New approaches to predicting the distribution and dynamics of seafloor habitats

by

Alexander James Rattray B.Env.Sc. (Hons)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Deakin University

July, 2013
I am the author of the thesis entitled

New approaches to predicting the distribution and dynamics of seafloor habitats

submitted for the degree of Doctor of Philosophy

This thesis may be made available for consultation, loan and limited copying in accordance with the Copyright Act 1968.

'I certify that I am the student named below and that the information provided in the form is correct'

Full Name:  Alexander James Rattray.

Signed:  

Signature Redacted by Library

Date: 12/07/2013
I certify that the thesis entitled

‘New approaches to predicting the distribution and dynamics of seafloor habitats’

submitted for the degree of Doctor of Philosophy

is the result of my own work and that where reference is made to the work of others, due acknowledgment is given.

I also certify that any material in the thesis which has been accepted for a degree or diploma by any university or institution is identified in the text.

'I certify that I am the student named below and that the information provided in the form is correct'

Full Name: Alexander James Rattray

Signed: [Signature Redacted by Library]

Date: 26/10/2010
Abstract

A general and pervasive decline in the health of marine ecosystems has been brought about by a synergy of human-induced pressures spanning local to global scales. Nowhere are the effects of these pressures more evident than in coastal regions which are home to a large and growing proportion of the world’s population. The paradigm for management of marine resources in these areas is necessarily evolving to one that recognises the importance of interactions between all facets of marine biodiversity and the totality of the surrounding abiotic and anthropogenic environment. Provision of information to government bodies responsible for management of marine resources in the face of often competing environmental, economic and social needs has been historically limited by the costs and technical difficulties associated with acquiring data at the seabed. The evolution of sophisticated acoustic survey technologies, global positioning systems and remote observational methods has fundamentally enhanced our ability to study the seafloor and the biodiversity that exists there. The advances made in the short time that these tools have been available for marine ecological research have accentuated how little is known about the distribution of subtidal marine biodiversity and revealed the considerable potential of multidisciplinary acoustic approaches to address these knowledge gaps.

This thesis was motivated by the recognition that our understanding of the distribution of benthic biodiversity is fragmentary and in general unable to meet the information requirements for effective spatial management of coastal
marine environments. The work presented here addresses a range of issues and applications of hydroacoustic habitat discrimination pertinent to a better understanding of both the distribution of benthic habitats and the manner in which they are characterised. This thesis is organised into a series of four related studies, principally focusing on exploring applications of some of the contemporary sampling tools and analytical techniques used to derive benthic habitat maps, testing assumptions of data quality, and assessing strengths of association between benthic communities and their measurable physical surrogates.

The first research chapter (Chapter 2) presents an analysis of the application of underwater video data collected for training and validating spatially explicit benthic habitat models. Specifically, two major sources of error pertaining to collection of this type of reference data are quantified:

- Spatial error - a spatial error budget was developed for a hybrid positioning system used to co-register video data to their corresponding locations at the seafloor.
- Thematic error - variability in interpretation between trained operators assessing the same video frames was compared between times, over 3 levels of a benthic classification scheme.

Results indicated that potential propagated error in the positioning system described is highly correlated with depth of operation, and varied from 1.5m near the surface to 5.7m in 100m of water. In order of decreasing classification hierarchy, mean overall observer agreement was found to be 98% (range 6%), 82% (range 12%) and 75% (range 17%) for the 2, 4 and 6 class levels of the
scheme respectively. Patterns in between observer variation demonstrate that levels of agreement between observers are related to the level of detail imposed by each hierarchy of the classification scheme, the feature of interest, and also potentially to the level of observer experience.

In Chapter 3 multibeam echosounder (MBES) survey data were integrated with georeferenced video observations to quantify benthic biotic communities at Cape Nelson, Victoria, Australia. Using an automated decision tree (DT) classification approach, 5 representative biotic groups defined from video analysis were related to hydro-acoustically derived variables in the Cape Nelson survey area. Using a combination of MBES bathymetry, backscatter and derivative products produced highest overall accuracy (87%) and kappa statistic (0.83). This study demonstrated that DT classifiers are capable of integrating variable data types for mapping distributions of benthic biological assemblages, which are important in maintaining biodiversity and other system services in the marine environment.

Building on techniques and limitations identified in the first two research chapters, Chapter 4 describes the use of landscape transition analysis as a means to differentiate systematic signals of habitat change separate from the high level of habitat persistence at a shallow (10-50m depth) 18km2 site on the temperate Australian continental shelf between the years 2007 and 2008. Supervised classifications for each year were accomplished using independently collected MBES and video-derived reference data. Of the 4 representative biotic classes considered, signals of directional systematic changes were observed to occur between a shallow kelp dominated class, a
deep sessile invertebrate dominated class and a mixed class of kelp and sessile invertebrates. These signals of change are interpreted as inter-annual variation in the density and depth related extent of a canopy forming kelp species \textit{Ecklonia radiata} at the site, a phenomenon reported in smaller scale temporal studies of the same species. The methods applied in this study provide a detailed analysis of the various components of the traditional change detection cross tabulation matrix allowing identification of the strongest signals of systematic habitat transitions across broad geographic regions. Identifying clear patterns of habitat change is an important first step in linking habitat patterns to the processes that drive them.

The final research chapter investigated the role of wave exposure (defined by a fine-scale numerical wave model) on the distribution of near-shore benthic habitats. Comparison of classifications implemented using the Random Forests algorithm, an ensemble variant of the DT approach used in Chapters 3 and 4, established that significantly more accurate characterisations of habitat were obtained using the exposure model than MBES acoustic data alone. Variable importance measures and map interpretation indicated that the exposure model was most influential in discriminating habitat classes containing the canopy forming kelp \textit{E. radiata} in areas in highly exposed areas. The study demonstrated that combination of ancillary environmental data with hydroacoustic variables is an effective approach to improving model accuracy and validity.
## Contents

Abstract .............................................................................................................................................. i

Contents ............................................................................................................................................... v

List of Figures ....................................................................................................................................... x

List of Tables ....................................................................................................................................... xv

Acknowledgements ........................................................................................................................... xviii

1. General Introduction ............................................................................................................................ 1

   1.1 Background .................................................................................................................................. 2

   1.2 Overview of marine benthic habitat mapping .............................................................................. 5

       1.2.1 Defining habitat .................................................................................................................. 5

       1.2.2 Benthic habitat mapping .................................................................................................. 6

       1.2.3 Tools for benthic habitat mapping .................................................................................... 7

       1.2.4 Integrating acoustic and biological data ........................................................................... 8

       1.2.5 Unsupervised approaches to data aggregation ................................................................ 9

       1.2.6 Supervised approaches to data aggregation .................................................................... 10

   1.3 Thesis focus, objectives and orientation ...................................................................................... 12

2. Sources of uncertainty in video derived reference data. ................................................................. 18

   2.1 Introduction .................................................................................................................................. 19

       2.1.1 Spatial Error ....................................................................................................................... 21

       2.1.2 Thematic error .................................................................................................................... 22
2.2 Methods .................................................................................................. 24

