sa | 2 | 8 | 6 | 0 | 16 | 32 |
| | Sum | 14 | 18 | 25 | 27 | 32 | |



Table 4. Map accuracy by agreement class. Count gives the number samples in the test set.

Figure 8. Ensemble map based on predictions of Methods A to E: (**A**) shows modal class, which has the highest agreement between methods, as shown in (**B**). Areas with agreement below 3 were assigned no class.

4. Discussion

The overall accuracy of the maps from the various mapping methods were lower than expected, although significantly higher than the NIR. These five attempts to map the site use various approaches that are commonly applied, yet accuracies of $\approx 50\%$ suggest limited success. An 85% target accuracy has often been adopted in thematic mapping for remotely sensed data but stems from assumptions that pixels are purely a single thematic unit, areal extent is definable and that all groundtruth data are positionally accurate and error-free [56]. The accuracy of the five maps would make monitoring change and implementing management strategies difficult, as the map error would be considered so great that actual seabed change would need to be enormous to be detected. While these data types (acoustics, grabs and videos of rock) and the classification scheme (EUNIS) are commonly used in conservation and spatial management and the mapping methods have been widely applied with success in other studies, supporting their use, the results suggest that these are not always suitable. It appears timely to assess what issues would be preventing better maps from being achieved and consider alternative approaches for challenging sites such as this one. While there are other sources of error, we consider the main issues for this site to fit under four main categories: (i) Sample type and classification scheme; (ii) acoustic discrimination; (iii) scale and (iv) outputs. Each of these sources of error are discussed below along with potential options that could be used in the future to overcome these issues.

4.1. Sample Type and Classification Scheme

Model development and accuracy assessments used reference data in the form of grab and video samples to determine what the 'true' substrate type was for a given location. Therefore, it is worth considering how representative the reference data are that have been classified. For example,

the variation in sediment composition across the site is gradational and the classification based on a rigid scheme might appear arbitrary (Figure 3). Small variations, within the margin of error for different analysis techniques [57], could result in two samples of the same sediment, in terms of broad character and associated biota, being classified as different classes. The need to fit continuous data into predefined categories can therefore be problematic as samples that differ by only a few percent may be acoustically similar but classified as different sediment types. This is a common issue of broad scale classification schemes such as the EUNIS level III, that was adopted from a geologically minded classification scheme, and not based on class descriptions that are defined from acoustic data. It also means the majority of sediment information (originally measured to half-phi grain size distributions) was discarded prior to any attempt to model sediments. Further considerations for repeatability of sediment sampling relate to how subsamples of gravelly sediments are taken, how shell fragments are treated during grain size analysis, and the potential loss of parts of the fine fraction during retrieval of grab samples from the seabed.

For a map to support marine spatial planning and environmental conservation it must have a classification scheme that is relevant to the environment that is being measured and these classes must be mappable. There are many alternative classification methods and schemes to EUNIS level III from which maps of higher accuracy could be created. For example, in order to define sediment classes that are distinctly different one option may be to group sediment fractions in a way that maximises entropy between the groups [58]. It is also worth keeping in mind that if the maps are created to support our understanding of how the biology varies across a site, then a focus on physical properties may not be ideal. The question may be how large a variation in grain size fractions is required to change the biological community, given changes of a few percent (of mud, sand, and gravel fractions) do not appear to impact the acoustic signature? Cooper et al. [59] have suggested defining meaningful habitat groups based on infaunal species assemblages and then developing maps from these classes. It is possible that some of the issues encountered in this challenging site would become irrelevant, for example if the infaunal species assemblages observed in similar classes (like CS and Mx) supported the same communities of species.

The EUNIS level III scheme was developed at a transnational level, and has succeeded in supporting harmonisation of samples and maps across research institutes, governments and organisations [60,61]. Site specific classifications, such as maximum entropy and infaunal assemblage classes, may not be desirable where maps form part of larger programmes and management strategies. More detailed classification schemes are available (i.e., further subdividing the sediment classes such as Folk 11 or Folk 16 [60,62]; however, increased classes may also require more samples as each class would become "rarer". On the other hand, reducing the number of classes by aggregating "overlapping" categories (in this case merging Mx with CS and Sa with Mu) has the potential to improve classification accuracy by compromising thematic resolution. It is therefore up to the manager to determine whether less classes would still be relevant for managing the area at an ecological scale.

4.2. Acoustic Discrimination

A discussion of defining mappable classes leads to the second issue: insufficient information in the two acoustic datasets used by all analysts (bathymetry and backscatter) limited the ability to discriminate between the thematic classes. In this study the bathymetric range was only 10 m over an area of 400 km² and therefore much of the interpretation was dependent on the backscatter strength. Other studies have noted that a single-frequency backscatter layer contains insufficient information to accurately predict the full Folk sediment spectrum, and class-aggregation may be required [63–65]. Attempts to map this site encountered the same issue. Differences between R and areas of Sa or Mu were evident in the backscatter (likely resulting from the contribution of the differences in acoustic impedance contrasts), but backscatter failed at separating Mx from CS as well as having considerable overlap between Mu and Sa classes (Figure 4). It is clear from both pioneering [66] and more recent research e.g., [67] that the backscatter strength registered by an MBES does not exclusively relate to the relative percentages of the sediment fractions. Rather, it relates to differences in a combination of acoustic impedance contrasts, sediment, and topographic roughness (i.e., at the level of the grain size which exhibits intrinsic roughness and at the level of sub-beam topographic roughness such as oscillatory ripples) and sediment volume inhomogeneities [68]. Refining the relationship between frequency, seafloor type (including porosity, compactness and permeability) and backscatter response is continually being improved [14,16,63,65,69,70]. Further developments to better utilise backscatter data such as angular response analysis (ARA) [71–73] and the hyper-angular cube (HAC) [21,74] may also increase the predictive power of acoustic data [75]. The recent development of MBES systems that can acquire multi-frequency backscatter simultaneously also has the potential to improve mapped accuracy [29,76]. As the low and high frequencies interact with soft sediments differently [77] this could potentially prove most beneficial for complex but morphologically flat sites such as this one, essentially doubling the amount of information available compared to a single frequency MBES.

