Estuarine, Coastal and Shelf Science 192 (2017) 57–71

Contents lists available at ScienceDirect

Estuarine, Coastal and Shelf Science

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

Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows

Kristjan Herkül [a], [*], Anneliis Peterson [a], Sander Paekivi [b]

[a] *Estonian Marine Institute, University of Tartu, Mäealuse 14, 12618 Tallinn, Estonia*
[b] *School of Natural Sciences and Health, Tallinn University, Narva mnt 29, 10120 Tallinn, Estonia*

ARTICLE INFO

Article history:
Received 19 October 2016
Received in revised form
17 April 2017
Accepted 19 April 2017
Available online 27 April 2017

Keywords:
Multibeam sonar
Mathematical modeling
Baltic Sea
Seabed mapping
Benthos

ABSTRACT

Both basic science and marine spatial planning are in a need of high resolution spatially continuous data on seabed habitats and biota. As conventional point-wise sampling is unable to cover large spatial extents in high detail, it must be supplemented with remote sensing and modeling in order to fulfill the scientific and management needs. The combined use of *in situ* sampling, sonar scanning, and mathematical modeling is becoming the main method for mapping both abiotic and biotic seabed features. Further development and testing of the methods in varying locations and environmental settings is essential for moving towards unified and generally accepted methodology. To fill the relevant research gap in the Baltic Sea, we used multibeam sonar and mathematical modeling methods – generalized additive models (GAM) and random forest (RF) – together with underwater video to map seabed substrate and epibenthos of offshore shallows. In addition to testing the general applicability of the proposed complex of techniques, the predictive power of different sonar-based variables and modeling algorithms were tested. Mean depth, followed by mean backscatter, were the most influential variables in most of the models. Generally, mean values of sonar-based variables had higher predictive power than their standard deviations. The predictive accuracy of RF was higher than that of GAM. To conclude, we found the method to be feasible and with predictive accuracy similar to previous studies of sonar-based mapping.

1. Introduction

Marine benthic habitat maps are essential tools for marine spatial planning, planning and monitoring of marine protected areas (Ward et al., 1999), carrying out environmental impact assessments related to maritime construction (wind parks, mineral extraction, dredging etc.), and fulfilling the requirements of the European Union habitat directive. The understanding of the distribution and extent of marine habitats has been very limited until the advancement of acoustic sensors. The most widely used benthic sampling devices such as grabs, trawls and underwater video or photography (Eleftheriou and McIntyre, 2005) yield information only from the visited sites, leaving most of the study area unsampled (Herkül et al., 2013). Using only point-wise sampling methods, it is difficult to detect wider spatial distributions of benthic habitats and to determine the actual borders between them. Interpolation of data between sampling points has been used to produce data layers with full spatial coverage. Interpolation helps to depict the general distribution of habitats, but fails to adequately reflect the natural patchiness of seabed when the study area is sparsely and irregularly sampled.

Implementation of acoustic methods (sonars) in benthic habitat mapping has greatly enhanced the quality of mapping products. Sonar is an active hydro-acoustic device, which uses sound waves to determine water depth. Besides depth measurements, sonars can also be used to measure the difference between transmitted and scattered energy – the backscatter intensity (ICES, 2007). As backscatter intensity is highly dependent on the properties of reflecting surface, it can be used to distinguish between different seabed substrate types. Depth and bathymetry-derived variables such as seabed slope, aspect and roughness, together with backscatter intensity, form a valuable set of information for mapping seabed substrate, habitats, and biota (Diesing et al., 2014).

The main approaches of how to derive meaningful seabed substrate and habitat variables from sonar data (backscatter intensity and depth) are (1) visual interpretation by an expert, (2)

---

[*] Corresponding author.
E-mail address: kristjan.herkul@ut.ee (K. Herkül).