2.2.1 Study Area .................................................................................. 24

2.2.2 Towed video system ................................................................... 26

2.2.3 Spatial Error Propagation ........................................................... 26

2.2.4 Vertical Reference Unit Error .................................................... 30

2.2.5 Gyrocompass Error ..................................................................... 31

2.2.6 USBL Acoustic Error ................................................................. 31

2.2.7 Comparison of Observer Variability .......................................... 33

2.2.8 Sampling Design ........................................................................ 34

2.2.9 Video Attribution Protocol ......................................................... 36

2.2.10 Statistical Analyses ..................................................................... 39

2.3 Results .............................................................................................. 44

2.3.1 Spatial Error Budget ................................................................... 44

2.3.2 Thematic Error ........................................................................... 45

2.3.3 Comparison of between class observer agreement ..................... 46

2.3.4 Tests of Marginal Homogeneity ................................................. 48

2.4 Discussion ....................................................................................... 51

2.4.1 Spatial Error Budget ................................................................... 51

2.4.2 Thematic Error ........................................................................... 55

2.5 Conclusions ..................................................................................... 58

3. Linking bathymetry and backscatter for benthic habitat characterisation 61
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.7</td>
<td>Model validation</td>
<td>102</td>
</tr>
<tr>
<td>4.2.8</td>
<td>Change assessment</td>
<td>102</td>
</tr>
<tr>
<td>4.2.9</td>
<td>Habitat Persistence</td>
<td>103</td>
</tr>
<tr>
<td>4.2.10</td>
<td>Random and systematic transitions</td>
<td>103</td>
</tr>
<tr>
<td>4.3</td>
<td>Results</td>
<td>105</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Temporal habitat classifications</td>
<td>105</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Summary of habitat transitions</td>
<td>107</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Persistence of habitats</td>
<td>111</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Systematic and random transitions</td>
<td>113</td>
</tr>
<tr>
<td>4.4</td>
<td>Discussion</td>
<td>117</td>
</tr>
<tr>
<td>5.</td>
<td>Wave driven exposure as a surrogate for benthic habitat distribution</td>
<td>122</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>123</td>
</tr>
<tr>
<td>5.2</td>
<td>Methods</td>
<td>127</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Study area</td>
<td>127</td>
</tr>
<tr>
<td>5.2.2</td>
<td>MBES data acquisition and processing</td>
<td>129</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Exposure model</td>
<td>131</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Regional scale model parameterisation</td>
<td>131</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Model Calibration</td>
<td>132</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Local scale model parameterisation</td>
<td>134</td>
</tr>
<tr>
<td>5.2.7</td>
<td>Bottom orbital velocity</td>
<td>134</td>
</tr>
<tr>
<td>5.2.8</td>
<td>Observation Data</td>
<td>135</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>5.2.9</td>
<td>Habitat classification</td>
<td>136</td>
</tr>
<tr>
<td>5.2.10</td>
<td>Variable importance measures</td>
<td>137</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>139</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Exposure Model</td>
<td>139</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Classification comparisons</td>
<td>141</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Variable importance measures</td>
<td>145</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>152</td>
</tr>
<tr>
<td>6</td>
<td>Summary and Key findings</td>
<td>159</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Conclusion</td>
<td>167</td>
</tr>
<tr>
<td>7</td>
<td>References cited</td>
<td>169</td>
</tr>
</tbody>
</table>
List of Figures

**Figure 2.1** Site location in Bass Strait, Australia showing artificially
illuminated MBES bathymetry. Major reef systems of the site extend from
Cape Patton (A) and Point Hawdon (B). .......................................................... 25

**Figure 2.2** Positional referencing parameters for gyrocompass, vertical
reference unit and ultra-short baseline system. Gravity reference frame is
shown in blue and vessel reference frame in red............................................. 28

**Figure 2.3** Classification GUI showing: (a.) Video frame grab with satellite
derived time-stamp; (b.) Drop down lists of physical and biological habitat
descriptors; (c.) Positional information and coupled time/date strings used to
georeference classification details for each video frame; (d.) Combined
positional and thematic information appended to spread sheet format and (e.)
aggregated thematic attributes exported to a GIS. ........................................... 37

**Figure 2.4** a) Three level hierarchical classification scheme used in the study
with b). associated decision rules for aggregation of observations ............... 38

**Figure 2.5** Cumulative contribution of each component of the relative
positioning system to total propagated error with depth............................... 45

**Figure 2.6** Proportion of observer agreement between habitat classes using a 4
category classification scheme. Pairwise comparisons within and between
observers are indicated by legend below figure........................................... 47
Figure 2.7 Proportion of observer agreement between habitat classes using a 6
category classification scheme. Pairwise comparisons within and between
observers are indicated by legend below figure.............................................. 47

Figure 3.1 Hill shaded MBES bathymetric coverage of Cape Nelson (Victoria,
Australia) study area with towed video transects overlaid............................ 66

Figure 3.2 Conceptual diagram of data analysis procedures used for the
development of classified habitat maps............................................................. 67

Figure 3.3 (a) Multibeam bathymetry and (b) backscatter intensity products. 69

Figure 3.4 Multibeam derivatives: (a) Complexity; (b) Maximum curvature; (c)
HSI (R); (d) Northness; (e) Slope; (f) HSI (G); (g) Benthic Position Index; (h)
Rugosity; (i) HSI (B)....................................................................................... 72

Figure 3.5 Comparison of averaged User’s and Producer’s accuracy values
obtained from decision trees incorporating; backscatter and derivatives,
bathymetry and derivatives and combined backscatter/bathymetry and
derivatives. Habitat classes are mixed brown algae and invertebrates
(ALG/INV), mixed brown algae (MB), mixed red algae and invertebrates
(MRI), mixed brown and mixed red algae (MBMR) and unconsolidated
sediments (SED).............................................................................................. 80

Figure 3.6 Biotic classification map created using combined bathymetry,
backscatter and derived products. Habitat classes are mixed brown algae and
invertebrates (ALG/INV), mixed brown algae (MB), mixed red algae and
invertebrates (MRI), mixed brown and mixed red algae (MBMR) and
unconsolidated sediments (SED)....................................................................... 83
**Figure 4.1** Site location in Bass strait, Australia showing bathymetric zones overlaying artificially illuminated multibeam bathymetry. Video transect locations for each year are displayed as black (2006) and white (2007) lines. Detail shows bathymetric profile of advancing barchans dune waves between the years 2006 and 2007. 95

**Figure 4.2.** Habitat maps of the study area from a) 2006 and b) 2007 derived using an automated decision tree approach. (See figure 4.1 for analysis extents) 106

**Figure 4.3** Spatial representation of gross gains, gross losses and persistence between the years 2006 and 2007 for classes (a) ALGDOM, (b) ALG/INV, (c) INVDOM and (d) SED. Gains are represented in yellow, losses in red and persistence in blue. 110

**Figure 4.4** Reciprocal systematic gain/loss relationships occurring at the study site between the years 2006 and 2007. 116

**Figure 5.1** Location map of the study site at Cape Otway south-eastern Australia superimposed with hill-shaded bathymetry. Numerals represent each of the four MBES survey blocks undertaken at the site. Letters are associated with detail insets represent major reef systems that are referred to throughout the chapter. 128

**Figure 5.2** Model domain of the regional spectral wave model showing the triangular mesh created from 0.0025° national bathymetry grid. Wave buoy locations are Point lonsdale (Victoria) and Cape Sorrel (Tasmania). Extent of study area (Figure 5.1) are detailed in box. 132
**Figure 5.3.** Results of the Cape Otway coast spectral wave model hindcast for the year 2000. Summaries of significant wave height ($H_{\text{sig}}$) and direction are shown for waters a) west and b) east of the southernmost point of the cape under typical prevailing south-westerly swell conditions. Extents of study area (Figure 5.1) are detailed in box. .......................................................... 133

**Figure 5.4.** Distribution of modelled maximum orbital velocity ($U_{\text{max}}$) values across the Cape Otway study site. Boxes delineate major reef systems at the site and are analogous to those presented in Figure 1. ................................. 140

**Figure 5.5** Out of Bag (OOB) cross-validation error rates plotted against the number of trees permuted for each model used in the study (exposure and acoustic variables in red; acoustic variables only in blue) ................................. 142

**Figure 5.6.** A sample of 400 randomly selected *in situ* video data points were used to illustrate how values from the exposure model ($U_{\text{max}}$) (x-axis) varies with depth (y-axis) and habitat class. ................................................................. 146

**Figure 5.7** Variable importance by mean decrease accuracy of the acoustic model (top) and exposure model (bottom). Variable order is detailed in Table 5.1. Interpretation of habitat classifications ................................................. 147

**Figure 5.8** Habitat characterisations of representative reef systems *west of* Cape Otway overlaid with 10m isobaths and classified observation data from towed video transects. Insets shown are identified in Figures 5.1 and 5.44. Classifications on the left hand side of the figure were derived using acoustic only variables, while those on the right were derived using an additional exposure dataset. ............................................................................................ 150
Figure 5.9 Habitat characterisations of representative reef systems east of Cape Otway overlaid with 10m isobaths and classified observation data from towed video transects. Insets shown are identified in Figures 5.1 and 5.4. Classifications on the left hand side of the figure were derived using acoustic only variables, while those on the right were derived using an additional exposure dataset.
List of Tables

Table 2.1 Manufacturer’s accuracy statements for components of the relative positioning system (95% probability). Acronyms are: DGPS – differential GPS; USBL – ultra short baseline system; VRU – Vertical reference unit. .......... 29

Table 2.2 Levels of experience of observers participating in the study at the first classification attempt (T1), and at the second classification attempt (T2). 33

Table 2.3 General contingency matrix for comparing observer agreement. Values on the major diagonal representing agreement between observers are highlighted in bold. ................................................. 40

Table 2.4 Aggregated general contingency matrix used for McNemar tests of marginal homogeneity (refer to Table 2.3). ......................................................... 43

Table 2.5 Summary of overall agreement, Kappa coefficient (K) of agreement and McNemar tests (p-value) for proportional homogeneity. ...................... 50

Table 3.1 Derivative products from MBES bathymetry and backscatter. ....... 71

Table 3.2 Biotic classes extracted from towed video data collected at the Cape Nelson site, enumerating corresponding training/validation pixels and mean depth distributions (m) (± s.d.). Habitat classes are mixed brown algae and invertebrates (ALG/INV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED). ................................................................. 78

Table 3.3 Confusion matrix for biotic classification of Cape Nelson site, including User’s and Producer’s (underlined italics) accuracies (%). Habitat
classes are mixed brown algae and invertebrates (ALGINV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED).

