

Seafloor mapping based on multibeam echosounder bathymetry and backscatter data using Object-Based Image Analysis: a case study from the Rewal site, the Southern Baltic

by

Łukasz Janowski^{1,*}, Jarosław Tęgowski¹,
Jarosław Nowak²

DOI: [10.1515/ohs-2018-0024](https://doi.org/10.1515/ohs-2018-0024)

Category: **Original research paper**

Received: **October 4, 2017**

Accepted: **January 16, 2018**

¹*Institute of Oceanography, University of Gdańsk, Al. M. Piłsudskiego 46, 81-378 Gdynia, Poland*

²*Maritime Institute in Gdańsk, ul. Długi Targ 41/42, 80-830 Gdańsk, Poland*

Abstract

Seafloor mapping is a fast developing multidisciplinary branch of oceanology that combines geophysics, geostatistics, sedimentology and ecology. One of its objectives is to isolate distinct seabed features in a repeatable, fast and objective way, taking into consideration multibeam echosounder (MBES) bathymetry and backscatter data. A large-scale acoustic survey was conducted by the Maritime Institute in Gdańsk in 2010 using Reson 8125 MBES. The dataset covered over 20 km² of a shallow seabed area (depth of up to 22 m) in the Polish Exclusive Economic Zone within the Southern Baltic. Determination of sediments was possible based on ground-truth grab samples acquired during the MBES survey. Four classes of sediments were recognized as muddy sand, very fine sand, fine sand and clay. The backscatter mosaic created using the Angular Variable Gain (AVG) empirical method was the primary contribution to the image processing method used in this study. The use of the Object-Based Image Analysis (OBIA) and the Classification and Regression Trees (CART) classifier makes it possible to isolate the backscatter image with 87.5% overall and 81.0% Kappa accuracy. The obtained results confirm the possibility of creating reliable maps of the seafloor based on MBES measurements. Once developed, the OBIA workflow can be applied to other spatial and temporal scenes.

Key words: habitat mapping, multibeam echosounder, image processing, Object-Based Image Analysis, Classification and Regression Trees, Southern Baltic, Angle Varied Gain, feature selection

* Corresponding author: ocelj@ug.edu.pl

Introduction

The rapidly evolving seabed mapping technology, related to multibeam echosounders and sidescan sonars, generates large amounts of data, which require time-consuming traditional analyses. For this reason, there is a need for automatic classification of sediment and seabed morphological forms, using the analysis of reflected acoustic signals from the bottom (e.g. Tęgowski & Łubniewski 2002), analysis of images created by the bathymetric model of the seabed, or maps of backscattered intensity of acoustic signals (e.g. Montereale Gavazzi et al. 2016). The analysis of seabed sediment composition, as a branch of marine geology, has recently had more and more in common with hydroacoustic swath measurements. Multibeam echosounder (MBES) surveys allow large areas of the seafloor to be investigated, which is not comparable with traditional sampling technologies. MBES collects two kinds of acoustic data: a bathymetric map and a backscatter intensity image of seafloor areas. While the former generates a precise bathymetric Digital Elevation Model (DEM) that represents seafloor geomorphology, the latter is equal to seafloor acoustic reflectivity, which is characterized by scattering of the acoustic wave by the seabed type (Lurton & Lamarche 2015).

Beyond the MBES bathymetry and backscatter characteristics of the seafloor, other important information on the seafloor is provided by traditional ground-truth samples acquired at point locations. In order to obtain a meaningful map of the seabed, samples often need to be processed at various stages of seafloor mapping. According to Diesing et al. (2016), seafloor mapping usually includes the following steps: pre-processing of a dataset, feature extraction, feature selection, classification, post-classification and evaluation of classifier performance. In this study, pre-processing of data involves processing of hydroacoustic data. The next two steps are described together in "Image processing". Although this study deals with the broad application of Object-Based Image Analysis ("Image processing"), the use of segmentation-classification algorithms was sufficient to create a final map of seabed sediment composition, and therefore no post-classification of data was necessary. The final step, which involves an accuracy assessment, is described in "Accuracy assessment" section.

The dataset presented in this paper comes from the survey that was conducted by the Maritime Institute in Gdańsk in 2010 on board the r/v Imor ship. Multibeam echosounder measurements and acquisition of ground-truth samples covered 20 km² of the seafloor

in the southern part of the Baltic Sea. Analysis of ground-truth samples allows to define four classes of the seabed, three of which were characterized by overlapping backscatter distribution. Despite difficulties, it was possible to process the data and obtain a good classification performance, which was confirmed by an accuracy assessment based on the error matrix.

Study Site

The hydroacoustic survey was conducted in the area located in the Polish Exclusive Economic Zone (EEZ), in the southern part of the Baltic Sea, north of the coastal village of Rewal. The location and basic parameters of the area are shown in Figure 1.

According to Mojski (1995), the area is covered by recent marine fine- and medium-grained sands, mud and clay from different phases of the Baltic Sea development. General geomorphological features are the result of the glacier impact during the last glaciation, especially the deglaciation and current marine processes that have formed seabed deposits since the Littorina transgression until now (Pieccka 1980; Gudelis & Jemielianow 1982). Geomorphologic structures include accumulation or abrasive plains, as well as shoals with varying thickness of recent

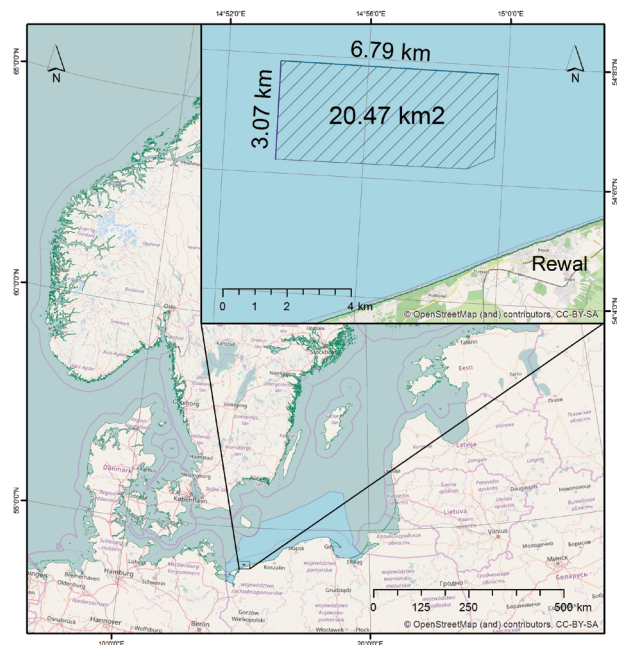


Figure 1

Location of the Rewal study site within the Polish EEZ. Sources: our study, OpenStreetMap, European Environment Agency

