Marine Geology 357 (2014) 37–52

Contents lists available at ScienceDirect

Marine Geology

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

Predictive mapping of seabed substrata using high-resolution multibeam sonar data: A case study from a shelf with complex geomorphology

Huang Zhi *, Justy Siwabessy, Scott L. Nichol, Brendan P. Brooke

Geoscience Australia, GPO Box 378, Canberra, ACT 2601, Australia

A R T I C L E I N F O

Article history:
Received 15 May 2013
Received in revised form 11 July 2014
Accepted 25 July 2014
Available online 4 August 2014

Keywords:
Multibeam
Seabed mapping
Bathymetry
Backscatter
Angular response curves
Feature extraction
Prediction

A B S T R A C T

This study explores the full potential of high-resolution multibeam data for the automated and accurate mapping of complex seabed features under a predictive modelling framework. For an area of seabed on the Carnarvon shelf in Western Australia, morphometric variables and textural measures were derived from multibeam bathymetry and backscatter data. Several feature extraction approaches were applied to backscatter angular response curves to obtain new features. These derivatives and new features were used separately and in combination in the predictions. Despite the complex distribution of various hard substrata within the study area, we achieved a nearly perfect prediction of "hard vs soft" seabed types with an AUC (Area Under Curve) close to 1.0. The predictions were also satisfactory for gravel, sand and mud content, with R^2 values that range from 0.55 to 0.73. This study demonstrates that using a full range of derivatives and new features from both multibeam bathymetry and backscatter data optimises the accuracy of seabed mapping. From the modelled relationships between sediment properties and multibeam data, we confirmed that coarser sediment generally generates stronger backscatter return. Importantly, the results again highlight the advantages of applying proper feature extraction approaches over using original backscatter angular response curves.

1. Introduction

Accurate information on seabed substrata is key for effective benthic habitat mapping (e.g., Lanier et al., 2007; Erdey-Heydorn, 2008; Brown et al., 2011), benthic biodiversity/species prediction (e.g., Brown et al., 2002; Holmes et al., 2008; McGonigle et al., 2009, 2011), and management of marine protected areas (MPAs) (e.g., Huang et al., 2011). Traditionally, seabed information was only available for a limited number of points collected during marine surveys. However, the rapid development of remote sensing technologies provides great potential for automated and accurate mapping of the seabed across large areas. In particular, (active) acoustic remote sensing techniques such as sidescan and multibeam sonar utilise the propagation of acoustic signals through the water column and their return from the seabed interface to map large areas of seabed in water depths up to several thousand metres (e.g., Dartnell and Gardner, 2004; McGonigle et al., 2009; Brown et al., 2011; Huang et al., 2012a). In the last two decades, multibeam sonar (echo-sounder) has become the preferred seabed mapping tool because it can collect simultaneous and co-registered bathymetry and backscatter data (Hewitt et al., 2010; Brown et al., 2011; Micallef et al., 2012). Modern high-frequency multibeam sonar systems transmit pulses of sound and receive backscatter signals from hundreds of narrow-angle beams that generate small footprints on the seabed. Therefore, they can produce bathymetry and acoustic backscatter data with a spatial resolution similar to airborne remotely sensed data (e.g., less than 3 m at a depth up to 150 m).

The overall aim of this study was to map (and classify) an area of geomorphically complex seabed from high-resolution multibeam data using a robust predictive mapping approach. Both unsupervised and supervised classification techniques, as well as hybrid approaches, have previously been used for seabed mapping. Two examples of unsupervised techniques used to classify acoustic data are QTC-Multiview (McGonigle et al., 2009; Preston, 2009) and the CLARA clustering algorithm (Hamilton and Parnum, 2011). In these approaches, data are classified into seabed acoustic classes and ground-truth samples are used to attribute the acoustic classes into meaningful seabed substrata. In contrast, predictive modelling methods such as classification trees, Neural Networks and rule-based approaches have been used to classify acoustic data in a supervised manner (e.g., Dartnell and Gardner, 2004; Zhou and Chen, 2005; Lathrop et al., 2006; Rooper and Zimmermann, 2007; Huang et al., 2012a, 2013). These supervised approaches used ground-truth data to develop a predictive model which was then used to predict the whole study area. In this study, we use the supervised method for its ability to provide reliable accuracy assessments and for investigating modelled relationships between explanatory variables and target variables.