Table 4.1 Summary of the four category classification scheme employed with associated geo-physical (MBES) descriptive statistics for each category.

Table 4.2. Confusion matrices for the classified images, 2006 and 2007. Each column corresponds to the ground reference pixels used for accuracy assessment for a single class. The values in the column indicate the number of those ground observation pixels classified into each class, while the values on the main diagonal (italicised) indicate agreement between ground points and classified maps.

Table 4.3 Basic change transition matrix (% of study area) for the years 2006 and 2007. The major diagonal of matrix (italicised) represents persistence (no change) of classes between years while other values represent ‘from to’ changes between habitat categories.

Table 4.4 Summary of classification changes between the years 2006 and 2007 expressed as percentage of the study area. Note: \( g_p = \text{gain/persistence}, \quad l_p = \text{loss/persistence}, \quad n_p = \text{net change/persistence} \).

Table 4.5 Inter-category gains 2006 to 2007.

Table 4.6 Inter-category losses 2006 to 2007.

Table 5.1 Variables used in the Random Forest models describing scale of analysis and spatial grain.
Table 5.2 Confusion matrices for the classified images derived from the acoustic model (top) and the exposure model. Each column corresponds to the ground reference pixels used for accuracy assessment for a single class. The values in the column indicate the number of those ground observation pixels classified into each class, while the values on the main diagonal (italicised) indicate agreement between observations and classified maps ....................... 144

Table 5.3 Class area estimations derived from the model containing only acoustic variables (Acoustic) and the model containing both acoustic and exposure variables (Exposure). ................................................................. 148
Acknowledgements

I extend my heartfelt gratitude and appreciation to the many people who generously and selflessly donated their time, expertise and support to help me realise this work.

I gratefully acknowledge the help and support of:

- My principal supervisor Dr Daniel Ierodiaconou who has supplied boundless support, energy and guidance throughout the journey.

- My associate supervisors Assoc. Prof. Laurie Laurenson and Prof. Gerry Quinn for offering pragmatic advice (and sympathy) when it was needed, and for fostering an environment of approachability and mutual support.

- Dr Jac Monk (and many others) at the Warrnambool Geospatial Facility where most of the work in this thesis was completed, for helping to improve the good ideas and identify the bad.

- It is only thanks to the academic, technical and administrative staff at Deakin University’s Warrnambool campus, and their willingness to provide help at every turn that it was possible to complete this project. My sincerest thanks everyone.

- The crew of the Australian Maritime College vessel ‘Bluefin’ and Fugro Survey for their work in collecting the multibeam echosounder data used in this research.
This work was financially supported by National Heritage Trust as part of the Victorian Marine Habitat Mapping Project and an Australian Postgraduate Award scholarship. I gratefully acknowledge the hard work of all those in partner institutions - the Glenelg Hopkins Catchment Management Authority, Department of Sustainability and Environment, Parks Victoria, University of Western Australia and the Coastal CRC.

Shoaib Burq and Marcus Reston from Victorian Partnership for Advanced Computing who helped implement the decision tree methods used in Chapters 3 and 4.

Tim Womersley from Water Technology, who was instrumental in developing the spectral wave model used in Chapter 5.

And finally to my wife and boys who have persevered with and occasionally suffered from my frequent absences, both mental and physical. This wouldn’t have been possible without your unswerving love and support. Thank you for constantly reminding me of the important things in life.
1. General Introduction
1.1 Background

Human reliance on the resources and ecosystem services provided by the world’s oceans is growing as the global population exceeds 7 billion (UNFPA 2011). Regional scale anthropogenic pressures on our oceans through coastal development, overfishing, resource extraction and urban and rural runoff have significantly affected the structure and function of marine systems and their ability to provide essential goods and services (Worm et al. 2006). Overarching this is the mounting evidence that the global effects of human induced climate change such as rising ocean temperatures, ocean acidification and alteration of oceanographic dynamics may impact directly on marine species and communities and also reduce their resilience to regional scale stressors (Poloczanska et al. 2007, Brown et al. 2012a).

Increasing recognition of the value and vulnerability of marine systems has instigated a general paradigm shift in the principles guiding the approach to their conservation and management. Where approaches to marine management have traditionally focused on a single species, sector or activity (Cogan et al. 2009), attention is now being directed at maintaining the health, productivity and resilience of an ecosystem as a whole. The fundamentals of ecosystem based management (EBM) recognise the complex functional dynamics and interactions of species and communities with their physical, chemical and biological (including anthropogenic) environments. Underpinning the set of ecological principles that define EBM is a core concept of ‘place based’
management, as marine ecosystems are in most cases fixed in space (Katsanevakis et al. 2011). A spatially explicit knowledge of the distribution and composition of marine resources within an ecosystem can therefore be seen as a primary requirement for its management. A distinct gap exists however between the information needed to implement the principles of EBM though effectively targeted marine spatial management and our understanding of the distribution of marine biodiversity (Arkema et al. 2006).

A major difficulty faced by managers in developing policy and implementing measures to safeguard ecologically important areas of the oceans is the relative paucity of scientific information available to direct and inform such initiatives. In comparison to terrestrial ecosystems, spatial management of marine ecosystems has been constrained by the lack of high quality, spatially explicit data describing the basic patterns of their biophysical constituents. This is for the most part a function of the inherent difficulties and costs associated with collecting biological data in the marine environment. As a result quantitative spatial information in marine ecosystems is typically sparse, localised and patchily distributed through space and time (Kostylev & Hannah 2007, Foster et al. 2009).

Terrestrial remote sensing of the composition, extent and spatial arrangement of land cover types has advanced considerably over the past few decades concurrent to the availability and ease of access to high resolution, accurately georeferenced spectral imagery. Thematic maps representing land cover, resources and land use are now considered an essential tool supporting spatial management of natural resources across regional, national and international
scales. Spectral methods have also been applied to resource mapping in marine environments where the water column is shallow enough to allow light to be reflected from the seafloor and measured by a range of passive (Mumby & Edwards 2002, Phinn et al. 2005) and active (Wang & Philpot 2007, Costa et al. 2009) electromagnetic sensors.

It is only comparatively recently however, that workers in the marine field have had access similar full coverage datasets at a comparable grain and spatial accuracy to their terrestrial counterparts beyond the shallow sub-littoral (Hughes Clarke et al. 1996, Kenny et al. 2003) The emergence of remotely sensed acoustic technologies coupled with the ability to collect seabed information with georeferenced towed camera systems, opens the possibility of surveying large areas of seafloor and producing high resolution maps of topography, subsurface structures, and benthic habitats. Because many benthic habitats are defined by their geology (along with depth, associated biotic communities and other attributes), geophysical techniques play an important role in determining habitat type. Major developments in seabed mapping technology have occurred that meet many of the requirements of the marine scientific community. Principal among these are multibeam sonar acquisition systems (Anderson et al. 2008, Brown & Blondel 2009), improved navigation and positioning (Philip 2003, Brown et al. 2012b), underwater video systems (Van Rein et al. 2011, Seiler et al. 2012) and advances in digital data processing and interrogation techniques (Ierodiaconou et al. 2011, Hasan et al. 2012). The intersection of technological advancements and the pressing need for detailed information regarding the distribution of marine ecological
resources has led to the rapid evolution of efforts to map distinct geological and biological characteristics of the seafloor at scales relevant to management.