marine sand sediments. Clay in abrasive plain areas is often exposed, whereas muddy-clayey sediments from terminal post-glacial basins also occur locally. Accumulation plains are usually filled with muddy-clayey sediments from various phases of the Baltic Sea development, partly covered by sandbanks and sand waves, related to marine transgressions and recent processes that have not yet been fully understood (Piecza 1980; Gudelis & Jemielianow 1982).

Materials and methods

Data acquisition and processing

Data acquisition was performed from board of the research vessel Imor of the catamaran type, equipped with the multibeam echosounder (MBES) Reson Seabat 8125, the Trimble AG132 GPS positioning receiver (accuracy below 1 m), and the DMS-05 motion reference unit (MRU).

MBES transmitted one wide acoustic beam with dimensions of $120 \times 0.5^\circ$ and receives 248 beams, allowing to perform swath measurements with a maximum extent of 120 m width. With an operating frequency of 455 kHz, it covers a maximum swath width of $3.4 \times$ operating depth. In perfect environmental conditions, it is possible to perform measurements with a resolution of up to 0.02 m. Backscatter signals registered by Reson Seabat 8125 MBES have relative values. The shape of the angular dependence of the backscattered acoustic signal is strongly correlated with the sediment composition and seafloor morphology. Therefore, it was assumed that unitless features of angular dependence are still strongly correlated with the type of the bottom. All measurements were acquired and processed using the QPS QINSy (ver. 7.5) software.

In order to properly interpret results of hydroacoustic measurements, they need to be supported by ground-truth data. A total of 19 samples were acquired during the hydrographic survey, using a Van Veen grab sampler and a VKG-3 vibrocorer. Due to practical reasons, locations of all samples were determined in target areas, focusing on all expected sediment classes. Table 1 gives a brief description of all ground-truth samples with their locations.

Hydroacoustic data

Bathymetric and backscatter data from MBES measurements were processed using the QPS QINSy (ver. 8.15) software with a grid resolution of $0.5 \times$

Table 1

List of all ground-truth samples with their descriptions and locations. Sources: our study

ID	Sediment type	Class	Type	Latitude	Longitude
VR 201	Very Fine Sand	VFS	training	54.105263	14.911653
VR 202	Fine Sand	FS	training	54.108030	14.911664
VR 203	Muddy Sand	MS	training	54.120947	14.911780
VR 204	Muddy Sand	MS	validation	54.127028	14.911643
VR 205	Very Fine Sand	VFS	training	54.111783	14.926790
VR 206	Very Fine Sand	VFS	training	54.114840	14.926752
VR 207	Fine Sand	FS	training	54.126588	14.926762
VR 208	Very Fine Sand	VFS	validation	54.120305	14.942225
VR 209	Fine Sand	FS	validation	54.110749	14.942321
VR 210	Very Fine Sand	VFS	validation	54.106412	14.972621
VR 211	Very Fine Sand	VFS	validation	54.122082	14.972553
VR 212	Muddy Sand	MS	validation	54.130222	14.972554
VR 214	Fine Sand	FS	validation	54.116665	14.987440
VR 215	Very Fine Sand	VFS	training	54.111829	14.987819
VR 216	Very Fine Sand	VFS	validation	54.109825	14.987764
VR 225	Muddy Sand	MS	training	54.125345	14.896156
VR 226	Very Fine Sand	VFS	training	54.118745	14.896162
VR 227	Clay	C	training	54.109162	14.896139
VR 228	Muddy Sand	MS	training	54.106042	14.896208

0.5 m. Bathymetric processing was conducted in accordance with international standards, i.e. removal of acoustic artefacts. On the other hand, the backscatter mosaic was created using the combined QINSy-ArcGIS method.

In order to obtain a mosaic of backscattering intensity independent of incidence angles, the empirical method of Angle Varied Gain (AVG) was applied in the Rewal study area. AVG considers a defined number of seafloor scans in order to calculate an averaged angular response curve (QPS 2015). The curve is calculated based on averaged parts of backscatter values, separated for defined ranges of incidence angles. The outcome value of the corrected backscatter is then calculated based on values computed for a defined normalization range that should cover a portion of incidence angles, for which the angular response curve is assumed to be flat for all types of the seabed. Depending on the survey and research, the normalization range can take different sizes, from narrow (Lamarche et al. 2011) to wide

(Fonseca et al. 2009). In this research, good results have been achieved by applying angles between 40 and 60 degrees. In addition, the AVG filter was set to separate angles between 0.1 degree and to calculate an angular response curve based on 300 pings.

All backscatter intensity grids were then exported to a GeoTIFF format. The ArcGIS software was used to create an image mosaic from these files, which can be created from a dataset of grids that often overlap (Blondel 2009). Therefore, a few parameters of mosaicking must be defined. The following parameters were defined in this study: the “north–west” mosaic method, the “ascending” order and the “mean” mosaic operator (Esri 2016). The mosaic method defines which grid will be placed as the top image in the mosaic. The “north–west” mosaic method means that the image closest to the northwest corner of the mosaic would be positioned at the top. Another parameter, the “ascending” order, defines the type of sorting of image grids. There is also an option of reverse sorting, depending on the specific arrangement of image grids. The “mean” mosaic operator means that pixel values of all overlapping areas will be calculated as an average from all overlapping grids. Methods and parameters of mosaicking used in this study are obviously related to a specific arrangement of backscatter image grids in this dataset (Esri 2016).