One yet to be fully realised advantage of multibeam data is that we can derive a large number of additional variables from both bathymetry

* Corresponding author at: GPO Box 378, Canberra, ACT 2601, Australia. Tel.: +61 2 62495876.
E-mail address: Zhi.Huang@ga.gov.au (H. Zhi).

http://dx.doi.org/10.1016/j.margeo.2014.07.012
0025-3227/

and backscatter data (Huang et al., 2012a, 2013). Only a few studies have used both multibeam bathymetry and backscatter data for seabed mapping (Mitchell and Hughes Clarke, 1994; Dartnell and Gardner, 2004; Huang et al., 2012a), but not to their full potential. For example, Mitchell and Hughes Clarke (1994) used the backscatter data and two derivatives of the bathymetry data. Dartnell and Gardner (2004) used the backscatter data and three derivatives of the backscatter and bathymetry data. Huang et al. (2012a) used a larger number of derivatives of the backscatter and bathymetry data. None of them, however, used the backscatter angular response curves (see Section 2 for details), which have been demonstrated to be very useful for seabed mapping (e.g., Hughes Clarke, 1994; Fonseca and Mayer, 2007; Hamilton and Parnum, 2011; Huang et al., 2013).

The main objective of this study is to explore the full potential of multibeam data for the automated and accurate mapping of seabed substrata by using a large number of derivatives of bathymetry and backscatter data, and new features obtained from backscatter angular response curves. Although these datasets do not exhaust all possible variables, they represent a full range of information from both primary sources of multibeam data (i.e. bathymetry and backscatter). We hypothesise that, with additional input variables of different sources, the accuracy of the seabed mapping will be improved. Another objective of this study is to demonstrate the value of using this robust modelling technique to quantitatively investigate the underlying relationships between multibeam data and seabed properties. We chose a study area that has complex seabed geomorphology ranging from hard substrate to various soft sediment types to investigate the above two objectives.

2. Background — multibeam data and seabed mapping theory and applications

Multibeam bathymetry and backscatter data, and new variables derived from them provide complementary information for accurate seabed mapping. Their chief utility lies in the capacity to describe seafloor morphology and seabed texture that are proxies of oceanographic processes and seabed physical properties.

2.1. Backscatter

Acoustic backscatter data, acquired by multibeam sonar, records acoustic returns scattered back from the seabed. The recorded backscatter intensity is a function of the absorption and scattering of water and the seabed interface, the angle of incidence and the seafloor topography (De Moustier and Matsumoto, 1993). After radiometric and geometric corrections (Hughes Clark et al., 1996; Mitchell, 1996; Beaudoin et al., 2002), the level of backscatter intensity is largely controlled by three seabed physical properties: the acoustic impedance contrast ("hardness"), the apparent surface roughness (relative to sonar frequency) and volume heterogeneity (Goff et al., 2000; Kloser et al., 2001; Ferrini and Flood, 2006; Parnum et al., 2006; Siwabessy et al., 2006; Fonseca et al., 2009). These three parameters are seabed-type dependent. Studies indicate that acoustic backscatter strength correlates with sediment mean grain size (e.g., Urick, 1983; Davis et al., 1996; Ferrini and Flood, 2006; De Falco et al., 2010). Hard substrata, such as reef, could also be easily differentiated from soft sediment using backscatter data because of their very different hardness (e.g., Lucieer, 2008; Huang et al., 2013).