1.2 Overview of marine benthic habitat mapping

1.2.1 Defining habitat

The concept of habitat lies at the core of ecological theory. Ambiguity around its denotation has been the subject of much debate, and repeated calls have been made for a standardised definition in both the terrestrial (Murphy & Noon 1991, Hall et al. 1997) and marine literature (Olenin & Ducrotoy 2006, Dauvin et al. 2008a). In the field of benthic habitat mapping the definition of habitat has evolved to reflect the objectives and applications of the data (Dauvin et al. 2008b). Definitions of habitat from earlier studies underlie the field’s origins in marine geology and geophysics. Acoustic geophysical tools including echosounders and later, multibeam echosounders, allowed the delineation of meaningful geological facies based the acoustic response of the seabed, supported by appropriate physical samples. Benthic taxa often exhibit strong links with seafloor geology, for example kelp species are generally associated with hard reef. The biological component of habitat is therefore often inferred directly under the assumption that geologically defined substrate is the primary determinant of the species and community types that develop there (e.g. Greene et al. 1999). Subsequent adoption of this technology by ecologists for mapping from a biophysical perspective has given rise to an increasingly biocentric notion of habitat (Dauvin et al. 2008a, Brown et al. 2011). This reflects both the needs of natural resource management agencies for primarily biological information, and also the recognition that many other physical,
chemical and biological determinants are also central to patterns of biological
distribution (McArthur et al. 2010). In this thesis habitat is defined inclusive of
its physical, chemical and biological components following Kostylev et al.,
(2001) as:

‘a spatially defined area where the physical, chemical, and biological
environment is distinctly different from the surrounding environment’

1.2.2 Benthic habitat mapping

Diaz et al., (2004) highlight the need to create biologically and ecologically
meaningful maps of marine habitats to fulfil the demands of managers using
increasingly sophisticated strategies to identify and track trends in living
resources. Models depicting the spatial arrangement of the components of
marine ecosystems ranging from the physical (e.g. Harris & Baker 2011) to
species and communities (e.g. Brown et al. 2012b) underpin the notion of
marine spatial planning for EBM (Cogan et al. 2009). By mapping distinct
patterns of the components of ecosystems we acquire requisite information for
the spatial partitioning of the marine environment into ecologically meaningful
units across the many scales that management is applied. Marine benthic
habitat maps also hold a broader scientific value as they provide baseline
information that can be used for further work relating ecological patterns to the
processes that drive them; a fundamental requirement in understanding,
predicting and mitigating anthropogenic impacts on the marine environment
(Pittman et al. 2011).
1.2.3 Tools for benthic habitat mapping

Recent advances in benthic habitat mapping have paralleled the evolution of acoustic surveying technologies that allow the remote sensing of geophysical properties of the seafloor, providing information on morphology and surficial geology. The three principle acoustic technologies applied for benthic habitat mapping comprise single beam acoustic ground discrimination systems (AGDS) (e.g. Anderson et al. 2002, Freitas et al. 2003), side scan sonar systems (SSS) (e.g. Cochrane & Lafferty 2002, Collier & Humber 2007) and multibeam echosounders (MBES) (e.g. Kostylev et al. 2001, Wilson et al. 2007). Due to the ability to record co-registered travel time of the acoustic signal (bathymetry) and amplitude of the acoustic return (backscatter) while achieving continuous coverage of the seafloor, MBES is gaining widespread adoption as a tool for benthic habitat mapping (Brown et al. 2012a, Kostylev 2012). A review paper by Brown et al., (2011) highlights both the growth in benthic habitat mapping studies and the increasing adoption of MBES as a mapping tool. Of the 144 applications of acoustic technologies cited by the authors between the years 1997 and 2011 more than 70% were published in the 5 years from 2007 to 2011. Applications of MBES for benthic habitat mapping correspondingly increased from 24% of studies between 1997 and 2006, to 54% between 2007 and 2011.

Acoustic methods alone are unable to directly measure biological attributes of the benthos with the exception of biogenic structure formed by colonial organisms such as corals (Collier & Humber 2007), reef forming polychaetes (Degraer et al. 2008), or aggregations of shellfish (Wildish et al. 1998, van
Some form of in situ data is therefore required to provide the biological information needed for benthic habitat mapping. A range of physical and visual sampling methods are available and their application is largely dependent on study scale, the nature of the habitats of interest and their association with hard or soft substrate types (see Van Rein et al. 2009 for review).

Underwater video supported by relative positioning systems represents a substantial advance in the collection of observational data used to inform and validate models of benthic community distribution. Video affords a permanent, retrievable and non-destructive visual record of the seafloor and associated macro-benthos without the depth and time constraints associated with diver-based underwater visual census (UVC) (Van Rein et al. 2009) or the spatial limitations of point-based methods. To date however, there has been little investigation of the potential limitations in the use of underwater video.

### 1.2.4 Integrating acoustic and biological data

A diversity of approaches has emerged to relate characteristics of the acoustic return to patterns of biological distribution. Subjective interpretation methods where boundaries between features are manually digitised over an image were ubiquitous before computer aided analytical techniques became widely accessible and are still used due to the high degree of precision achievable by an expert interpreter. The preponderance of ever larger and sometimes highly dimensional environmental datasets and the desire to reduce potential subjectivity in the mapping process has led to the widespread adoption of automated analytical methods based on mathematical relationships between
environmental and biological variables. These are broadly categorised by the order in which relationships between environmental and biological variables are derived in the map production process.

1.2.5 *Unsupervised approaches to data aggregation*

Unsupervised (top down) approaches to data segmentation are largely driven by patterns inherent in the acoustic data. Biological data, where they are used, serve to provide ecological meaning *post hoc* to an essentially geophysical classification of seafloor units. A range of clustering algorithms have been implemented to reduce acoustic data layers to homogeneous seafloor categories including modified *k*-means (McGonigle et al. 2009), *fuzzy c-means* (Lucieer & Lucieer 2009) and *artificial neural networks* (Chust et al. 2010). These methods are, in general, reliant on the assumption that areas of the seafloor displaying similar acoustic characteristics support similar species or communities. Stevens and Connolly (2004) note that this assumption may not always be supported and identify two types of error that may arise - *false homogeneity*, where sites with similar environmental conditions support different biological distributions, or *false heterogeneity*, where sites with different environmental conditions support similar biological distributions. Similarly Post (2007) and Kostylev (2012) suggest that inconsistencies in assumed bio-physical relationships have the potential to lead to misinterpretation of patterns of biological distribution. In some cases where the geographic region of interest is so large that collection of representative *in situ* data is not logistically feasible, modelled surrogates of key environmental drivers of habitat distribution (e.g. hydrodynamic regime, temperature, nutrient
availability) are often used in place of direct sampling. For example Kostylev and Hannah (2007) combine MBES survey data with broad scale environmental datasets to derive a process driven model of marine habitats on the Canadian continental shelf.

1.2.6 Supervised approaches to data aggregation

Supervised (bottom up) approaches to partitioning acoustic data rely on establishing relationships between environmental and biological variables before the classification procedure. These relationships are then used to extrapolate biological information to areas where environmental data are present but where no \textit{in situ} data exists. Two broad approaches have been undertaken to mapping benthic habitats in this respect.

Species-based approaches are used to model the response (realised niche) of a focal species to a set of environmental predictor variables across a geographic region of interest (Guisan & Zimmermann 2000). Widely applied in terrestrial ecology to test and predict species-environment relationships (Elith & Leathwick 2009), such methods are being increasingly applied to predict distributions of commercially important molluscs (Brown et al. 2012b), fish (Monk et al. 2010), crustaceans (Galparsoro et al. 2009) and rare marine species (Guinan et al. 2009).

Community-based approaches address marine habitat mapping from a broader ecological perspective by - biological distribution at a collective level. This is achieved either by combining individual species distribution models and defining community boundaries based on probability thresholds (predict first assemble later), or by aggregating \textit{in situ} biological data into community level
entities that are then modelled in relation to environmental predictors (assemble first predict later) (Ferrier & Guisan 2006). It is the latter approach (assemble first predict later) that has been most commonly applied in bottom up benthic habitat mapping studies and a diversity of statistical methods are available, many of which were originally developed for terrestrial modelling.
1.3 Thesis focus, objectives and orientation

This thesis focuses on exploring applications of some of the contemporary sampling tools and analytical techniques used to derive marine habitat maps. By testing assumptions of data quality and strengths of association between benthic communities and their measurable physical surrogates, improvements to the field of benthic habitat mapping and its application to the broader context of marine spatial planning and management can be proposed.

Research in this thesis was carried out in the nearshore marine environments of the state of Victoria on the South Eastern continental shelf of Australia. Multibeam Echosounder (MBES) and underwater video data collected to fulfil the requirements of the Victorian Marine Habitat Mapping Program form the basis of the work. The major objectives of this work are each aligned with one of four core research chapters as follows.

**Objective 1**

- To quantify and critically evaluate potential sources of uncertainty in the interpretation and integration of underwater video used to inform supervised benthic habitat classifications.

The first core chapter (*Chapter 2- Sources of uncertainty in video derived reference data*) presents an analysis of the application of underwater video data collected for training and validating spatially explicit benthic habitat models.
Specifically this chapter quantifies the 2 major sources of error pertaining to collection of this type of reference data:

Spatial error -a spatial error budget is developed for a hybrid positioning system used to co-register video data to their corresponding locations at the seafloor. The system consists of a number of different technologies and is therefore subject to measurement error which has the potential to impact on subsequent assessment of model reliability.