4.3. Scale

Scale issues are of fundamental importance in seabed mapping but are frequently not explicitly dealt with [18]. Scale mismatches frequently occur between observations and the resolution of predictor variables. In this study, the pixel size of the predictor variables was 2 m by 2 m. This compares with an area of 0.1 m² (1/40 of the area covered by a pixel) sampled with a mini Hamon grab as employed in this study. Furthermore, only a sub-sample of approximately 500 mL was used for subsequent grain-size analysis. This indicates a scale mismatch between the response and predictor variables. The implications are that grab samples might not necessarily be representative of the environment to which they are related. The field of view of the optical data is dependent on the type of equipment. In the case of a drop camera, the field of view varies with the altitude above the seabed between maybe 0.2–1 m². The towed sledge system did provide a constant field of view of approximately 0.7 m². Such fields of view are much closer to the size of a pixel of 4 m², hence optical data might be more representative when compared to grab sample data.

Several methods used superpixels or image objects rather than pixels as the unit of analysis. The size of the created polygons is however variable and dependent on the segmentation parameters (which differ between approaches) and the spatial heterogeneity of the predictor variable that is being segmented. Despite this, it follows that the mismatch between the area covered by image objects or superpixels and the groundtruthing data is even more pronounced. Object-based methods have been advocated in the literature over the last 20 years or so [78,79], due to their perceived strengths [80]. GEOBIA approaches have also been applied to marine datasets for more than ten years [29,30,48,81–85]. However, it should be kept in mind that it is much easier in terrestrial mapping to obtain a holistic view of the ground conditions than in marine mapping. Drop-camera systems with the ability to rotate and tilt [31], thereby capturing a larger expanse of the seabed might be a way to bridge the gap between the area covered by seabed imagery and predictor variables.

A critical issue in this study was the high complexity and spatial heterogeneity of the site, with changes in substrate types at scales of several metres to a few tens of metres. For the towed camera sledge and the drop camera, which were both fitted with HiPAP systems, we assume positioning errors of 5–6 m in approximately 100 m water depth [86]. For the grab samples, the position was recorded as the offset to the side gantry relative to the vessel's GPS. Additional offset from this recorded position due to drift might be on the order of 10 m. Given these estimated positioning accuracies, it is likely that in many cases a sample was associated with incorrect values of the response variables. Achieving high map accuracy requires a positioning accuracy in line with the spatial heterogeneity of the site. In situations like those encountered here, this might mean pushing the positioning accuracy requirements to the limits of what is technically feasible. An alternative might be a simulation approach by which many models were fitted to training data that used samples that were randomly relocated based on a bivariate probability function around the original location [28].

In addition to spatial scale, there is the possibility that temporal scale may have been a source of error due to the time interval between the acquisition of MBES data and subsequent groundtruth sampling (up to three years). Movement of benthic sediment through major storms or currents are common and large moving dune features have been observed at depths > 100 m, particularly where strong tidal currents are present ([87]: maximum ~2.0 ms⁻¹). However, our analysis assumed no change had occurred to sediment distributions between surveys and features sampled in the MBES data were represented in the groundtruth samples. This assumption was based on: the absence of obvious mobile sedimentary features in the bathymetry, such as sandwaves; relatively slow currents predicted across the site (Figure 2C,D); and the minimum depth (94 m) suggesting wave action would be of limited influence. Modelled data from Aldridge et al. [88] supports this assumption, as annual maximum disturbances for this area were <7.5 cm for the most mobile sediment component (sand). Although given the observed heterogeneity of the site, the actual annual maximum disturbance is likely to be even less. Nevertheless, we cannot rule out that some change may have occurred and where possible, avoiding this temporal scale issue is desirable.

4.4. Output

In the previous sections we have discussed how changes to the classification scheme, either by adopting a site-specific scheme or by merging classes, may have improved map accuracy. However, this is often out of control for the map producer, as it is commonly the stakeholders who specify the format of map outputs. Another option would be to aggregate the information contained in five maps into a single map of the modal class associated with spatial information on agreement between predictions. Such ensemble maps have been proposed as an effective way of improving classification performance [89,90]. While this was not supported by our results (Figure 5) and other studies [91], the ensemble map provides additional information on the agreement between the outputs derived with different methods. Intuitively, this could be interpreted as a spatially-explicit measure of confidence [92]. Maps like Figure 8 could fulfil the recommendation to include spatial representations of confidence [20], thereby allowing future map users to interpret the outputs based on their needs. However, a more detailed analysis (Table 3) revealed that map accuracy increased markedly only in the case of complete agreement. Nevertheless, such information is valuable for the map user: for example, if the goal were to investigate or monitor a specific substrate type, then samples could be positioned to target the relevant class in areas of complete agreement. Conversely, if the aim were to improve map accuracy then reclassifying areas of low agreement as hybrid classes, or even acquiring additional samples may be preferred.