http://dx.doi.org/10.1016/j.ecss.2017.04.026
0272-7714/

differentiating seabed types using the substrate specific relationships between incidence angle and backscatter intensity (angular range analysis, ARA), (3) unsupervised classification based on sonar data followed by *in situ* determination of the seabed classes, (4) supervised classification and regression. Visual interpretation and visual classification by an expert was the earliest method and it has proven to be effective where distinct broad scale seabed features display characteristic backscatter responses, or where there are sharp demarcations between neighboring seabed types (Brown et al., 2005; Hasan et al., 2012a). The lack of objectivity, time cost related to manual digitation of polygons of seabed types, and difficulties in demarcation of seabed types in the case of high heterogeneity and smooth transitions between seabed types renders the visual classification infeasible in most use cases (Brown et al., 2004). The drawback of using the ARA method is that the across-swath resolution of the output is low as one side of one ping (from nadir to the outermost beam for both sides) is the unit of classification (Sternlicht and de Moustier, 2003; Fonseca et al., 2009). This is because the shape of the function is created using the whole side of the measured data, so that one substrate type is determined for all the pixels in one side of a ping. If there are several different seabed types present across one side of a ping, no clear signature can be matched, which is especially problematic in shallow heterogeneous areas (Preston, 2009). Both unsupervised classification and supervised classification and regression include segmentation of a scanned area (usually rectangular grid) and calculation of backscatter and bathymetric statistics in the segments. In unsupervised models, the segments are automatically clustered into distinct groups based on similarity of the values of the segment statistics. Non-acoustic ground truth sampling (e.g. underwater video, bottom grab sampling) will then be carried out to assign the actual bottom type to each cluster (Preston, 2009). Compared to unsupervised methods, seabed variables from ground truth sampling (video, grabs) are used as dependent variables in model calibration of supervised models, and values of the seabed variables are predicted based on the values of the segment-based sonar-derived variables (Lucieer et al., 2013). Thus, the relationships between seabed types and sonar variables are directly formalized in supervised modeling and the model predictions are in the same categories as the input data.

Given the peculiarities of different methods described above, the supervised modeling has the potential to deliver results with the highest accuracy. The development of novel machine learning algorithms, like random forest and support vector machines, have further contributed to the success of supervised modeling in sonar based seabed mapping (Hasan et al., 2012b; Lucieer et al., 2013; Stephens and Diesing, 2014). However, the use of supervised mathematical modeling in sonar based mapping is a very recent approach. Scientific studies on using side-scan sonar data to predict seabed types started to emerge around the year 2000 (e.g. Greene et al., 1999; Brown et al., 2002). The studies addressing multibeam sonar-based mapping methods have emerged only during the last decade (e.g. Gonzalez-Mirelis et al., 2011; Lucieer et al., 2013; Hill et al., 2014; Stephens and Diesing, 2014; Lark et al., 2015; Montereale Gavazzi et al., 2016). Given the novelty of this methodological approach, studies replicated in a multitude of environmental settings, using various sonar systems, and applying different modeling algorithms are needed to further elaborate this approach. Furthermore, to date, the main commercially available software for mapping seabed substrate types is based on the ARA method (Geocoder module, Fonseca and Calder, 2005) that has low across-swath resolution and cannot be directly used for mapping seabed biota. Further studies are needed to enable generalizations about methodical aspects (cell size, sonar equipment, sonar based variables, mathematical algorithms) in relation to environmental settings of a study site (depth, properties of substrate and biota).

In addition to contributing to the overall knowledge base, this study is aimed to advance regional expertise in sonar based mapping of seabed. Regardless of the recent advancements in mathematical modeling methods and increased usage of modern multibeam sonars in hydrographic surveys, sonar based mapping studies that address seabed substrate and biota, are very scarce in the Baltic Sea. The primary methods that have been used for mapping include *in situ* sampling (underwater video, bottom grabs, scuba diving) together with simple interpolation or more sophisticated predictive mathematical modeling to fill in the gaps between sampling sites. The obvious drawback of using only mathematical methods to fill in the gaps between sparsely located sampling sites is that the mathematical methods usually fail to reflect the actual natural patchiness of the seabed. To date we are not aware of any published scientific papers addressing acoustic seabed scanning together with mathematical modeling and *in situ* sampling. There is only a recent study by Bučas et al. (2016) where echograms of a simple single beam echo sounder were used to visually distinguish macrophytes in a shallow lagoon. To fill this gap, this study aimed to:

1) identify sonar-based variables that best explain the distribution of seabed substrate and biotic variables to contribute to the knowledge base for effective and accurate seabed mapping;
2) assess the prediction accuracy of different supervised modeling methods in predicting the distribution of substrate and biotic variables.

In addition to the methodological aims, the study contributes to gaining knowledge about the seabed substrates and biota of the offshore shallows of the northeastern Baltic Sea that have been little studied.

2. Material and methods

2.1. Study area

The Baltic Sea is a tideless and brackish water body. This study was conducted in offshore waters located to the west of the Saaremaa and Hiiumaa Islands, northeastern Baltic Sea (Fig. 1.). The examined areas are exposed to the open Baltic Proper and have a wave fetch of hundreds of kilometers; they are strongly influenced by the environmental conditions of the Baltic Proper. The salinity is between 6 and 7 PSU. The study area covered a depth range of 12–116 m. The chosen areas represent the most shallow areas in western Estonian offshore waters that had not been previously studied. Some of these areas are proposed as potential wind park areas.

As most of the study area is aphotic and macroinvertebrates dominate benthic communities. Based on previous records from similar depths and wave exposure (macrobenthos database of the Estonian Marine Institute), the blue mussel (*Mytilus trossulus*) and hydrozoans of the Hydroidolina subclass are the dominant taxa.

2.2. Acoustic data

A 240 kHz multibeam sonar Reson SeaBat 7101-Flow,