Thematic error –human interpretation of video information is required to render it suitable for use in habitat distribution models. Variability in interpretation is compared between trained operators assessing the same video frames between times, over 3 levels of a benthic habitat classification scheme.

Objective 2

- To test the influence of MBES backscatter and bathymetry on the accuracy of predictive benthic habitat distribution models.

Bathymetry and backscatter are the primary data products resulting from MBES survey. Much research attention has focussed on image classification using MBES backscatter (and morphological derivatives of bathymetry) due to its utility in describing geophysical characteristics of the seafloor. Bathymetry is seen as an important predictor in mapping biological components of the seafloor, especially on the shallow continental margin as it acts as a surrogate for direct drivers of community distribution such as light availability and exposure to wave energy. Despite their obvious associations with habitat distribution, backscatter and bathymetric depth are often considered separately.
Chapter 3, (Linking bathymetry and backscatter for benthic habitat characterisation) introduces a supervised machine learning approach to segmenting MBES derived predictor variables into discrete habitat classes using both bathymetry and backscatter information. A decision tree approach was used to assess the performance of a model containing both data types as inputs to that of a model using bathymetric derivatives alone.

**Objective 3**

- To assess the applicability of supervised acoustic remote sensing methods for broad-scale habitat change assessment using time series MBES and video surveys.

Detecting change in biological habitat distribution using remotely sensed data is becoming common practice in shallow marine environments, where a variety of data types from satellite and airborne sensors is available. Beyond the shallow sub-littoral, acoustic sensors are now commonly used to obtain full coverage information describing geophysical characteristics of the seafloor at a comparable resolution to that obtained in optically shallow waters. Few studies exist that describe spatially explicit change in benthic biological habitats beyond the range of optical sensors. Chapter 4 (Assessing change in benthic community distribution using acoustic methods) presents a post-classification comparison study of a site where MBES and in situ video surveys were carried out in consecutive years. Drawing on the outcomes of chapter 2 and methods developed in chapter 3, independent map products for each year are compared to assess change in terms of habitat gains and losses. A matrix based approach
is applied to differentiate signals of systematic habitat gains and losses from those occurring randomly.

**Objective 4**

- To assess the impact of hydrodynamic energy on the accuracy of benthic habitat classification.

The majority of the temperate southern coastline of Australia is exposed to unimpeded oceanic swells occurring as a result of synoptic storm events in the Southern Ocean. As a result, the benthic taxa of the shallow continental margin display a range of adaptations to living in a high energy environment. While energy induced spatial variation at a species and community level has been well documented on the rocky shore, there has been little work carried out beyond the eulittoral.

From a habitat classification standpoint, bathymetry and orientation are often cited as proxies for the effects of exposure to hydrodynamic energy. At sites that are subject to a gradient of exposure (i.e. sheltered to exposed) these proxies may not necessarily hold true, as habitats occurring at similar depths and orientation may be subject to differing hydrodynamic regimes. In the final core chapter of this thesis (*Wave driven exposure as a surrogate for benthic habitat distribution*) a site specific spectral wave model is developed for an area of seafloor along Cape Otway, the major coastal feature of the west coast of Victoria. Due to the orientation of the Cape and the prevailing direction of wind and oceanic swells, the site experiences a gradient of wave exposure from fully to moderately exposed. A wave exposure proxy is incorporated into a
benthic habitat characterisation of site to test the effects of including this ancillary information on classification accuracy. Variable importance is determined using a Random Forest approach.

The final chapter of the thesis (Chapter 6 – Summary and key findings) summarises the key findings of each research chapter and relates them to the four objectives outlined above (Section 1.3 of Chapter 1). The chapter assesses the relevance and contribution of the research outcomes to the field of benthic habitat mapping and provides future directions for research arising from the work.

The research chapters in this thesis represent standalone manuscripts that are presented in a format suitable for submission to peer-reviewed academic journals.

Chapter 2 has been accepted for publication in ‘Marine Geodesy’.

Rattray, A., Ierodiaconou, D., Monk, J. and Laurenson, L.B.J. Quantification of Spatial and Thematic Uncertainty in the Application of Underwater Video for Benthic Habitat Mapping (Lead contributor: Alex Rattray)

Chapter 3 is published in ‘Estuarine, Coastal and Shelf Science’

Chapter 4 is published in ‘Marine Ecology Progress Series’


Chapter 5 is in preparation for submission to ‘Remote Sensing of Environment’

Rattray A., Ierodiaconou, D., Womersley, T., Monk, J. & Che Hasan, R. Wave-driven exposure as a surrogate for benthic habitat distribution. (Lead contributor: Alex Rattray)
2. Sources of uncertainty in video derived reference data.
2.1 Introduction

Growing awareness of human impacts on the marine environment and the resources that it provides has instigated efforts to improve management strategies in marine systems across the world (Halpern et al. 2008, Cogan et al. 2009). These efforts are increasingly borne out by the implementation of legislation aimed at delineating areas of the seafloor into conservation and management zones, where human activities are restricted. With the growing paradigm of marine spatial planning (Douvere 2008) there has been an urgent need to fill knowledge gaps regarding the composition and extents of benthic habitats (Lourie & Vincent 2004). A suite of new technologies has emerged that facilitate the mapping of benthic habitats on the continental shelf at a resolution approaching that obtained by optical sensors commonly used to map terrestrial land cover and shallow marine habitats. Principal among these are high-resolution swath acoustic systems, which, combined with recent advances in positioning, motion sensing and computer processing power (Kenny et al. 2003), provide highly accurate geo-physical representations of large areas of the seafloor. In order to map marine habitats in an ecological sense, however, accurate in situ reference information is required to provide biological context to the structural information that these systems provide (Brown et al. 2011).

Video surveys are now a common element of marine habitat mapping programs and have become an essential tool in providing biological data to inform the acoustic classification process and also to assess the map quality.
Video has advantages over point-based sampling methods as data density is generally higher (allowing for data redundancy), large areas of seafloor are able to be sampled relatively quickly and transition zones in the benthic environment are more likely to be sampled (Parsons et al. 2004). In the marine habitat mapping context, video information is used to supplement derivation of species or habitat classes from acoustic datasets in 3 broadly defined ways:

1. To inform the benthic habitat modelling process before image classification (supervised ‘bottom up’ approach) (e.g. Guinan et al. 2009, Shumchenia & King 2010, Ierodiaconou et al. 2011)

2. To examine linkages between biotic observations and acoustic (bathymetry/backscatter) data (supervised ‘top down’ approach) (e.g. Kloser et al. 2010, Anderson et al. 2011)

3. To derive measures of confidence around model outputs (model validation) (e.g. Kendall et al. 2005, Holmes et al. 2008, Rattray et al. 2009)

A key attribute of observation data used for these applications is that they are free from error. There is a growing body of evidence in the terrestrial literature, however, to suggest that observation data are commonly subject to error and therefore have the potential to introduce error into spatial habitat models both in the training/interpretation phase and also in assessment of model validity (Foody 2002, Foody 2009, Gardin et al. 2011). The presupposition that reference data are correct is an important underlying assumption in its application to marine habitat mapping. This issue has been the subject of much ongoing research in the terrestrial remote sensing literature where Lunetta et al., (1991) describe a range of error sources in spatial data from the acquisition
to the decision making phase, including error arising from data processing, analysis, conversion, error assessment and final product presentation.

Congalton and Green (2009) identify two major sources of error in remote sensing reference data; spatial error, where attributes of a point on a map are not correctly co-located with their corresponding position on the ground, and thematic error, where labels on a map differ from attributes of their corresponding position on the ground. Due to manner in which underwater video information is collected and interpreted, it is reasonable to assume that both of these types of error may exist. It is therefore important to apply similar rigour in identifying and mitigating such error in marine applications. Failure to recognise and account for the effects of spatial and thematic error in reference data may lead to spurious results, or undermine confidence in subsequent maps of habitat distribution (Newton et al. 2009). While spatial and thematic errors are not independent from one another (Cherrill & McClean 1999, Congalton & Green 2009), as a thematic error may arise from a positioning error, for reasons of clarity these error types will be treated separately in this study.

2.1.1 Spatial Error

One of the advantages of using vessel mounted multibeam echosounders (MBES) over towed systems is the ability to directly utilise the vessel’s offset GPS positional information without applying corrections for layback of a towed body, resulting in improved positional accuracy of soundings and associated backscatter values. This, along with increasing sounding density especially in shallow waters (McGonigle et al. 2010), has led to highly
spatially accurate geophysical products that can be displayed in a GIS at horizontal resolutions in the order of <1m. To make use of these datasets at the highest resolution possible, there is a need to accurately co-register individual video frames with corresponding physical data.