In this exercise scientists were tasked with creating a thematic map for the site. However, considering the limited success, an alternate approach may have been to generate predictions as class-specific probability maps or quantitative sediment composition maps [31,93,94]. This could be achieved by generating maps in two stages. Initially, R substrate could be delineated using a presence/absence mapping approach [95,96]. Sediment substrates could then be mapped using methods similar to Misiuk et al. [93] and Mitchell et al. [94]. Spatial predictions like these may better capture the heterogeneity of the seafloor, highlighting areas that are relatively fine or coarse, even though that variation may be within a single EUNIS level III class. Unlike thematic mapping based on predefined substrate classes, quantitative modelling is also able to map rare and unsampled classes [93]. In addition, should stakeholders and other map users require thematic maps these could be calculated from the predictions by applying the desired classification scheme.

4.5. Way Forward

Whether field sampling, sample processing, remote sensing or subsequent classification, the approaches used in seabed mapping are continually improving. Here we have presented a range of methods and discussed some of the main issues and possible solutions as they relate to a heterogenous and complex site. However, this is not an exhaustive list (for a detailed error analysis see Strong [97]).

Improvements to the seabed mapping workflow are not typically great leaps forward (although a small number of exceptions may exist such as single beam echosounders to MBES) but rather a series of incremental changes to each component of the workflow. However, what may work in one location does not always transfer to all environments. In the terrestrial remote sensing community, the creation of benchmark datasets allows users to compare approaches on common data to understand the relative strengths and weaknesses of new innovations [98–100]. For seabed mapping, different benchmark datasets should cover different environments, MBES systems (including multi-spectral backscatter) and sampling methods, with sufficient observations (already split into training and testing data) to allow comparison. The data used in this study are published alongside this manuscript and are accessible for use under Crown Copyright to serve as one such example of a benchmark dataset.

5. Conclusions

The aim of this paper was to test a suite of well-established supervised classification methods on a common dataset. We highlighted the challenges summarising a complex seabed into five thematic classes using a classification scheme that targets harmonisation of maps across marine institutions at a European level. Although the backscatter mosaic is one of the most frequently used [17] and important predictor variables in marine habitat mapping [28,29], it should not be relied on solely to discriminate the subtle differences between sediment types with only fractional differences in sediment composition at such fine scales. Promising new possibilities at various stages of the classification process are currently being developed. These include developments of remote sensing technology and improvements of backscatter processing approaches, including the merging of uncalibrated backscatter mosaics [101], calibrated backscatter [102], multispectral backscatter [103], ARA [104] and HAC [21], acquiring more representative samples by applying rigorous sampling designs [105,106] and the derivation of novel features from the primary MBES data with increased discriminatory power for sediment classification [107,108]. We expect progress to mainly occur through refinement at each stage of the mapping process as discussed above. Alternatively, class-specific probability maps or quantitative sediment composition maps may overcome some of the issues encountered with maps of low thematic accuracy. The rapid uptake of maps has shown their importance, but further comparative exercises like this are required to identify research gaps at the various stages of the classification process focussing on challenging sites. Progress in seafloor substrate and habitat mapping will ultimately improve decision making in marine spatial planning and conservation.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/12/20/3398/s1, Figure S1: Substrate map of Method A. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S2: Substrate map of Method B. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S3: Substrate map of Method C. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S5: Substrate map of Method E. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S5: Substrate map of Method E. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S5: Substrate map of Method E. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S5: Substrate map of Method E. Projection UTM Zone 29N, Ellipsoid WGS 84, Figure S5: Substrate map of Method E. Projection UTM Zone 29N, Ellipsoid WGS 84, Table S1: Contingency table of Method A. OE—omission error; CE—commission error; Table S3: Contingency table of Method B. OE—omission error; CE—commission error; CE=commission error; CE=commission error; CE=commission

Author Contributions: Conceptualization, M.D.; Methodology, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Software, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Validation, M.D., P.J.M., E.O., G.O.A.M.G., T.L.B.; Formal Analysis, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Investigation, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Data Curation, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing—Original Draft Preparation, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Visualization, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Visualization, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing, W.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Visualization, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Visualization, M.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing, W.D., P.J.M., E.O., G.O.A.M.G. and T.L.B.; Writing, Writing, W.D., W.S., Witing, W.D., W.S., W.S

Funding: P.J.M. was funded in part by the UK MPA Programme. M.D. was funded by the MAREANO programme. T.L.B. was funded by the NERC CLASS NC programme and GCRF grant SOLSTICE-WIO (NE/P021050/1). E.O. was funded by the INFOMAR project. G.O.A.M.G. was funded in part by the Belgian Scientific Policy Office (BELSPO-NR BR/143/A2/INDI67).