Processed backscatter data can be represented as backscatter mosaics or backscatter angular response curves. A backscatter mosaic is an accurately registered spatial layer derived from normalising backscatter intensity against backscatter intensity at a chosen incidence angle (or average of several angles). Textural measures such as the first-order variance and second-order GLCMs (Gray Level Co-occurrence Matrices, Haralick et al., 1973) can then be derived from a backscatter mosaic. The textural analysis of backscatter data (most often used for sidescan-derived mosaics) has demonstrated its value in seabed mapping (e.g., Reed and Hussong, 1989; Cochrane and Lafferty, 2002; Lathrop et al., 2006; Lucieer, 2008; Preston, 2009; Huang et al., 2012a). Backscatter angular response curves, however, maintain backscatter information at a full range of incidence angles for individual patches of the seabed. They normally have lower spatial resolution but higher spectral resolution than backscatter mosaics (Huang et al., 2013). New features such as statistical parameters and first derivatives can be extracted from angular response curves for seabed mapping (e.g., De Moustier and Matsumoto, 1993; Hughes Clarke, 1994; Keeton and Searle, 1996; Hughes Clarke et al., 1997; Parnum et al., 2004; Fonseca and Mayer, 2007; Hamilton and Parnum, 2011; Lamarche et al., 2011; Hasan et al., 2012; Huang et al., 2013).

2.2. Bathymetry

Bathymetry information is not only needed for the correction of backscatter data (De Moustier and Matsumoto, 1993) but also provides information on the morphology of the seafloor in unprecedented detail (Goff et al., 1999). Many morphometric measures such as slope, Benthic Position Index (BPI) and curvature can be derived from the bathymetry data (Lundblad et al., 2006; Wilson et al., 2007). They are often good indicators of substrate type. For example, hard substrates such as reefs are often located on high-relief seabed (Dartnell and Gardner, 2004). Seabed morphological features and water depth are also effective proxies for oceanographic processes that influence sediment transport, thus distributions of soft sediment habitats (Nittrouer et al., 1998; Sternberg, 2005).

3. Materials and methods

In this section, we first describe the study area, and the acquisition of the multibeam data (Section 3.1). We then detail the methods used to process the raw multibeam data, obtain derivatives of the bathymetry and backscatter data, and derive new features from the backscatter angular response curves (Sections 3.2–3.5). Finally, we present the predictive modelling techniques for the mapping of the seabed substrate type (Section 3.6). Fig. 1A shows the flowchart of the overall data processing, analysis and predictive modelling steps used in this study. In the first step, processed multibeam data were obtained from calibrating and cleaning the raw data (Section 3.2). Next, GIS and image analysis techniques were used to derive new data and features from the processed multibeam data (Fig. 1B). Finally, both the processed multibeam data and the new derivatives were fed into the predictive models to obtain seabed classification and prediction maps (Section 3.6). The processed multibeam data include bathymetry and backscatter data. From the bathymetry data, we derived a range of terrain variables detailed in Section 3.3. From the backscatter data, we derived 48 mosaics according to the sonar incidence angles; from each mosaic, a number of textural measures were obtained (Section 3.4). In addition, we derived backscatter angular response curves from the mosaics; from these angular response curves, a number of new features were extracted (Section 3.5).

3.1. Study area and survey

The study area covers the width of the continental shelf offshore from Point Cloates, central Western Australia (Fig. 2). Here, the shelf extends ~25 km from Ningaloo Reef and lagoon, a World Heritage-listed area with significant biodiversity value. In 2008, an area covering 281 km^2 was mapped and sampled for the purpose of collecting co-located information on seabed habitats for surrogacy analysis (Brooke et al., 2009). A Kongsberg EM 3002 (300 kHz) multibeam echosounder system was used to acquire both bathymetry and backscatter data across a continuous grid. Sediment samples were collected from 89 representative locations of the seabed using a Smith–McIntyre grab (Fig. 2).