Traditionally the position of a deployed camera system is often inferred either directly from vessel position (drop-camera systems) or through an estimation of layback from the vessel’s GPS antenna (towed camera systems). Recently, positioning systems that measure the relative 3-dimensional vector of an acoustic transponder attached to the camera system have become available (Philip 2003). These systems improve the accuracy with which the position of the camera system can be defined relative to the vessel without some of the uncertainty associated with inference methods, such as the action of surface and sub-surface currents and surface wind conditions. The ultra-short baseline (USBL) configuration has been used to position camera systems for marine habitat mapping purposes across a range of depths, from the continental shelf (<500m) (Rooper & Zimmerman 2007, McGonigle et al. 2009, Kloser et al. 2010, Lucieer et al. 2012), to deep sea environments (Van Rooij et al. 2010, Foubert et al. 2011, Howell et al. 2011, Van Rooij et al. 2011). There is currently limited discussion in the habitat mapping literature of the potential for error in these systems and how that might impact on their application with broad-scale swath sonar datasets.

2.1.2 Thematic error

While many (semi) automated segmentation procedures for geophysical data are available, the interpretation of video imagery remains a largely manual
process and is therefore prone to human subjectivity (Culverhouse et al. 2003, Hearn et al. 2011). Interpretation of underwater video data is time consuming, cognitively demanding and inevitably requires long periods of sustained concentration where complex and sometimes unavoidably subjective decisions are routinely required to be made. An inevitable result therefore of visual image interpretation is a level of uncertainty in the assignation of class labels to video frames. Studies comparing interpretation of terrestrial imagery have shown that the rate of between-observer variation is often high. For example, Powell et al., (2004) found that trained image analysts disagreed on thematic class allocations from airborne video imagery almost 30% of the time. In a similar study Wulder et al., (2007) reported that mean interpreter disagreement in allocating primary class labels to airborne video frames was between 30 and 40%. Converting video imagery into a format suitable for analysis calls for considerable post processing time. Quantifying repeatability of video interpretation is therefore especially relevant as programs to map and monitor the seafloor increase in size, often requiring the use of multiple interpreters to process large video datasets.

Although there are no published studies that primarily address uncertainty inherent in the collection and interpretation of underwater video, the potential for subjective bias in human image interpretation is acknowledged throughout the literature. There is an implicit assumption that video derived data used to validate habitat distribution models is free from error, despite the evidence supporting the argument that this is rarely, if ever the case. This study addresses two fundamental issues surrounding the use of video derived
biological data to inform and assess site specific models of benthic habitat distribution. Firstly, a theoretical spatial error budget is developed for a system used to establish geodetic position of a towed body used for video survey. Inputs to the spatial error budget were derived from both manufacturers’ accuracy statements for each component of the system and from vessel attitude data collected during a field survey. Second, we test the repeatability of using multiple observers to attribute thematic information to a set of video derived images of benthic habitats across 3 levels of a hierarchical benthic classification scheme.

2.2 Methods

2.2.1 Study Area

Data used in the study were taken from towed video surveys carried on the eastern side of Cape Otway on the west coast of the state of Victoria, Australia (Figure 2.1). Video transect surveys were carried out to provide in situ reference data for a 2007 MBES survey encompassing an area of 132km². Approximately 30 linear kilometres were surveyed by video from depths of 10 – 79m over 17 transects in January 2008. Video transects were planned to capture the diversity of site morphological and acoustic variation evident in the MBES bathymetry and backscatter data and were run in a shore-normal direction from shallow to deep.
Figure 2.1 Site location in Bass Strait, Australia showing artificially illuminated MBES bathymetry. Major reef systems of the site extend from Cape Patton (A) and Point Hawdon (B).

Topographic features of the site are characterised by three major reef complexes extending from Point Hawdon, Cape Patton and the adjacent headlands to the west. These reef systems are generally high relief and sediment free at their shallowest extents becoming increasingly sand inundated with depth. Benthic characterisation of the site (unpublished data) found unconsolidated sandy sediments comprise 76% of site area with the remainder attributed to high profile reef (2%) and complexes of low profile reef interspersed with sandy sediments (22%).

Reef epibiota was found to be typical of the region, characterised by dense stands of the canopy forming kelp species *Ecklonia radiata* in the bathymetric highs (<25m) with an emergent red algal understory as the canopy receded.
with depth. Invertebrate communities found on deeper reefs comprised predominately sponges interspersed with bryozoans, stalked ascidians, whip gorgonians and hydroids.

\[2.2.2 \quad \text{Towed video system}\]

The reference data were acquired using an acoustically positioned towed video sled. A VideoRay Remote Operated Vehicle (ROV) was adapted as a towed camera platform for video survey. Survey speeds were generally between 0.5 – 1.0 ms\(^{-1}\) (1-2 knots). The towed video platform was maintained at approximately 1m above the seafloor by a winch operator observing real-time video on board the survey vessel in order to ensure a consistent field of view (approximately 2m\(^2\)) and maintain continuous spatial resolution for video classification.

An Ultra Short BaseLine (USBL) transponder attached to the video unit allowed 3-dimensional positioning of the video unit relative to the vessel’s dGPS antenna which was located directly above the pole mount housing the USBL transceiver on the port side of the vessel. Angular rates of roll, pitch and azimuth (±0.1°) at the dGPS antenna were measured and corrected using a KVH motion sensor mounted directly above the USBL transceiver.

\[2.2.3 \quad \text{Spatial Error Propagation}\]

The positioning system, although commonly referred to as USBL is reliant on a number of other technologies for its application in the field. Repeatable accuracy of the acoustic USBL component of the system is expressed as a percentage of slant range (straight line distance from shipboard transceiver to
deployed transponder beacon) under static conditions. Confusion often arises when this accuracy measure is assumed to represent repeatable accuracy of the relative positioning system as a whole (Philip 2003). Changes in vessel attitude during survey must be measured and this information used to dynamically correct for divergence of the vessel’s reference frame (and therefore that of the USBL transceiver) from the gravity reference frame (Figure 2.2). Each component used to accomplish this has some potential for measurement error. The following section details each of the components in the system used to establish geodetic position of towed video sled. The potential measurement error and its effect on determination of planar position for each of these components are examined independently, and then a repeatable error term is propagated for the system as a whole over its depth of operation.
A theoretical spatial error budget was created using manufacturers’ accuracy statements for each component of the system (Table 2.1). Constants of vessel attitude, relative heading and absolute (GPS) position used in the elements of the error budget are as follows:

- A conservative approximate of horizontal DGPS accuracy was deemed to be $\pm 1.5\text{m}$.
- Vessel attitude information from the vertical reference unit (VRU) was used to determine maximum angular values above and below horizontal for
pitch and roll for a single transect done in the middle of video survey operations at the site. The quadratic mean of maximum amplitude values was taken resulting in RMS values of 19° for pitch and 24° for roll. These values were deemed representative of maximum values for each of these parameters for the transect, while allowing for potential outliers in the dataset.

- The video sled was assumed to be behind the vessel, following in line with vessel heading. A constant vector of 170° from the USBL transceiver was assumed for calculations as the transceiver was situated on a pole mount on the port side of the vessel.

- While slant range (distance from transceiver head to beacon) varied according to changeable wind and current conditions, it was assumed to be 2 times depth of operation based on line markers used during survey.

Table 2.1 Manufacturer’s accuracy statements for components of the relative positioning system (95% probability). Acronyms are: DGPS –differential GPS; USBL – ultra short baseline system; VRU – Vertical reference unit.

<table>
<thead>
<tr>
<th>Component</th>
<th>Manufacturer/model</th>
<th>Error Type</th>
<th>Error parameter</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGPS</td>
<td>OmniStar Inc.</td>
<td>Planar</td>
<td>Lat/Long</td>
<td>±1.5m</td>
</tr>
<tr>
<td>USBL</td>
<td>LinkQuest Inc./Tracklink 1500MA</td>
<td>Angular</td>
<td>Azimuth</td>
<td>±1.0°</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Planar</td>
<td>Range</td>
<td>±0.2m</td>
</tr>
<tr>
<td>VRU</td>
<td>KVH Industries, Inc./GyroTrac.</td>
<td>Angular</td>
<td>Pitch</td>
<td>±1.0°</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angular</td>
<td>Roll</td>
<td>±1.0°</td>
</tr>
<tr>
<td>Gyro compass</td>
<td>KVH Industries, Inc./GyroTrac.</td>
<td>Angular</td>
<td>Heading</td>
<td>±1.0°</td>
</tr>
</tbody>
</table>
The theoretical spatial error budget was calculated assuming a constant sound velocity profile of the water column (i.e. no salinity or temperature induced ray-bending effects), that instrument positional offsets had been applied correctly, and that instruments were correctly calibrated prior to survey. It is noted that the assumption of a constant sound velocity profile would not be valid in areas with strong freshwater influence or other circulation effects.