Acknowledgments: The authors would like to thank the UK MPA Programme for providing the data used in this study and staff aboard the RV *Cefas Endeavour* who participated in sample collection and analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Kenny, A.J.; Jenkins, C.M.; Wood, D.; Bolam, S.G.; Mitchell, P.J.; Scougal, C.; Judd, A.D. Assessing cumulative human activities, pressures, and impacts on North Sea benthic habitats using a biological traits approach. *ICES J. Mar. Sci.* **2018**, *75*, 1080–1092. [CrossRef]
- 2. Thrush, S.F.; Dayton, P.K. Disturbance to marine benthic habitats by trawling and dredging: Implications for marine biodiversity. *Annu. Rev. Ecol. Syst.* **2002**, *33*, 449–473. [CrossRef]
- Halpern, B.S.; Walbridge, S.; Selkoe, K.A.; Kappel, C.V.; Micheli, F.; D'Agrosa, C.; Bruno, J.F.; Casey, K.S.; Ebert, C.; Fox, H.E.; et al. A global map of human impact on marine echosystems. *Science* 2008, *319*, 948–953. [CrossRef]
- 4. Ware, S.; Downie, A.-L. Challenges of habitat mapping to inform marine protected area (MPA) designation and monitoring: An operational perspective. *Mar. Policy* **2019**, *111*, 103717. [CrossRef]
- 5. Rattray, A.J.; Ierodiaconou, D.; Womersley, T. Wave exposure as a predictor of benthic habitat distribution on high energy temperate reefs. *Front. Mar. Sci.* **2015**, *2*, 1–14. [CrossRef]
- 6. Smith, G.; Yesilnacar, E.; Jiang, J.; Taylor, C. Marine habitat mapping incorporating both derivatives of LiDAR data and hydrodynamic conditions. *J. Mar. Sci. Eng.* **2015**, *3*, 492–508. [CrossRef]
- 7. De Moustier, C.; Matsumoto, H. Seafloor acoustic remote sensing with multibeam echo-sounders and bathymetric sidescan sonar systems. *Mar. Geophys. Res.* **1993**, *15*, 27–42. [CrossRef]
- 8. Pratson, L.F.; Edwards, M.H. Introduction to advances in seafloor mapping using sidescan sonar and multibeam bathymetry data. *Mar. Geophys. Res.* **1996**, *18*, 601–605. [CrossRef]
- 9. Kostylev, V.E.; Todd, B.J.; Fader, G.B.J.; Courtney, R.C.; Cameron, G.D.M.; Pickrill, R.A. Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Mar. Ecol. Prog. Ser.* **2001**, *219*, 121–137. [CrossRef]
- 10. Dartnell, P.; Gardner, J.V. Predicting seafloor facies from multibeam bathymetry and backscatter data. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 1081–1091. [CrossRef]
- 11. Brown, C.J.; Blondel, P. Developments in the application of multibeam sonar backscatter for seafloor habitat mapping. *Appl. Acoust.* **2009**, *70*, 1242–1247. [CrossRef]
- 12. Heap, A.D.; Harris, P.T. Geological and biological mapping and characterisation of benthic marine environments—Introduction to the special issue. *Cont. Shelf Res.* **2011**, *31*, S1–S3. [CrossRef]
- 13. Populus, J.; Rodrigues, A.M.; Freitas, R.; Quintino, V.; McGrath, F.; Tempera, F.; Galparsoro, I.; Goncalves, J.; Sanz Alonso, J.L. Preface to "MeshAtlantic: Mapping Atlantic area seabed habitats for better marine management". *J. Sea Res.* 2015, *100*, 1. [CrossRef]
- 14. Lamarche, G.; Lurton, X. Introduction to the Special Issue "Seafloor backscatter data from swath mapping echosounders: From technological development to novel applications". *Mar. Geophys. Res.* **2018**, *39*, 1–3. [CrossRef]
- 15. Lucieer, V.; Lecours, V.; Dolan, M. Charting the Course for Future Developments in Marine Geomorphometry: An Introduction to the Special Issue. *Geosciences* **2018**, *8*, 477. [CrossRef]
- 16. Anderson, J.T.; Holliday, D.V.; Kloser, R.J.; Reid, D.G.; Simrad, Y. Acoustic seabed classification: Current practice and future directions. *ICES J. Mar. Sci.* **2008**, *65*, 1004–1011. [CrossRef]
- 17. Brown, C.J.; Smith, S.J.; Lawton, P.; Anderson, J.T. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuar. Coast. Shelf Sci.* **2011**, *92*, 502–520. [CrossRef]
- Lecours, V.; Devillers, R.; Schneider, D.; Lucieer, V.L.; Brown, C.J.; Schneider, D.C.; Lucieer, V.L.; Brown, C.J.; Edinger, E.N. Spatial scale and geographic context in benthic habitat mapping: Review and future directions. *Mar. Ecol. Prog. Ser.* 2015, 535, 259–284. [CrossRef]
- 19. Lecours, V.; Dolan, M.F.J.; Micallef, A.; Lucieer, V.L. A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 3207–3244. [CrossRef]
- 20. Diesing, M.; Mitchell, P.J.; Stephens, D. Image-based seabed classification: What can we learn from terrestrial remote sensing? *ICES J. Mar. Sci.* 2016, *73*, 2425–2441. [CrossRef]

