Continental Shelf Research ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Continental Shelf Research

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

Research papers

Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches

Markus Diesing [a,*], Sophie L. Green [b], David Stephens [a], R. Murray Lark [c], Heather A. Stewart [b], Dayton Dove [b]

Q1

[a] Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road, Lowestoft, Suffolk NR33 0HT, United Kingdom
[b] British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA, United Kingdom
[c] British Geological Survey, Environmental Science Centre, Nicker Hill, Keyworth, Nottinghamshire NG12 5GG, United Kingdom

ARTICLE INFO

Article history:
Received 7 January 2014
Received in revised form
30 April 2014
Accepted 3 May 2014

Keywords:
Marine
Benthic
Habitat
Sediment
Mapping
North Sea

ABSTRACT

Marine spatial planning and conservation need underpinning with sufficiently detailed and accurate seabed substrate and habitat maps. Although multibeam echosounders enable us to map the seabed with high resolution and spatial accuracy, there is still a lack of fit-for-purpose seabed maps. This is due to the high costs involved in carrying out systematic seabed mapping programmes and the fact that the development of validated, repeatable, quantitative and objective methods of swath acoustic data interpretation is still in its infancy. We compared a wide spectrum of approaches including manual interpretation, geostatistics, object-based image analysis and machine-learning to gain further insights into the accuracy and comparability of acoustic data interpretation approaches based on multibeam echosounder data (bathymetry, backscatter and derivatives) and seabed samples with the aim to derive seabed substrate maps. Sample data were split into a training and validation data set to allow us to carry out an accuracy assessment. Overall thematic classification accuracy ranged from 67% to 76% and Cohen's kappa varied between 0.34 and 0.52. However, these differences were not statistically significant at the 5% level. Misclassifications were mainly associated with uncommon classes, which were rarely sampled. Map outputs were between 68% and 87% identical. To improve classification accuracy in seabed mapping, we suggest that more studies on the effects of factors affecting the classification performance as well as comparative studies testing the performance of different approaches need to be carried out with a view to developing guidelines for selecting an appropriate method for a given dataset. In the meantime, classification accuracy might be improved by combining different techniques to hybrid approaches and multi-method ensembles.

1. Introduction

Worldwide, the oceans and marginal seas are under increasing pressure from human activities (Halpern et al., 2008) and there is an ever greater need for good seabed habitat maps, both to underpin environmental and socio-economic impact assessments and to assist in the development of effective management measures that will contribute to our responsible stewardship of the marine environment and the sustainable use of its resources. The development of seabed mapping is now driven more by specific policy needs than our innate desire to explore our world. Several global, European and national initiatives aiming at maintaining biodiversity and conserving habitats and species (Convention on Biological Diversity, OSPAR Convention, EU Habitats, Birds and Marine Strategy Framework Directives, UK Biodiversity Action Plan, Marine and Coastal Access Act etc.) require better seabed habitat maps than exist at present to support assessments of the status of the seabed. In Europe, this need is currently addressed in part, through the European Marine Observation and Data Network (EMODNet), which, among other outputs, has so far compiled and harmonised available seabed sediment information to deliver a map of seabed substrates at a scale of 1:1 million. However, the resultant map has a smallest cartographic unit of approximately 4 km^2 and hence might be too generalised for detailed analysis on a more local level, although higher resolution seabed maps will be produced in a subsequent phase.

The advent of swath acoustic techniques has revolutionised seabed mapping science, as we are now able to map the seabed at high spatial resolution and accuracy. The development of swath acoustic systems dates back as early as the 1940s (Kenny et al., 2003) and was initially driven by sidescan sonar. Multibeam echosounders (MBES), with their ability to simultaneously record

* Corresponding author. Tel.: +44 1502 524266; fax: +44 1502 513865.
E-mail address: markus.diesing@cefas.co.uk (M. Diesing).

http://dx.doi.org/10.1016/j.csr.2014.05.004
0278-4343/© 2014 Published by Elsevier Ltd.

Please cite this article as: Diesing, M., et al., Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Continental Shelf Research (2014), http://dx.doi.org/10.1016/j.csr.2014.05.004

bathymetry and backscatter strength, have become the system of choice for detailed, high-resolution seabed mapping (Brown et al., 2011a). Despite the fact that we now have the technical ability to map the seabed with high detail and accuracy, we are a long way from achieving accurate and fit-for-purpose seabed habitat maps. This can be attributed to two key reasons:

First, only a limited number of countries have so far initiated and executed large-scale seabed mapping programmes (e.g. 'Irish National Seabed Survey', 'Integrated Mapping for the Sustainable Development of Ireland's Marine Resource' and 'Marine Area Database for Norwegian waters') due to the high costs involved. However, a larger number of countries are collecting swath acoustic data for hydrographic charting purposes in a systematic way and possess a wealth of legacy data sets, mainly physical seabed sampling (grabs and cores) and seabed observations (videos and stills). Making best use of these available data sets is becoming increasingly important due to limited financial resources. In the United Kingdom, the Marine Environmental Mapping Programme (MAREMAP) aims to achieve common, national objectives in seabed and shallow geological mapping addressing themes such as habitat mapping, Quaternary science, coastal and shelf sediment dynamics and the assessment of human impacts and geohazards in the marine environment. MAREMAP makes use of data that are primarily collected for other purposes (e.g. MBES data of the Civil Hydrography Programme) or already existing (e.g. legacy grain-size data of the British Geological Survey (BGS)).