2.2.4 Vertical Reference Unit Error

The VRU provides vessel attitude information (pitch and roll) along the $x$ and $y$ axes respectively which, integrated with vessel heading azimuth and USBL derived acoustic angle, is used to establish the true relative relationship between the baseline in the transducer head and the acoustic beacon. Inaccuracies in the angular measurement of pitch and roll parameters lead to positional uncertainty in both the $x$ and $y$ directions. The impact of VRU roll angle error relative to the $x$ axis is given by:

$$\Delta x = \frac{\theta_{err} D}{\cos(\theta)^2}$$

Where: $\theta$ = roll angle (radians), and $\theta_{err}$ = angular error in roll (radians), $D$ = depth (m).

A similar argument can be proposed for the impacts of pitch angle error on beacon position relative to the $y$ axis:

$$\Delta y = \frac{\theta_{err} D}{\cos(\theta)^2}$$
Where $\theta = \text{pitch angle (radians)}$, $\theta_{err} = \text{angular error in pitch (radians)}$ and $D = \text{depth (m)}$.

### 2.2.5 Gyrocompass Error

Effects of gyro-compass error on the true estimate of vessel heading azimuth can impact the computation of beacon position in both the x and y directions. Uncertainty in position as it relates to the $y$ axis is given by:

$$
\Delta y = \sin(\theta)R \theta_{err}
$$

Where: $\theta = \text{heading azimuth}$, $\theta_{err} = \text{angular error in heading azimuth (radians)}$ and $R = \text{slant range to beacon (m)}$.

And in relation to the $x$ axis:

$$
\Delta x = \cos(\theta)R \theta_{err}
$$

Where: $\theta = \text{heading azimuth}$, $\theta_{err} = \text{angular error in heading azimuth (radians)}$ and $R = \text{slant range to beacon (m)}$.

### 2.2.6 USBL Acoustic Error

A baseline (typically < 0.1 m) between a pair of elements in the USBL transducer head is used to measure phase differences in the incoming acoustic signal in order to establish a baseline relative azimuth from the transducer head to the acoustic beacon. Positioning error introduced by inaccuracies in acoustic azimuth angle is calculated in a manner analogous to gyro-compass azimuth error in the $x$:
\[ \Delta x = \cos(\theta) R \theta_{err} \]  \hspace{1cm} (5)

and \( y \) axes:

\[ \Delta y = \sin(\theta) R \theta_{err} \]  \hspace{1cm} (6)

Where: \( \theta \) = acoustic azimuth, \( \theta_{err} \) = angular error in acoustic azimuth (radians)

and \( R \) = slant range to beacon (m).

Acoustic slant range, the Euclidean distance between the transducer head and the acoustic beacon, is determined by the two-way travel time of the acoustic signal between the transducer head and the acoustic beacon. As previously stated this solution does not incorporate any possible refraction effects related to differences in the velocity gradient due to variable temperature or salinity through the water column. Positional uncertainty in the \( x \) plane can be defined as:

\[ \Delta x = \cos(\theta) R_{err} \]  \hspace{1cm} (7)

And in the \( y \) plane:

\[ \Delta y = \sin(\theta) R_{err} \]  \hspace{1cm} (8)

Where: \( \theta \) = acoustic azimuth (radians), \( R_{err} \) = error in acoustic slant range

Total propagated error was calculated at 1m intervals over the bathymetric range of the study site. To preserve the probabilistic accuracies (95%) of the various components of the relative positioning system, error terms for each component were combined in quadrature using the equation:
\[ x_{rga} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \] (9)

This also has the effect of converting a generally elliptical pattern of error around the transponder beacon into a circular error term which facilitates reporting.

2.2.7 *Comparison of Observer Variability*

In this study 3 researchers trained in identification of sessile benthic taxa of the region independently analysed the same set of 198 video frame samples included in the analysis. Experience levels differed between observers with one observer experienced in interpreting underwater videography, one observer experienced in conducting diver based underwater visual census with limited video interpretation experience, and one observer experienced in both diver based and video based assessment (Table 2.2).

**Table 2.2** Levels of experience of observers participating in the study at the first classification attempt (T₁), and at the second classification attempt (T₂).

<table>
<thead>
<tr>
<th>Observer</th>
<th>Diver based census (UVC)</th>
<th>Video based census T₁</th>
<th>Video based census T₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>&gt; 100 hrs.</td>
<td>&gt; 300 hrs.</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 100 dives</td>
<td>&gt; 300 hrs.</td>
<td>&gt; 300 hrs.</td>
</tr>
<tr>
<td>3</td>
<td>&gt; 100 dives</td>
<td>-</td>
<td>&lt; 50 hrs.</td>
</tr>
</tbody>
</table>

Observers 1 and 2 were involved in the collection of the video reference data and the initial habitat classification and therefore had additional understanding of the spatial distribution and visual representation of the habitat cover classes.
Observers 1 & 2 first classified the video data set immediately after its collection in 2008 and then again in 2010. Observer 3 classified the video reference frames only once in 2010. This resulted in 2 intra-observer comparisons (observers 1&2) between T₁ (2008) and T₂ (2010) performed more than 1 year apart, and 3 inter-observer comparisons (observers 1, 2 & 3) at T₂.

2.2.8 Sampling Design

In terms of scientific objectivity and the ability to rigorously meet the criteria of the statistical analyses presented in the following section there is a general requirement that sampling be undertaken using a random sampling approach. However, limitations imposed by both the sampling methodology and by the nature of the habitats under investigation precluded the implementation of a simple random sampling design for 2 reasons common to the collection of map reference data. Firstly, the population of reference sample points were located along video transects which were not assigned random starting points, instead transects were planned using MBES information to capture the diversity of bathymetric, morphological and acoustic variation at the site using logistically realistic sampling effort. Secondly, the significant disparity in occurrence between the largest and smallest classes of interest would have invariably resulted in over representation of bare sediments, estimated to cover approximately 76% of the site, with few samples assigned to rarer biological classes under a simple random sampling strategy.

The first issue is addressed by noting that all of the biological classes of interest at the site are associated with hard substrata which provide attachment
points for macroalgal and sessile invertebrate species. Areas of reef and associated reef/sand complexes are clearly distinguishable from surrounding morphologically simple sandy sediments through visual inspection of high resolution MBES bathymetry, derived complexity and acoustic backscatter images. The deliberate focus of the video survey on these more complex areas was designed to capture a representative sample of epibenthic biological diversity at the site without over committing sampling effort to areas where prior surveys and field experience had not discerned visible epibiota. Although unconsolidated sediments were observed to vary across the site in their geological and morphological structure they were reduced to a single category for the purposes of the study. It was therefore assumed that despite the non-random nature of the video transect design, that the sub regions sampled were representative of the classes of interest within the site and could therefore be generalised to the full site extent in order to compare observer variability.

The second issue of disparate representation between small and large classes was addressed by implementing a stratified random sampling design based on locations where video transects coincided with an existing biological distribution model of the site developed by the Victorian Marine Habitat Mapping Project (unpublished data). The habitat map was used as a guide to attribute class information to video sampling points from which equal numbers of video frames were selected at random from each of the 6 habitat classes present in the habitat distribution model. This allowed, as far as possible, statistically valid samples to be drawn from rarer biological classes without the
potential imbalance introduced by the over inclusion of the unconsolidated sediment class ubiquitous to the study area.