- 21. Alevizos, E.; Greinert, J. The Hyper-Angular Cube Concept for Improving the Spatial and Acoustic Resolution of MBES Backscatter Angular Response Analysis. *Geosciences* **2018**, *8*, 446. [CrossRef]
- 22. Buscombe, D.; Grams, P.E. Probabilistic Substrate Classification with Multispectral Acoustic Backscatter: A Comparison of Discriminative and Generative Models. *Geosciences* **2018**, *8*, 395. [CrossRef]
- 23. Diesing, M.; Green, S.L.; Stephens, D.; Lark, R.M.; Stewart, H.A.; Dove, D. Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. *Cont. Shelf Res.* **2014**, *84*, 107–119. [CrossRef]
- Galparsoro, I.; Agrafojo, X.; Roche, M.; Degrendele, K. Comparison of supervised and unsupervised automatic classification methods for sediment types mapping using multibeam echosounder and grab sampling. *Ital. J. Geosci.* 2015, 134, 41–49. [CrossRef]
- 25. Hasan, R.C.; Ierodiaconou, D.; Monk, J. Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. *Remote Sens.* **2012**, *4*, 3427–3443. [CrossRef]
- 26. Ierodiaconou, D.; Monk, J.; Rattray, A.J.; Laurenson, L.J.B.; Versace, V.L. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. *Cont. Shelf Res.* **2011**, *31*, 28–38. [CrossRef]
- 27. Lucieer, V.L.; Hill, N.A.; Barrett, N.S.; Nichol, S.L. Do marine substrates "look" and "sound" the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle images. *Estuar. Coast. Shelf Sci.* **2013**, *117*, 94–106. [CrossRef]
- 28. Stephens, D.; Diesing, M. A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PLoS ONE* **2014**, *9*, e93950. [CrossRef]
- Janowski, L.; Trzcinska, K.; Tegowski, J.; Kruss, A.; Rucinska-Zjadacz, M.; Pocwiardowski, P. Nearshore Benthic Habitat Mapping Based on Multi-Frequency, Multibeam Echosounder Data Using a Combined Object-Based Approach: A Case Study from the Rowy Site in the Southern Baltic Sea. *Remote Sens.* 2018, 10, 1983. [CrossRef]
- Ierodiaconou, D.; Schimel, A.C.G.; Kennedy, D.M.; Monk, J.; Gaylard, G.; Young, M.A.; Diesing, M.; Rattray, A.J. 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.* 2018, *39*, 271–288. [CrossRef]
- 31. Kågesten, G.; Fiorentino, D.; Baumgartner, F.; Zillén, L. How do continuous high-resolution models of patchy seabed habitats enhance classification schemes? *Geosciences* **2019**, *9*, 237. [CrossRef]
- 32. Fakiris, E.; Blondel, P.; Papatheodorou, G.; Christodoulou, D.; Dimas, X.; Georgiou, N.; Kordella, S.; Dimitriadis, C.; Rzhanov, Y.; Geraga, M.; et al. Multi-Frequency, Multi-Sonar Mapping of Shallow Habitats—Efficacy and Management Implications in the National Marine Park of Zakynthos, Greece. *Remote Sens.* **2019**, *11*, 461. [CrossRef]
- 33. Gardline Geosurvey Limited. *Lot 8-East of Haig Fras-Acquisition Report;* Unpublished Report; Gardline Geosurvey Limited: Great Yarmouth, UK, 2012; Volume 9062.
- 34. Eggleton, J.; Downie, A.-L. East of Haig Fras rMCZ Post-Survey Site Report; Defra: London, UK, 2017.
- Clare, D.; Downie, A.-L.; Hawes, J.; Langton, B. East of Haig Fras Marine Conservation Zone (MCZ) Monitoring Report 2015; JNCC: Peterborough, UK, 2020. Available online: http://data.jncc.gov.uk/data/0a1035b1-3b40-4ada-a47a-7c8771aff3f1/JNCC-Cefas-29-FINAL-WEB.pdf (accessed on 22 April 2020).
- 36. Long, D. BGS Detailed Explanation of Seabed Sediment Modified Folk Classification. 2006; MESH (Mapping European Seabed Habitats). Available online: https://webarchive.nationalarchives.gov.uk/20101014090013/ http://www.searchmesh.net/PDF/GMHM3_Detailed_explanation_of_seabed_sediment_classification.pdf (accessed on 22 April 2020).
- 37. Tewkesbury, A.P.; Comber, A.J.; Tate, N.J.; Lamb, A.; Fisher, P.F. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens. Environ.* **2015**, *160*, 1–14. [CrossRef]
- Stephens, D.; Diesing, M. Towards quantitative spatial models of seabed sediment composition. *PLoS ONE* 2015, *10*, e0142502. [CrossRef]
- Montereale-Gavazzi, G.; Roche, M.; Lurton, X.; Degrendele, K.; Terseleer, N.; van Lancker, V. Seafloor change detection using multibeam echosounder backscatter: Case study on the Belgian part of the North Sea. *Mar. Geophys. Res.* 2018, 39, 229–247. [CrossRef]
- 40. Li, J.; Tran, M.; Siwabessy, J. Selecting optimal random Forest predictive models: A case study on predicting the spatial distribution of seabed hardness. *PLoS ONE* **2016**, *11*, e0149089. [CrossRef]