Second, the development of validated, repeatable, quantitative and objective methods of swath acoustic data interpretation is lagging behind the ability to collect high-quality swath acoustic data. Anderson et al. (2008) identified a lack of statistical, objective procedures as one of the most 'burning issues' of acoustic seabed classification. Expert interpretation of acoustic data 'by eye' is still relatively common. More recently, automated methods have been explored, driven largely by the advantages of using objective classification algorithms, thus minimising subjectivity (Brown et al., 2011a). A variety of approaches has been trialled including artificial neural networks (Marsh and Brown, 2009; Ojeda et al., 2004), Bayesian decision rules (Simons and Snellen, 2009), decision trees (Che Hasan et al., 2012a; Dartnell and Gardner, 2004; Ierodiaconou et al., 2011; Rattray et al., 2009; Rooper and Zimmermann, 2007), support vector machines (Che Hasan et al., 2012b), Random Forest (Che Hasan et al., 2012b; Lucieer et al., 2013), maximum likelihood classifier (Buhl-Mortensen et al., 2009; Che Hasan et al., 2012b; Ierodiaconou et al., 2011), clustering (Blondel and Gomez Sichi, 2009; Brown and Collier, 2008; Brown et al., 2012) and principal component analysis within commercial software QTC Multiview (Brown et al., 2011b; McGonigle et al., 2009; Preston, 2009). However, only three studies have been published that attempt to compare different automated seabed mapping approaches (Che Hasan et al., 2012b; Ierodiaconou et al., 2011; Stephens and Diesing, 2014).

It is against this background that we held a "MAREMAP Acoustic Data Interpretation Workshop" in Edinburgh in October 2012. The workshop was centred on a common data set exercise where MBES and physical ground-truthing data were made available prior to the workshop. Participants were asked to apply their preferred methodology to the data sets and derive a map of seabed substrates. Results were discussed and compared during the workshop. This paper summarises the main outcomes of the common data set exercise. The objectives of the exercise, and hence this paper, were (i) to compare the different methodologies with respect to their thematic accuracy and spatial representation of predicted seabed substrates and (ii) to assess the relative merits and limitations of the different approaches.

2. Data

The area selected for the common data set exercise lies in the western North Sea off the Scottish coast of the United Kingdom

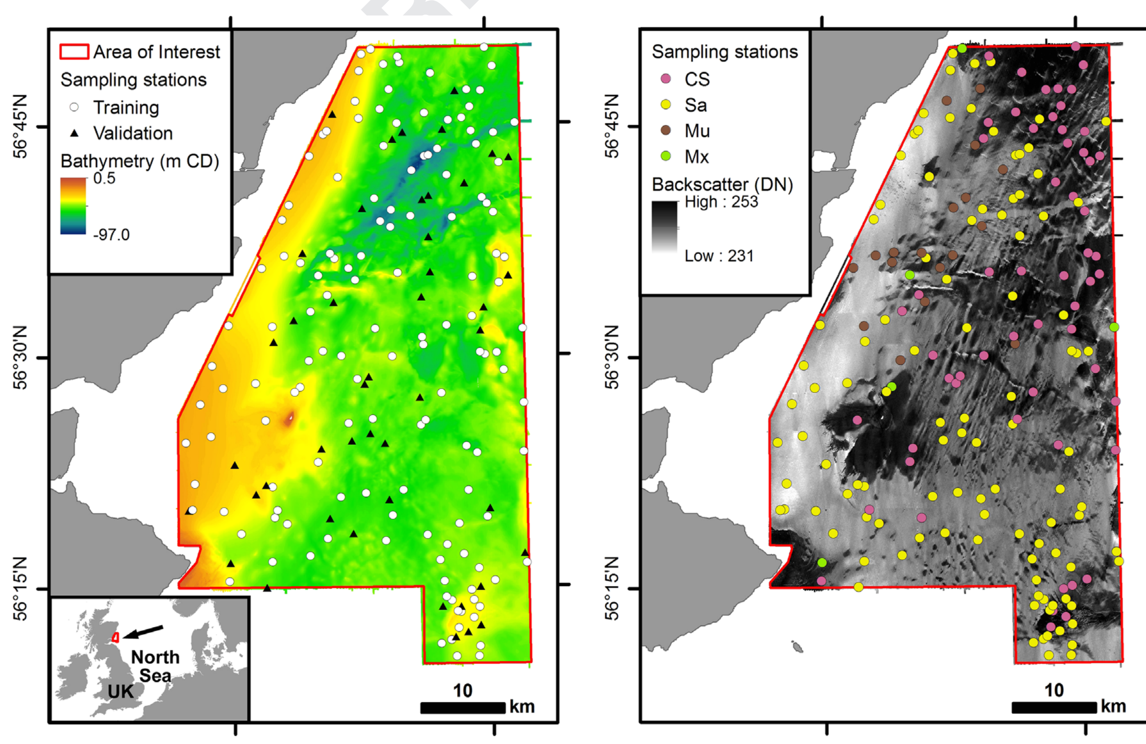

Fig. 1. Area selected for the common data set exercise in the North Sea (inset). Left panel shows bathymetry in relation to Chart Datum (CD) and the locations of training and validation sample data. Right panel shows backscatter strength as digital number (DN) and classified samples (CS—coarse sediment; Sa—sand and muddy sand; Mu—mud and sandy mud; Mx—mixed sediments).

Please cite this article as: Diesing, M., et al., Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Continental Shelf Research (2014), http://dx.doi.org/10.1016/j.csr.2014.05.004