Image Processing

Feature extraction and selection

A review of marine habitat mapping studies confirm that good predictive performance can often be achieved by considering secondary bathymetric and backscatter features and their proper selection (Diesing et al. 2016). Secondary features of bathymetry, like slope, rugosity (Sappington et al. 2007), aspect or the Bathymetric Position Index (BPI; Micallef et al. 2012), as well as secondary features of backscatter, like commonly used Grey Level Co-occurrence Matrices (GLCM; Haralick et al. 1973) increase the dimensionality, which means that they may capture greater variability in primary data and they may add unique geospatial information to the data (Diesing et al. 2016). Sixteen bathymetric secondary features were extracted in this study: rugosity, slope, variance, aspect, northness, eastness, curvature, profile curvature, planar curvature, small-scale BPI, broad-scale BPI, bathymetry standard deviation, rugosity standard deviation, slope standard deviation, small-scale BPI standard deviation and broad-scale BPI standard deviation. The BPI allows to measure differences in DEM among specific locations

and their neighborhood in relation to the overall bathymetric area (Wilson et al. 2007). The general attribute of BPI is its scale, which corresponds to certain regions with different morphological sizes. The units of scale are expressed in meters.

An additional number of 9 backscatter secondary features were extracted in this study, including: backscatter standard deviation, GLCM parameters as mean, standard deviation, entropy, homogeneity, contrast, correlation, angular second moment and dissimilarity. In order to obtain a good classification accuracy, it was necessary to select important and eliminate correlated features. From many available techniques of feature selection, the embedded algorithm of Classification and Regression Trees (CART) was used in this study (Breiman et al. 1984). The applied feature selection algorithm attempts to find an optimal subset of features during the training part of the CART classifier. Statistical importance of a certain geostatistical or textural feature is computed for all splits of this feature in the decision tree. The results of the algorithm are presented in an array, showing the decisive power of features in a range within [0, 1], called the importance scores (Breiman et al. 1984).

Object-Based Image Analysis

The constantly developing Object-Based Image Analysis is a relatively young branch in the recent marine habitat mapping literature (i.e. Lucieer 2008; Che Hasan et al. 2012a; Lucieer et al. 2013; Diesing et al. 2014; Montereale Gavazzi et al. 2016). Image objects allow to analyze geospatial data, considering not only pixel-based information, but also other measures of objects, like their texture, geometry, statistics and hierarchy, which may be especially useful in high-resolution geophysics and hydroacoustic data with large noise. The advantage of image objects is that it resembles the way people look at images (Hay & Castilla 2006).

In general, the OBIA consists of two steps. The first step is to create image objects (or segments), applying segmentation algorithms. The second step is to classify generated segments, using various classification methods.

The multiresolution segmentation algorithm, available in the Trimble eCognition software (Benz et al. 2004), is the main method that was used in this study to create image objects, similarly to some other object-based marine habitat mapping studies (i.e. Lucieer 2008; Lucieer & Lamarche 2011; Che Hasan et al. 2012). Image objects in this approach are created from one-pixel objects using the bottom-up region merging technique, based on their primary

features such as greyscale or shape (Benz et al. 2004). Each merging step is performed on a pair of adjacent image objects with the lowest increase in heterogeneity. Scale is the main parameter of multiresolution segmentation, i.e. it is responsible for stopping the process of fusion of image objects after reaching the homogeneity criterion (Benz et al. 2004). This criterion can also be expressed as the minimum standard deviation of heterogeneity, which is defined as the relation between color, shape, compactness and smoothness of image objects. The parameters are grouped into weighted pairs: color/shape and smoothness/compactness. In this study, the color parameter corresponds to the relative value of backscatter intensity within the considered image object. The shape parameter consists of two remaining parameters: smoothness and compactness. While smoothness corresponds to the ratio between the border length of an image object and its bounding box, the compactness is related to the ratio between the border length of an image object and the square root of the pixel count inside the image (Benz et al. 2004). Parameters of both weighted pairs can be assigned to values from 0.1 to 0.9 and the total value of each weighted pair is equal to 1. In practice, the eCognition software allows to define values of two parameters: shape and compactness, which are related to the other two parameters (i.e. shape 0.3 and compactness 0.6 mean the following values of all 4 parameters: shape 0.3, color 0.7, compactness 0.6, smoothness 0.4). Detailed equations of all heterogeneity parameters, including the scale parameter, are described in Benz et al. 2004. The scale parameter defines the size of created image objects in such a way that if it is larger, more objects can be merged together and larger objects (defined by heterogeneity parameters) can be enlarged (Baatz & Schäpe 2000). Therefore, the scale parameter is not characterized by any units.

In order to obtain reliable results, many scales of multiresolution segmentation have been tested, including intervals of 50 in the range of 50–1500. The following methodological steps (described further in this section) have been performed for the resulting dataset, up to the accuracy assessment step. Based on the best results of the error matrices, the most promising range of scale was selected to repeat the test for a more detailed scale interval. For the resulting narrower range of multiresolution segmentation scales, between 200 and 750, an interval of 20 was applied. At the end, including previous testing, for some specific ranges of scales, the results allow us to analyze the scale interval of 10 (for example, for scales 240, 250 and 260). Other multiresolution segmentation

parameters – shape and compactness were set to 0.1 and 0.5, respectively. The same values were applied in other marine habitat mapping studies using the multiresolution segmentation method (Lucieer et al. 2013; Diesing et al. 2014; Montereale Gavazzi et al. 2016). Once created, segments can be classified based on specific features of objects using more or less complex classification algorithms.

Five various methods of supervised classification were tested in this study, one of which gave the best results. The majority of them belong to machine learning techniques and they include: Classification and Regression Trees (Breiman et al. 1984), Random Forest (Breiman et al. 2001), Support Vector Machine (Cortes & Vapnik 1995), K-Nearest Neighbor and Bayes. The common part of all algorithms is the two-step classification scheme: training and application of the classifier. The training means that the classifier algorithm learns the relationships between the backscatter image and the labelled ground-truth training data. The application means that the classifier uses the inferred function to map unclassified areas implicitly (Mehryar et al. 2012).