- 41. Kursa, M.; Rudnicki, W. Feature selection with the Boruta Package. J. Stat. Softw. 2010, 36, 1–11. [CrossRef]
- 42. The R Foundation. The R Project for Statistical Computing. 2020. Available online: https://www.r-project.org/ (accessed on 22 April 2020).
- 43. Liaw, A.; Wiener, M. Breiman and Cutler's Random Forests for Classification and Regression. 2018. Available online: https://cran.microsoft.com/snapshot/2018-03-27/web/packages/randomForest/index.html (accessed on 22 April 2020).
- 44. Calvert, J.; Strong, J.A.; Service, M.; McGonigle, C.; Quinn, R. An evaluation of supervised and unsupervised classification techniques for marine benthic habitat mapping using multibeam echosounder data. *ICES J. Mar. Sci.* **2014**. [CrossRef]
- 45. Dolan, M.F.J.; Buhl-Mortensen, P.; Thorsnes, T.; Buhl-Mortensen, L.; Bellec, V.K.; Bøe, R. Developing seabed nature-type maps offshore Norway: Initial results from the MAREANO programme. *Nor. Geol. Tidsskr.* **2009**, *89*, 17–28.
- 46. Diesing, M.; Thorsnes, T. Mapping of Cold-Water Coral Carbonate Mounds Based on Geomorphometric Features: An Object-Based Approach. *Geosciences* **2018**, *8*, 34. [CrossRef]
- 47. Le Bas, T. RSOBIA—A new OBIA Toolbar and Toolbox in ArcMap 10.x for Segmentation and Classification. In *GEOBIA 2016: Solutions and Synergies*; Kerle, N., Gerke, M., Lefevre, S., Eds.; University of Twente Faculty of Geo-Information and Earth Observation: Twente, The Netherlands, 2016; p. 4.
- Innangi, S.; Tonielli, R.; Romagnoli, C.; Budillon, F.; Di Martino, G.; Innangi, M.; Laterza, R.; Le Bas, T.; Lo Iacono, C. Seabed mapping in the Pelagie Islands marine protected area (Sicily Channel, southern Mediterranean) using Remote Sensing Object Based Image Analysis (RSOBIA). *Mar. Geophys. Res.* 2019, 40, 333–355. [CrossRef]
- 49. Shepherd, J.D.; Bunting, P.; Dymond, J.R. Operational Large-Scale Segmentation of Imagery Based on Iterative Elimination. *Remote Sens.* **2019**, *11*, 658. [CrossRef]
- 50. Lecours, V.; Devillers, R.; Simms, A.E.; Lucieer, V.L.; Brown, C.J. Towards a framework for terrain attribute selection in environmental studies. *Environ. Model. Softw.* **2017**, *89*, 19–30. [CrossRef]
- 51. Mitchell, P.J.; Downie, A.-L.; Diesing, M. How good is my map? A tool for semi-automated thematic mapping and spatially explicit confidence assessment. *Environ. Model. Softw.* **2018**, *108*, 111–122. [CrossRef]
- 52. Breiman, L. Random forests. Mach. Learn. 2001, 5-32. [CrossRef]
- 53. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]
- 54. Story, M.; Congalton, R.G. Accuracy Assessment: A User's Perspective. *Photogramm. Eng. Remote Sens.* **1986**, 52, 397–399.
- 55. Kuhn, M. Building Predictive Models in R Using the caret Package. J. Stat. Softw. 2008, 28, 1–26. [CrossRef]
- 56. Foody, G.M. Harshness in image classification accuracy assessment. *Int. J. Remote Sens.* **2008**, *29*, 3137–3158. [CrossRef]
- 57. Blott, S.J.; Pye, K. Particle size distribution analysis of sand-sized particles by laser diffraction: An experimental investigation of instrument sensitivity and the effects of particle shape. *Sedimentology* **2006**, *53*, 671–685. [CrossRef]
- 58. Stewart, L.K.; Kostylev, V.E.; Orpin, A.R. Windows-based software for optimising entropy-based groupings of textural data. *Comput. Geosci.* 2009, *35*, 1552–1556. [CrossRef]
- Cooper, K.M.; Bolam, S.G.; Downie, A.-L.; Barry, J. Biological-based habitat classification approaches promote cost-efficient monitoring: An example using seabed assemblages. J. Appl. Ecol. 2019, 56, 1085–1098. [CrossRef]
- 60. Kaskela, A.M.; Kotilainen, A.T.; Alanen, U.; Cooper, R.; Green, S.L.; Guinan, J.C.; Van Heteren, S.; Kihlman, S.; Van Lancker, V.; Stevenson, A. EMODnet Geology Partners Picking up the pieces—Harmonising and collating seabed substrate data for European maritime areas. *Geosciences* **2019**, *9*, 84. [CrossRef]
- 61. Strong, J.A.; Clements, A.; Lillis, H.; Galparsoro, I.; Bildstein, T.; Pesch, R. A review of the influence of marine habitat classification schemes on mapping studies: Inherent assumptions, influence on end products, and suggestions for future developments. *ICES J. Mar. Sci.* **2018**, *76*, 10–22. [CrossRef]
- 62. Folk, R.L. The distinction between grain size and mineral composition in sedimentary-rock nomenclature. *J. Geol.* **1954**, *62*, 344–359. [CrossRef]