Ground-truth data

Ground-truth samples were processed using granulometric analysis and sieves with mesh sizes of 16.0, 8.0, 4.0, 2.0, 1.0, 0.5, 0.25 and 0.125 mm. On the basis of grain-size composition, the sediments were classified according to the Wentworth scale (Wentworth 1922). For the purpose of supervised classification, ground-truth data were divided into two subsets: training and validation data. Of the total number of 19 samples, 11 samples (58%) were selected as a training subset and 8 samples (42%) as a validation subset in the most representative way possible in terms of the distribution of backscatter intensity values. The training subset was used as an input for the feature selection algorithm and supervised classification. The validation subset was used as an input to assess the accuracy of classification results.

Accuracy assessment

Accuracy assessment was applied to evaluate the classification results. Error matrices were calculated for all classes showing cross-tabulation between the classification results and ground-truth samples of each class (Foody 2002). The following ordinary measures of accuracy assessment were defined: user's and producer's accuracy, overall and Kappa accuracy (Cohen 1960). User's accuracy is the relation between correctly classified objects and all ground-truth

possibilities. Producer's accuracy is the relation between correctly classified reference pixels and all classified pixels (Story & Congalton 1986; Congalton 1991). The overall accuracy is equal to the sum of all correctly classified instances in relation to all instances in the error matrix, whereas Kappa takes into account the possibility of the agreement occurring by chance (Cohen 1960).

Results

Seabed composition classes

Based on the granulometric analysis of ground-truth samples and applying the Wentworth scale (Wentworth 1922), we distinguished 4 different classes of sediments: clay (C), muddy sand (MS), very fine sand (VFS) and fine sand (FS). The position of each sample was compared to its backscatter intensity characteristics, so it was possible to determine basic statistical parameters of backscatter intensity averaged for all classes. The result of the analysis is presented as a boxplot shown in Figure 2. Backscatter intensity values are presented in unitless numbers, as they were given by the manufacturer of the echosounder (see "Data acquisition and processing").

Although there was only one sample of clay, all boxplot statistics for this class are equal. Nevertheless, the backscatter intensity for this class was the highest and amounted to 3826.68, making the class of clay distinctly different from the other seabed composition classes. Medians of other classes were similar, i.e.

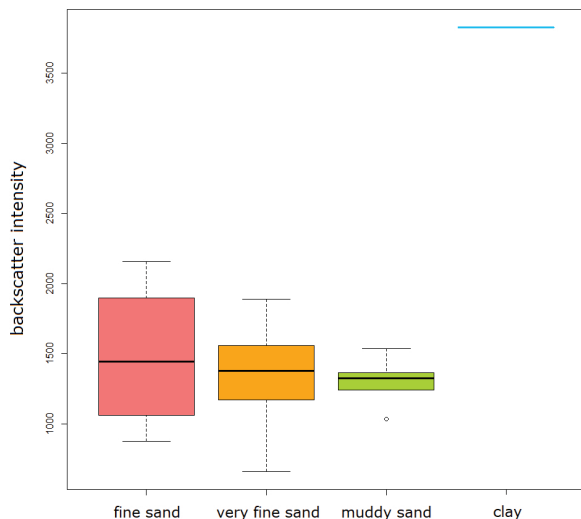


Figure 2
Distribution of the mean backscatter intensity in four ground-truth classes. Sources: our study

1324.48 for muddy sand and 1445.44 for fine sand. The main difference was then visible in the boxplot spread for all classes. The class of fine sand has the thinnest spread. The spread increases in the case of the very fine sand class, being the widest in the class of muddy sand. There is an overlapping distribution of backscatter intensity between the three mentioned classes. The class of fine sand is completely contained in the two remaining classes. After omitting the outstanding whiskers, the class of very fine sand is also contained in the muddy sand class. Therefore, except for the clay class, it is not possible to clearly separate the backscatter mosaic based on the statistical characteristics of the sediment classes. In this case, the use of secondary features of bathymetry and backscatter as well as advanced methods of classification is justified.

Object-Based Image Analysis and supervised classification

The resulting bathymetry of the Rewal study site is presented in Figure 3A. The workflow described in section "Hydroacoustic data" enables the creation of the corresponding backscatter intensity layer presented in Figure 3B. The same figure presents the location of samples. According to section "Image processing", 25 secondary features of bathymetry and backscatter were extracted. They were treated by the embedded feature selection algorithm of CART (Classification and Regression Trees; Breiman et al. 1984) together with primary features to select the most relevant features. In this study, the result of the feature selection algorithm indicates two most important layers: backscatter and rugosity (Figs 3B and C). They reached the importance scores of 0.65 and 0.35, respectively (in the range from 0 to 1). Other features were not relevant, so they were not included in further analysis.

Various scales of multiresolution segmentation were tested and the best result was obtained for the scale parameter of 280. Out of the five tested methods of supervised classification, Classification and Regression Trees gave the best result (Fig. 3D). As shown in Figure 4, the CART algorithm produced a decision tree that separates the training samples from backscatter and rugosity layers. The plot shows the prediction that the C (clay) class is to be assigned to intensity values of the backscatter layer higher than or equal to 2788.15. For lower values, it is necessary to consider the secondary feature – rugosity. If the value of rugosity is lower than 1.59×10^{-6} (-5.7977 in the logarithmic scale, according to Figure 4), then the object should be classified as the MS (muddy