- 63. Snellen, M.; Gaida, T.C.; Koop, L.; Alevizos, E.; Simons, D.G. Performance of Multibeam Echosounder Backscatter-Based Classification for Monitoring Sediment Distributions Using Multitemporal Large-Scale Ocean Data Sets. *IEEE J. Ocean. Eng.* **2019**, *44*, 142–155. [CrossRef]
- 64. Fogarin, S.; Madricardo, F.; Zaggia, L.; Sigovini, M.; Montereale-Gavazzi, G.; Kruss, A.; Lorenzetti, G.; Manfé, G.; Petrizzo, A.; Molinaroli, E.; et al. Tidal inlets in the Anthropocene: Geomorphology and benthic habitats of the Chioggia inlet, Venice Lagoon (Italy). *Earth Surf. Process. Landf.* **2019**, *44*, 2297–2315. [CrossRef]
- Gaida, T.; Tengku Ali, T.; Snellen, M.; Amiri-Simkooei, A.; van Dijk, T.; Simons, D. A Multispectral Bayesian Classification Method for Increased Acoustic Discrimination of Seabed Sediments Using Multi-Frequency Multibeam Backscatter Data. *Geosciences* 2018, *8*, 455. [CrossRef]
- 66. Hamilton, E.L. Geoacoustic modeling of the sea floor. J. Acoust. Soc. Am. 1980, 68, 1313–1340. [CrossRef]
- 67. Ferrini, V.L.; Flood, R.D. The effects of fine-scale surface roughness and grain size on 300 kHz multibeam backscatter intensity in sandy marine sedimentary environments. *Mar. Geol.* 2006, 228, 153–172. [CrossRef]
- 68. Lurton, X. An Introduction to Underwater Acoustics; Springer: Berlin/Heidelberg, Germany, 2010.
- 69. Montereale-Gavazzi, G.; Roche, M.; Degrendele, K.; Lurton, X.; Terseleer, N.; Baeye, M.; Francken, F.; Van Lancker, V. Insights into the Short-Term Tidal Variability of Multibeam Backscatter from Field Experiments on Different Seafloor Types. *Geosciences* **2019**, *9*, 34. [CrossRef]
- Gaida, T.C.; van Dijk, T.A.G.P.; Snellen, M.; Vermaas, T.; Mesdag, C.; Simons, D.G. Monitoring underwater nourishments using multibeam bathymetric and backscatter time series. *Coast. Eng.* 2020, 158, 103666. [CrossRef]
- 71. Hasan, R.C.; Ierodiaconou, D.; Laurenson, L.J.B.; Schimel, A.C.G. Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. *PLoS ONE* **2014**, *9*, e97339. [CrossRef] [PubMed]
- 72. Hasan, R.C.; Ierodiaconou, D.; Laurenson, L.J.B. Combining angular response classification and backscatter imagery segmentation for benthic biological habitat mapping. *Estuar. Coast. Shelf Sci.* **2012**, *97*, 1–9. [CrossRef]
- 73. Rzhanov, Y.; Fonseca, L.; Mayer, L. Construction of seafloor thematic maps from multibeam acoustic backscatter angular response data. *Comput. Geosci.* **2012**, *41*, 181–187. [CrossRef]
- 74. Parnum, I.M. *Benthic Habitat Mapping Using Multibeam Sonar Systems;* Curtin University of Technology: Perth, Australia, 2007.
- 75. Schimel, A.C.G.; Beaudoin, J.; Parnum, I.M.; Le Bas, T.; Schmidt, V.; Keith, G.; Ierodiaconou, D. Multibeam sonar backscatter data processing. *Mar. Geophys. Res.* **2018**, *39*, 121–137. [CrossRef]
- Feldens, P.; Schulze, I.; Papenmeier, S.; Schönke, M.; Schneider von Deimling, J. Improved Interpretation of Marine Sedimentary Environments Using Multi-Frequency Multibeam Backscatter Data. *Geosciences* 2018, *8*, 214. [CrossRef]
- 77. Costa, B.M. Multispectral Acoustic Backscatter: How Useful Is it for Marine Habitat Mapping and Management? *J. Coast. Res.* 2019, 35, 1062. [CrossRef]
- Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* 2010, 65, 2–16. [CrossRef]
- Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; Van Der Meer, F.; Van Der Werff, H.; Van Coillie, F.; et al. Geographic Object-Based Image Analysis—Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* 2014, *87*, 180–191. [CrossRef] [PubMed]
- Hay, G.J.; Castilla, G. Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT). In Proceedings of the 1st International Conference on Object-Based Image Analysis (OBIA 2006), Salzburg, Austria, 4–5 July 2006; ISPRS: Hannover, Germany, 2006; Volume XXXVI-4/C42, p. 3.
- 81. Lucieer, V.L. Object-oriented classification of sidescan sonar data for mapping benthic marine habitats. *Int. J. Remote Sens.* **2008**, *29*, 905–921. [CrossRef]
- 82. Montereale-Gavazzi, G.; Madricardo, F.; Janowski, L.; Kruss, A.; Blondel, P.; Sigovini, M.; Foglini, F. Evaluation of seabed mapping methods for fine-scale classification of extremely shallow benthic habitats-Application to the Venice Lagoon, Italy. *Estuar. Coast. Shelf Sci.* **2016**, *170*, 45–60. [CrossRef]
- Lacharité, M.; Brown, C.J.; Gazzola, V. Multisource multibeam backscatter data: Developing a strategy for the production of benthic habitat maps using semi-automated seafloor classification methods. *Mar. Geophys. Res.* 2018, 39, 307–322. [CrossRef]