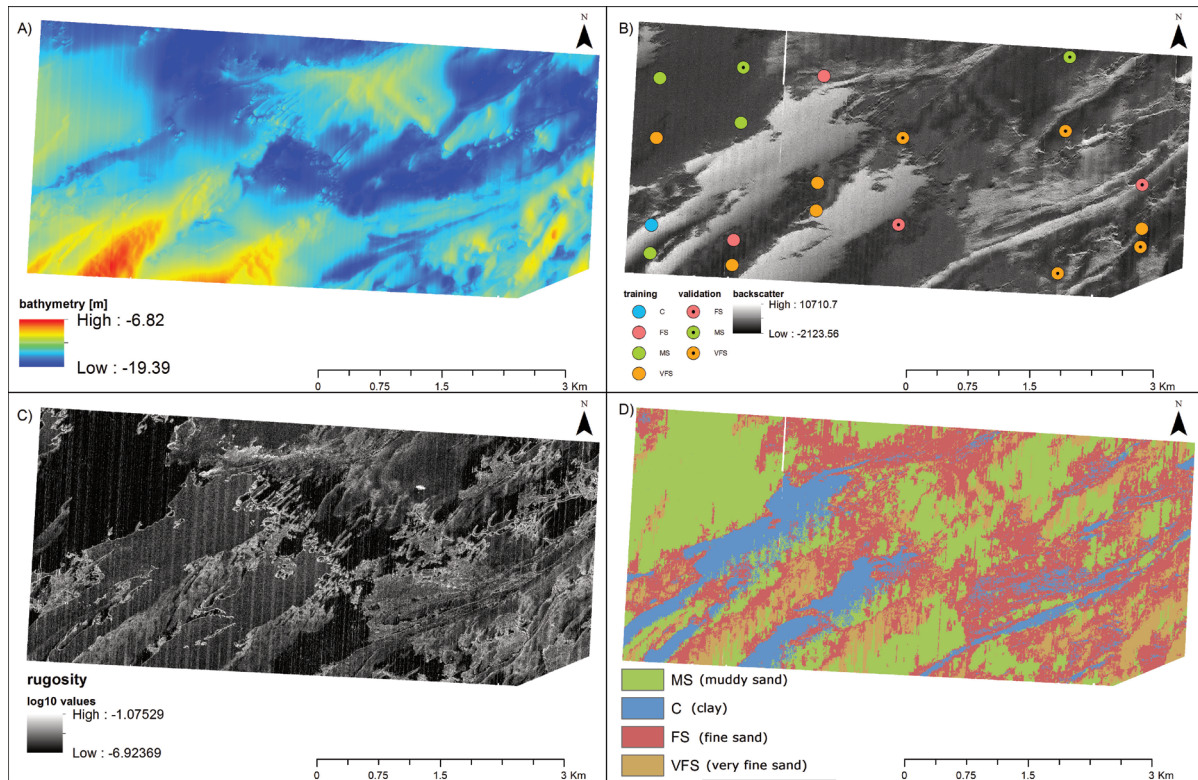


Figure 3

A) result of bathymetric processing; B) result of backscatter processing and location of ground-truth samples, division into training and validation ground-truth samples; C) result of rugosity extracted as a secondary derivative of bathymetry; D) result of CART classification based on Object-Based Image Analysis. Sources: our study

sand) class. If the value is higher or equal, then the backscatter layer should be considered again. For values higher than or equal to 1299.09, the FS (fine sand) class should be assigned. Lower values should be matched with the VFS (very fine sand) class. The numbers and percentage values of the plot are related to the amount and ratios of the training samples corresponding to each class (Fig. 4).

Validation of data processing results

The habitat map generated using the OBIA and CART classifier is shown in Figure 3D. The accuracy assessment of this map, based on the validation of ground-truth samples, is shown in the error matrix (Table 2). It is worth noting that there was no sample for the clay (C) class validation, so producer's and user's accuracies were not determined for this class. Cells of the error matrix that belong to the C class were preserved due to possible misclassification. The error matrix shows that all validation samples were classified properly, except one – i.e. VFS misclassified as MS. This obviously has an effect on the overall accuracy – 87.5%, and Kappa coefficient – 81.0%.

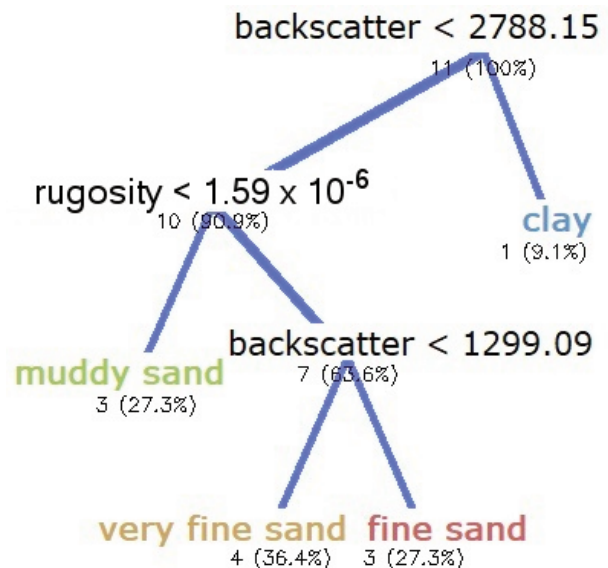


Figure 4

CART results. Sources: our study

Table 2

Error matrix of the CART classification based on the OBIA technique. MS – Muddy Sand, C – Clay, FS – Fine Sand, VFS – Very Fine Sand. Sources: our study

User	Reference Class				Sum
	MS	C	FS	VFS	
MS	2	0	0	1	3
C	0	0	0	0	0
FS	0	0	2	0	2
VFS	0	0	0	3	3
Sum	2	0	2	4	

Producer	1	undefined	1	0.75
User	0.666667	undefined	1	1
Overall Accuracy	0.875			
KIA	0.809524			

Discussion

The world's seafloor has been mapped only in 5–10% with a resolution comparable to onshore research (Wright & Heyman 2008), mainly in areas deeper than 10 m (Montereale-Gavazzi et al. 2016). Although the main device used for hydroacoustic measurements is a multibeam echosounder (MBES), bathymetry is not the most important feature for seabed mapping. The MBES provides additional information about the seafloor reflectivity, or backscatter of the returned acoustic signal, which is crucial for seabed mapping, but its availability is even more limited than that of bathymetric data. There are still no standards for MBES backscatter acquisition and processing, which is why hydroacoustic data collected during one survey is practically incomparable with others (Diesing et al. 2016).

This study deals with the application of a complete methodological approach in order to process the multibeam echosounder backscatter data in one study. The methodology presented in this paper assumes the use of the QINSy-ArcGIS processing method, Object-Based Image Analysis and additional data from ground-truth samples. The in situ approach allows realistic and consistent categorical seafloor mapping (Lurton & Lamarche 2015).

Backscatter grids from the multibeam echosounder are often created using various implementations of the Geocoder engine (Fonseca et al. 2009). This engine makes it possible to apply the Angle Varied Gain (AVG) correction to minimize the ship's along-track nadir effect and to maximize the contrast of the backscatter intensity image. Among others, the Geocoder engine is implemented in two most common commercial

developments: Fledermaus Geocoder Toolbox (FMGT), and CARIS HIPS and SIPS software (Lurton & Lamarche 2015). Both solutions are often used in habitat mapping. For example, FMGT was recently used to create backscatter images in the North Sea (Stephens & Diesing 2014), the Belgian part of the North Sea (Montereale Gavazzi et al. 2017), and the northern China (Li et al. 2017). Raw backscatter data were also processed using CARIS HIPS and SIPS, for example in the areas of Georges Bank in Canada (Brown et al. 2011), Maltese Islands (Micallef et al. 2012), the Tasman Peninsula (Lucieer et al. 2013) and the Lagoon of Venice (Montereale-Gavazzi et al. 2016; Madricardo et al. 2017). For comparison, in this research we applied the new Geocoder engine implementation in the QINSy software. All AVG and mosaicking parameters were tuned and selected carefully. Although the angular dependence of backscatter strength (Parnum 2007; Lurton & Lamarche 2015) could not be completely eliminated, the backscatter mosaic with angular dependence correction, created using this method, gave satisfactory results compared to the backscatter grid without any angular correction. The quality of backscatter mosaic was good enough to perform Object-Based Image Analysis on seabed sediments.

The scale of multiresolution segmentation is a key parameter that may have a major impact on the accuracy of results (Benz et al. 2004). This phenomenon has been observed at a small scale in the Tasman peninsula object-based habitat mapping by comparing the accuracies for 30 and 60 scales (Lucieer et al. 2013). The multiresolution segmentation scale is defined as imperfect, therefore it is necessary to perform many segmentation-classification tunings in order to achieve the desired effect. In order to clearly define a value of the scale parameter, some statistical trials of its estimation were made (Drăguț et al. 2010; Drăguț et al. 2014). Nevertheless, multiresolution segmentation results should be visually assessed (Diesing 2016). However, after some unsuccessful applications of the ESP/ESP2 tool in this study, we have decided to use an iterative method of scale selection, as described in "Materials and methods".

Although it is recommended to acquire and use as many ground-truth samples as possible, statistically more than 50 samples per class (Carlotto 2009), for practical and budget reasons this is not always possible during seabed mapping (Diesing et al. 2016). Therefore, some studies present classification results after using a much smaller but representative number of ground-truth samples (Micallef et al. 2012; Montereale Gavazzi et al. 2017). A similar situation occurs in this study. Nevertheless, attention should be paid to potential sources of errors when many

significantly different classes of the seabed occur (Micallef et al. 2012; Diesing et al. 2016).

In marine habitat mapping studies, the distribution of backscatter intensity between ground-truth classes usually shows some scattering (Stephens & Diesing 2014; Montereale Gavazzi et al. 2017). In such a case, the correct use of advanced supervised classifiers should not cause any difficulties and the results may represent a high level of accuracy. The situation is more complex when the backscatter distribution of most classes overlaps. The results of this study show that the proper application of segmentation and classifier algorithms in the Object-Based Image Analysis can contribute to a high accuracy.

The exceptional accuracy of OBIA results in land-cover remote sensing is not always directly reflected in marine habitat mapping (Diesing et al. 2014; Montereale Gavazzi et al. 2016). There are many reasons for this, often related to difficulties in conducting research in a significantly different environment. Nevertheless, the quality of seabed mapping studies is constantly improving due to adaptation of good practices from land cover mapping, greater awareness of image processing techniques, and increased use of geostatistical secondary features (Diesing et al. 2016; Lecours et al. 2016; Li et al. 2016). One of them is rugosity that was selected in this study by the CART embedded feature selection algorithm. As mentioned in section "Image processing", rugosity is a secondary feature of bathymetry, developed in this study on the basis of the methodology proposed by Sappington et al. (2007). The presented study confirms that even for a dataset with fuzzy boundaries between classes in terms of the distribution of backscatter intensity, the use of other secondary geostatistical layers can lead to a good accuracy assessment. Therefore, it is important to generate such features and use them in the proper selection.

In the past, different decision trees were commonly used in seabed mapping (Dartnell & Gardner 2004; Rattray et al. 2009; Ierodiaconou et al. 2011; Che Hasan et al. 2012a; Che Hasan et al. 2012b; Huang et al. 2012; Stephens & Diesing 2014; Montereale Gavazzi et al. 2016), using both pixel-based and object-based approaches. The term "accuracy" is typically associated with the predictive performance of a map and a reference ground-truth dataset (Foody 2002). A literature review of 20 publications on marine habitat mapping conducted by Diesing et al. (2016) shows that a validation subset of ground-truth data was used in 77% of the analyzed studies. The overall accuracy and the Kappa index were most frequently used. No error matrix was found in 2/3 of the studies.

Che Hasan et al. (2012a,b) used the Quick, Unbiased and Efficient Statistical Tree (QUEST), a decision tree method based on image segmentation, resulting in the overall accuracy of 80.2% and the Kappa statistic of 67%. The object-based CART method was investigated by Montereale Gavazzi et al. (2016), but the results were not described, because they were less promising compared to the K-NN (K-Nearest Neighbor) classifier. The remaining studies represent a pixel-based application of decision tree classifiers. For example, Dartnell & Gardner (2004) show results with the overall accuracy of 72%. The Kappa index was not calculated in this study. Ierodiaconou et al. (2001) evaluated the pixel-based type of the QUEST decision tree with the overall accuracy of 80% and the Kappa statistic of 75%. The classification tree (CT) was used by Stephens & Diesing (2014) with the same overall accuracy and a lower Kappa coefficient (48%, but the highest of all tested classifiers). Huang et al. (2012) used the Boosted Decision Tree (BDT), a slightly different pixel-based method to calculate seabed sediment parameters. The error matrix and the overall or Kappa measures were not included in this case. Depending on the study, decision trees produced different results. The high accuracy obtained in this research (overall accuracy = 87.5%, Kappa = 81%) confirms the high usefulness of the Object-Based Image Analysis and the CART classifier. The range of the Kappa index is equal to $[-\infty, 1]$, and the result below 0.8 is interpreted as good, while close to 1 as very good.