- 84. Costa, B.M.; Battista, T.A. The semi-automated classification of acoustic imagery for characterizing coral reef ecosystems. *Int. J. Remote Sens.* **2013**, *34*, 6389–6422. [CrossRef]
- 85. Fakiris, E.; Zoura, D.; Ramfos, A.; Spinos, E.; Georgiou, N.; Ferentinos, G.; Papatheodorou, G. Object-based classification of sub-bottom profiling data for benthic habitat mapping. Comparison with sidescan and RoxAnn in a Greek shallow-water habitat. *Estuar. Coast. Shelf Sci.* **2018**, 208, 219–234. [CrossRef]
- Rattray, A.J.; Ierodiaconou, D.; Monk, J.; Laurenson, L.J.B.; Kennedy, P. Quantification of spatial and thematic uncertainty in the application of underwater video for benthic habitat mapping. *Mar. Geod.* 2014, 37, 315–336. [CrossRef]
- Barrie, J.V.; Conway, K.W.; Picard, K.; Greene, H.G. Large-scale sedimentary bedforms and sediment dynamics on a glaciated tectonic continental shelf: Examples from the Pacific margin of Canada. *Cont. Shelf Res.* 2009, 29, 796–806. [CrossRef]
- Aldridge, J.N.; Parker, E.R.; Bricheno, L.M.; Green, S.L.; van der Molen, J. Assessment of the physical disturbance of the northern European Continental shelf seabed by waves and currents. *Cont. Shelf Res.* 2015, 108, 121–140. [CrossRef]
- 89. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* 2007, *28*, 823–870. [CrossRef]
- 90. Du, P.; Xia, J.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple Classifier System for Remote Sensing Image Classification: A Review. *Sensors* 2012, *12*, 4764–4792. [CrossRef]
- 91. Turner, J.A.; Babcock, R.C.; Hovey, R.; Kendrick, G.A. Can single classifiers be as useful as model ensembles to produce benthic seabed substratum maps? *Estuar. Coast. Shelf Sci.* **2018**, 204, 149–163. [CrossRef]
- 92. Diesing, M.; Stephens, D. A multi-model ensemble approach to seabed mapping. J. Sea Res. 2015, 100, 62–69. [CrossRef]
- Misiuk, B.; Diesing, M.; Aitken, A.; Brown, C.J.; Edinger, E.N.; Bell, T. A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using random forest. *Geosciences* 2019, 9, 254. [CrossRef]
- 94. Mitchell, P.J.; Aldridge, J.N.; Diesing, M. Legacy data: How decades of seabed sampling can produce robust predictions and versatile products. *Geosciences* **2019**, *9*, 182. [CrossRef]
- 95. Downie, A.-L.; Dove, D.; Westhead, R.K.; Diesing, M.; Green, S.L.; Cooper, R. Semi-automated mapping of rock in the North Sea. *JNCC Rep.* **2016**, 592, 26. [CrossRef]
- Brown, L.S.; Green, S.L.; Stewart, H.A.; Diesing, M.; Downie, A.-L.; Cooper, R.; Lillis, H. Semi-automated mapping of rock in the Irish Sea, Minches, western Scotland and Scottish continental shelf. *JNCC Rep.* 2017, 609, 29. [CrossRef]
- 97. Strong, J.A. An error analysis of marine habitat mapping methods and prioritised work packages required to reduce errors and improve consistency. *Estuar. Coast. Shelf Sci.* **2020**, 240, 106684. [CrossRef]
- Hu, J.; Jiang, T.; Tong, X.; Xia, G.; Zhang, L. A benchmark for scene classification of high spatial resolution remote sensing imagery. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium, Milan, Italy, 26–31 July 2015; Volume 14, p. 4.
- Helber, P.; Bischke, B.; Dengel, A.; Borth, D. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2019, 17, 2217–2226. [CrossRef]
- 100. Campos-Taberner, M.; Romero-Soriano, A.; Gatta, C.; Camps-Valls, G.; Lagrange, A.; Le Saux, B.; Beaupere, A.; Boulch, A.; Chan-Hon-Tong, A.; Herbin, S.; et al. Processing of Extremely High-Resolution LiDAR and RGB Data: Outcome of the 2015 IEEE GRSS Data Fusion Contest—Part A: 2-D Contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2016, 9, 5547–5559. [CrossRef]
- 101. Misiuk, B.; Brown, C.J.; Robert, K.; Lacharité, M. Harmonizing Multi-Source Sonar Backscatter Datasets for Seabed Mapping Using Bulk Shift Approaches. *Remote Sens.* **2020**, *12*, 601. [CrossRef]
- 102. Fezzani, R.; Berger, L. Analysis of calibrated seafloor backscatter for habitat classification methodology and case study of 158 spots in the Bay of Biscay and Celtic Sea. *Mar. Geophys. Res.* **2018**, *39*, 169–181. [CrossRef]
- 103. Brown, C.; Beaudoin, J.; Brissette, M.; Gazzola, V. Multispectral Multibeam Echo Sounder Backscatter as a Tool for Improved Seafloor Characterization. *Geosciences* **2019**, *9*, 126. [CrossRef]
- 104. Fonseca, L.; Brown, C.; Calder, B.; Mayer, L.; Rzhanov, Y. Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures. *Appl. Acoust.* 2009, 70, 1298–1304. [CrossRef]

- 105. Minasny, B.; McBratney, A.B. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci.* **2006**, *32*, 1378–1388. [CrossRef]
- 106. Stevens, D.L., Jr.; Olsen, A.R. Variance estimation for spatially balanced samples of environmental resources. *Environmetrics* **2003**, *14*, 593–610. [CrossRef]
- 107. Zhao, T.; Lazendić, S.; Zhao, Y.; Montereale-Gavazzi, G.; Pižurica, A. Classification of Multibeam Sonar Image Using the Weyl Transform. In *Image Processing and Communications*; Choraś Michałand Choraś, R.S., Ed.; Springer International Publishing: Cham, Switzerland, 2020; pp. 206–213.
- 108. Trzcinska, K.; Janowski, L.; Nowak, J.; Rucinska-Zjadacz, M.; Kruss, A.; von Deimling, J.S.; Pocwiardowski, P.; Tegowski, J. Spectral features of dual-frequency multibeam echosounder data for benthic habitat mapping. *Mar. Geol.* 2020, 427, 106239. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).