Conclusions

The paper presents the general workflow of backscatter mosaic processing with the angular dependence correction, using common processing software for a multibeam echosounder dataset. For the first time in the peer-reviewed literature, we have described the QINSy-ArcGIS method for multibeam backscatter data processing, including the application of the AVG correction of the Geocoder engine in the QINSy software. Depending on the size of the dataset, the QINSy-ArcGIS method can be time-consuming, but in some cases it may be the only way to produce or reprocess a good quality backscatter mosaic based on limited or few-years old backscatter MBES data. Once created, it can be successfully fitted to other MBES datasets, reducing the total processing time.

The results of this research confirm the high usefulness of the methods of Object-Based Image Analysis and decision trees, including Classification and Regression Trees (CART). Tuned and properly applied Object-Based Image Analysis can be used

as a powerful tool to create a meaningful map of seabed sediments, even for ground-truth classes with overlapping backscatter distribution. In this case, another layer beyond backscatter can improve the classification accuracy. The use of OBIA, including CART and the feature selection tool, produces good accuracy assessment results (overall accuracy = 87.5%, Kappa = 81%). Therefore, we propose to add the CART algorithm to other classification methods used in future comparative studies.

Acknowledgements

This work was technically supported by the Maritime Institute in Gdańsk, responsible for hydroacoustic and ground-truth data acquisition, and the Polish Geological Institute – National Research Institute in Warsaw, responsible for ground-truth data processing on request of the Maritime Authority in Szczecin. The research was funded by the Maritime Authority in Szczecin, pursuant to the agreement OW/5120/06/10 of 09.07.2010.

References

- Baatz, M. & Schäpe, A. (2000). Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In J. Stobl, T. Blaschke & G. Griesebner (Eds.), *Angewandte Geographische Informations – Verarbeitung XII* (pp. 12–23). Karlsruhe: Wichmann Verlag.
- Benz, U., Hofman, P., Willhauck, G., Lingenfelder, I. & Heyen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *Journal of Photogrammetry and Remote Sensing* 58(3): 239–258. DOI: 10.1016/j.isprsjprs.2003.10.002.
- Blondel, P. (2009). *The Handbook of Sidescan Sonar*. Springer Praxis: Heidelberg.
- Breiman, L. (2001). Random Forests. *Machine Learning* 45(1): 5–32. DOI: 10.1023/A:1010933404324.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984). *Classification and Regression Trees*. Belmont: Wadsworth.
- Brown, C.J., Todd, B.J., Kostylev, V.E. & Pickrill, R.A. (2011). Image-based classification of multibeam sonar backscatter data for objective surficial sediment mapping of Georges Bank, Canada. *Continental Shelf Research* 31(2): 110–119. DOI: 10.1016/j.csr.2010.02.009.
- Carlotto, M.J. (2009). Effect of errors in ground truth on classification accuracy. *International Journal of Remote Sensing* 30(18): 4831–4849. DOI: 10.1080/01431160802672864.
- Che Hasan, R., Ierodiaconou, D. & Laurenson, L. (2012a). Combining angular response classification and backscatter imagery segmentation for benthic biological habitat mapping. *Estuarine, Coastal and Shelf Science* 97: 1–9. DOI: 10.1016/j.ecss.2011.10.004.
- Che Hasan, R., Ierodiaconou, D. & Monk, J. (2012b). Evaluation of Four Supervised Learning Methods for Benthic Habitat Mapping Using Backscatter from Multi-Beam Sonar. *Remote Sensing* 4(11): 3427–3443. DOI: 10.3390/rs4113427.
- Cohen, J. (1960). A coefficient of agreement of nominal scales. *Educational and Psychological Measurement* 20(1): 37–46. DOI: 10.1177/001316446002000104.
- Congalton, R.G. (1991). A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment* 37(1): 35–46. DOI: 10.1016/0034-4257(91)90048-B.
- Cortes, C. & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning* 20: 273–297. DOI: 10.1023/A:1022627411411.
- Dartnell, P. & Gardner, J.V. (2004). Predicting seafloor facies from multibeam bathymetry and backscatter data. *Photogrammetric Engineering and Remote Sensing* 70: 1081–1091. DOI: 10.14358/PERS.70.9.1081.
- Diesing, M., Green, S.L., Stephens, D., Lark, R.M., Stewart, H.A. et al. (2014). Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. *Continental Shelf Research* 34: 107–119. DOI: 10.1016/j.csr.2014.05.004.
- Diesing, M. (2016). Application of geobias to map the seafloor. In *GEOBIA 2016: Solutions and Synergies, 14–16 September 2016* (pp. 3). Enschede, Netherlands: ITC / University of Twente.
- Diesing, M., Mitchell, P. & Stephens, D. (2016). Image-based seabed classification: what can we learn from terrestrial remote sensing? *ICES Journal of Marine Science*. 73(10): 2425–2441. DOI: 10.1093/icesjms/fsw118.
- Drăguț, L., Csillik, O., Eisank, C. & Tiede, D. (2014). Automated parametrisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing* 88(100): 119–127. DOI: 10.1016/j.isprsjprs.2013.11.018.
- Drăguț, L., Tiede, D. & Levick, S.R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24(6): 859–871. DOI: 10.1080/13658810903174803.
- Esri. (2016, February 5). *Mosaic dataset properties*. Retrieved June 27, 2017, from <http://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/mosaic-dataset-properties.html>
- Fonseca, L., Brown, C., Calder, B., Mayer, L. & Rzhano, Y. (2009). Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and Multibeam angular signatures. *Applied Acoustics* 70(10): 1298–1304. DOI: 10.1016/j.apacoust.2008.09.008.
- Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80(1): 185–

201. DOI: 10.1016/S0034-4257(01)00295-4.
- Gudelis, W.K. & Jemielianow, J.M. (1982). *Geologia Morza Bałtyckiego*. Warszawa: Wyd. Geologiczne. (In Polish).
- Haralick, R.M., Shanmugam, K. & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* SMC-3(6): 610–621. DOI: 10.1109/TSMC.1973.4309314.
- Hay, G.J. & Castilla, G. (2006). Object-Based Image Analysis: Strengths, Weaknesses, Opportunities, and Threats (SWOT). In 1st International Conference on Object-based Image Analysis, 4–5 July 2006 (pp. 3). Salzburg, Austria: International Society for Photogrammetry and Remote Sensing.
- Huang, Z., Nichol, S.L., Siwabessy, J.P., Daniell, J. & Brooke, B.P. (2012). Predictive modelling of seabed parameters using multibeam acoustic data: a case study on the Carnarvon Shelf, Western Australia. *International Journal of Geographical Information Science* 26: 283–307. DOI: 10.1080/13658816.2011.590139.
- Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L. & Versace, V.L. (2011). Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustic and video observations. *Continental Shelf Research* 31: 28–38. DOI: 10.1016/j.csr.2010.01.012.
- Lamarche, G., Lurton, X., Verdier, A.L. & Augustin, J.M. (2011). Quantitative characterization of seafloor substrate and bedforms using advanced processing of multibeam backscatter. Application to the Cook Strait, New Zealand. *Continental Shelf Research* 31(2): 93–109. DOI: 10.1016/j.csr.2010.06.001.
- Lecours, V., Dolan, M.F.J., Micallef, A. & Lucieer, V.L. (2016). A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences* 20: 3207–3244. DOI: 10.5194/hess-20-3207-2016.
- Li, J., Tran, M. & Siwabessy, J. (2016). Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. *Plos ONE* 11(2): 1–29. DOI: 10.1371/journal.pone.0149089.
- Li, D., Tang, C., Xia, C. & Zhang, H. (2017). Acoustic mapping and classification of benthic habitat using unsupervised learning in artificial reef water. *Estuarine, Coastal and Shelf Science* 185: 11–21. DOI: 10.1016/j.ecss.2016.12.001.
- Lucieer, V. (2008). Object-oriented classification of sidescan sonar data for mapping benthic marine habitats. *International Journal of Remote Sensing* 29(3): 905–921. DOI: 10.1080/01431160701311309.
- Lucieer, V. & Lamarche, G. (2011). Unsupervised fuzzy classification and object-based image analysis of Multibeam data to map deep water substrates, Cook Strait, New Zealand. *Continental Shelf Research* 31: 1236–1247. DOI: 10.1016/j.csr.2011.04.016.
- Lucieer, V., Hill, N.A., Barrett, N.S. & Nichol, S. (2013). So marine substrates 'look' and 'sound' the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle. *Estuarine, Coastal and Shelf Science* 117: 94–106. DOI: 10.1016/j.ecss.2012.11.001.
- Lurton, X. & Lamarche, G. (2015, May). *Backscatter measurements by seafloor-mapping sonars – Guidelines and Recommendations*. Retrieved June 27, 2017, from <http://geohab.org/publications/>
- Madricardo, F., Fogliani, F., Kruss, A., Ferrarin, C., Pizzeghello, N.M. et al. (2017). High resolution multibeam and hydrodynamic datasets of tidal channels and inlets of the Venice Lagoon. *Scientific Data* 4: 170121. DOI: 10.1038/sdata.2017.121.
- Mehryar, M., Rostamizadeh, A. & Talwalkar, A. (2012). *Foundations of Machine Learning*. Cambridge: The MIT Press.
- Micallef, A., Le Bas, T.P., Huvenne, V.A., Blondel, P., Hühnerbach, V. et al. (2012). A multi-method approach for benthic mapping of shallow coastal areas with high-resolution Multibeam data. *Continental Shelf Research* 39–40: 14–26. DOI: 10.1016/j.csr.2012.03.008.
- Mojski, J.E. (1995). *Atlas Geologiczny Południowego Bałtyku*. Sopot-Warszawa: PIG. (In Polish).
- Montereale Gavazzi, G., Madricardo, F., Janowski, L., Kruss, A., Blondel, P. et al. (2016). Evaluation of seabed mapping methods for fine-scale classification of extremely shallow benthic habitats – Application to the Venice Lagoon, Italy. *Estuarine, Coastal and Shelf Science* 170: 45–60. DOI: 10.1016/j.ecss.2015.12.014.
- Montereale Gavazzi, G., Roche, M., Lurton, X., Degrendele, K., Terseleer, N. et al. (2017). Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian part of the North Sea. *Marine Geophysical Research* 1–19. DOI: 10.1007/s11001-017-9323-6.
- Parnum, I.M. (2007). *Benthic habitat mapping using multibeam sonar systems*. Perth, Australia: Curtin University.
- Pieczka, F. (1980). *Geomorfologia i Osady Denne Bałtyku Południowego. Perribalticum - Problemy Badawcze Obszaru Bałtyckiego*. Gdańsk: Ossolineum. (In Polish).
- QPS. (2015, March 31). *QINSy 8.10.2015.03.31.1 Release Notes*. Retrieved June 27, 2017, from <https://confluence.qps.nl/display/dwn/QINSy+8.10.2015.03.31.1+Release+Notes>
- Rattray, A., Ierodiaconou, D., Monk, J., Versace, V.L. & Laurenson, L.J.B. (2013). Detecting patterns of change in benthic habitats by acoustic remote sensing. *Marine Ecology Progress Series* 477: 1–13. DOI: 10.3354/meps10264.
- Sappington, J.M., Longshore, K.M. & Thomson, D.B. (2007). Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management* 71(5): 1419–1426. DOI: 10.2193/2005-723.
- Stephens, D. & Diesing, M. (2014). A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. *Plos ONE* 9(4): 1–14. DOI: 10.1371/journal.pone.0093950.
- Story, M. & Congalton, R.G. (1986). Accuracy assessment:

A user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52: 397–399.

- Tęgowski, J. & Łubniewski, Z. (2002). Seabed characterisation using spectral moments of the echo signal. *Acta Acustica / Acustica* 88(5): 623–626.
- Wentworth, C.K. (1922). A Scale of Grade and Class terms for Clastic Sediments. *The Journal of Geology* 30(5): 377–392. DOI: 10.1086/622910.
- Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C & Grehan, A.J. (2007). Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Marine Geodesy* 30: 3–35. DOI: 10.1080/01490410701295962.
- Wright, D. & Heyman, W. (2008). Introduction to the special issue: marine and coastal GIS for geomorphology, habitat mapping, and marine reserves. *Marine Geodesy* 31: 223–230. DOI: 10.1080/01490410802466306.