2.2.9 Video Attribution Protocol

Still images extracted from video were supplied to each observer in digital format. Images were pooled and their order randomised before classification to reduce any potential bias due to recency effects (Culverhouse et al. 2003). No time limit or other restrictions were imposed during the evaluation process other than defining the evaluation protocol. The evaluation protocol was designed to record as many aspects of the biological composition of each frame of video as possible in a logistically practicable manner. Observers scored the video frame as populated or unpopulated and for populated frames recorded the 3 most prevalent macroalgal species and/or the 3 most prevalent taxa of sessile invertebrates providing a density estimate for each; sparse - < 25%, medium 25 – 75% and dense > 75%. A purpose-designed graphical user interface (GUI) was created using Microsoft Access which allowed observers to efficiently access the expected range of physical and biological attributes via dropdown menus and append this information along with the frame’s positional coordinates to spread sheet format (Figure 2.3). Biological attributes recorded by each observer were then aggregated to a 3 level hierarchical classification (Figure 2.4a) conforming to the Victorian Towed Video Classification Program (Ierodiaconou et al. 2007) for the decision rules presented in Figure 2.4b. This classification scheme was developed for the Victorian Marine Habitat Mapping Project.
Figure 2.3 Classification GUI showing: (a.) Video frame grab with satellite derived time-stamp; (b.) Drop down lists of physical and biological habitat descriptors; (c.) Positional information and coupled time/date strings used to georeference classification details for each video frame; (d.) Combined positional and thematic information appended to spread sheet format and (e.) aggregated thematic attributes exported to a GIS.
Figure 2.4 a) Three level hierarchical classification scheme used in the study with b). associated decision rules for aggregation of observations.
2.2.10 Statistical Analyses

To derive comparisons of observer variability we examined matched-pairs data consisting of two dependent samples where each observation in one sample matches an observation of the other (Agresti 2007). Comparisons between observers were carried out by cross tabulating each of \( n \) matched reference classifications in a contingency matrix where rows display results generated by observer X and columns display results generated by observer Y (or in the case of comparison of the same observer between times, observer X at \( T_1 \) and observer X at \( T_2 \) (Table 2.3). Row totals \((n_{i+})\) signify the frequency with which each of \( k \) habitat labels was assigned by observer X and column totals \((n_{+j})\) signify frequencies of instances assigned by observer Y.
Table 2.3 General contingency matrix for comparing observer agreement. Values on the major diagonal representing agreement between observers are highlighted in bold.

<table>
<thead>
<tr>
<th>Observer X</th>
<th>Observer Y</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Observer T₂)</td>
<td>(Observer T₁)</td>
<td>nᵢ+</td>
</tr>
<tr>
<td>i = Rows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>n₁₁</td>
<td>n₁₂</td>
</tr>
<tr>
<td>2</td>
<td>n₂₁</td>
<td>n₂₂</td>
</tr>
<tr>
<td>.....</td>
<td>.....</td>
<td>.....</td>
</tr>
<tr>
<td>k</td>
<td>nₖ₁</td>
<td>nₖ₂</td>
</tr>
<tr>
<td>Column Total</td>
<td>nᵢjc</td>
<td>nᵢjc</td>
</tr>
</tbody>
</table>

Cell values located along the major diagonal of the matrix indicate instances of labelling agreement between observers while all other values in the principal matrix are indicative of classification disagreement. Overall labelling agreement between two observers was calculated by summing the frequencies of the main diagonal of the table and dividing by the sample size n, given by the equation:
While overall agreement gives an indication of conformity between observers it is potentially misleading as only frequency agreement values are considered while marginal frequencies are ignored. The Kappa statistic \( K \) is generally thought to be a more conservative global estimate of agreement as marginal values are included in its propagation which controls the measure of agreement by adjusting for the proportion of agreements expected to occur by chance (Congalton & Mead 1983, Congalton et al. 1983, Congalton 1991). A maximum likelihood estimate of \( K \) (Cohen 1960, Congalton & Green 2009) was calculated for each observer comparison matrix by obtaining the ratio of the residual of the summed frequencies of observed agreement and chance (marginal) agreement, and the residual of perfect agreement \( (n^2) \) and chance agreement expressed as:

\[
K = \frac{\sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} n_{i+} n_{+i}}{n^2 - \sum_{i=1}^{k} n_{i+} n_{+i}}
\]

(11)

Class specific agreement measures were derived in order to determine observer conformity by class. With reference to Table 2.3, the proportion of agreement \( (P_A) \) specific to each category \( k \) was computed by dividing twice the frequency
of agreement between observers for a specified category \( i \) by the row and column totals for that category such that:

\[
P_A = \frac{2n_{ii}}{n_{i+} + n_{+j}}
\]  

(12)

Non-parametric McNemar tests (Agresti 2007) were used to test significance of marginal homogeneity separately for each habitat category under the null hypothesis that the discordant marginal frequencies of class labels assigned by each observer share the same distribution, that is, row marginal proportions are equal to corresponding column marginal proportions. For each of these tests the \( k \times k \) contingency table was collapsed to form a fourfold (2×2) table. Specifically, for each habitat category \( k \) (\( k = 1...n \)), all categories other than \( k \) are combined, producing a 2 way table for the class \( k \) versus others distinction.

An example case for testing cell \( n_{11} \) (with reference to Table 2.3) is given in Table 2.4. Due to the unbalanced and at times sparsely populated nature of the resultant contingency matrices, significance of pairwise habitat class comparisons was established using exact binomial probability calculations (Fisher 1922, Agresti 2007) with Bonferroni corrected alpha values for type I error control across multiple tests.
Table 2.4 Aggregated general contingency matrix used for McNemar tests of marginal homogeneity (refer to Table 2.3).

Observer Y

*(Observer T₂)*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n₁₁</td>
<td>n₁₂ + ... n₁ₖ</td>
</tr>
<tr>
<td>2</td>
<td>n₂₁ + ... n₂ₖ</td>
<td>n₂₂ + ... n₂ₖ</td>
</tr>
</tbody>
</table>

Observer X

*(Observer T₁)*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n₁₁</td>
<td>n₁₂ + ... n₁ₖ</td>
</tr>
<tr>
<td>2</td>
<td>n₂₁ + ... n₂ₖ</td>
<td>n₂₂ + ... n₂ₖ</td>
</tr>
</tbody>
</table>
2.3 Results

2.3.1 Spatial Error Budget

A theoretical error budget was developed for a USBL-based system used to derive the geodetic position of a towed camera sled. Total propagated error (TPE) of the relative positioning system showed a linear and highly correlated positive relationship with depth ($R^2 > 0.99$). At the shallowest operational depth of the USBL system (5m), TPE was 1.5m, increasing to 3.6m at 60m depth (corresponding to the deepest video survey at the present site), and 5.7m at 100m depth (the operational maximum within the context of the Victorian Marine Habitat Mapping Program).

The cumulative contributions of each component to TPE over depth (Figure 2.5) showed a diminishing influence of the constant error term apportioned to the dGPS from 63% at the surface to 10% at a depth of 100m. Likewise, the effect of the slant range constant decreased with depth from 8% ($x$ axis) and 3% ($y$ axis) to <1% in both cases at a depth of 100m. Gyro-compass heading error and USBL azimuthal error were the greatest contributors to TPE in the $x$ axis between 44-100m depth, although proportional error from both these sources in the $y$ axis remained low contributing a combined total of just 8% to TPE at 100m depth.
2.3.2 Thematic Error

Comparison of Overall Observer Agreement

Video interpretation data from each observer were aggregated according to the 3 levels of the benthic classification scheme (Figure 2.4) and arranged in cross validation matrices to examine pair-wise differences both between and within observers. Results show that classification agreement decreases with increasing complexity (decreasing hierarchy) of the classification scheme employed. At the highest (base habitat) level of the classification scheme (populated v unpopulated), observers were almost universally in agreement, the lowest rate of agreement (94%) occurred in the between years comparison of observer 2, while all others showed rates of agreement between 98 and 100%. Overall agreement of the 5 pair-wise comparisons ranged from 78 – 90% using the level of habitat type (4 classes), and 65 – 85% for the level of habitat sub-type.
Values of the kappa coefficient of agreement for classification comparisons showed a similar trend decreasing from a range 0.67–0.85 and mean value of 0.78 for 4 class classifications, to a range of 0.57–0.78 and mean value of 0.69 for 6 class classifications.

2.3.3 Comparison of between class observer agreement

Observer agreement between classes varied according to observer, biological habitat class of interest, and also the level of classification hierarchy employed. As the SED class corresponds directly to the unpopulated class outlined above it will not be considered here. At the 4 class habitat level (Figure 2.6), invertebrate dominated habitats (INVDOM) showed the highest average proportion of agreement of 0.87 (87%), while attribution of the mixed ALG/INV class was most commonly disagreed upon (73%). Average observer agreement at the 6 class level ranged between 58% for attribution of the mixed brown and red algal class (MBMR) and 76% for the sessile invertebrate (INV) class (Figure 2.7).
Figure 2.6 Proportion of observer agreement between habitat classes using a 4 category classification scheme. Pairwise comparisons within and between observers are indicated by legend below figure.

Figure 2.7 Proportion of observer agreement between habitat classes using a 6 category classification scheme. Pairwise comparisons within and between observers are indicated by legend below figure.
Averaged values of agreement, however, are potentially misleading if not compared within the context of the variability between observers. Agreement between observer 3 and other observers was found to be consistently lower than other comparisons across all classes at both habitat type and habitat sub-type levels of the classification hierarchy. These comparisons contributed the lowest agreement values of 58% and 65% at the 4 category level (Figure 2.6), both occurring in the ALG/INV class. At the 6 category level, proportional agreement between observers was lowest in the MRI class at 45% and 40% between observer 3, and observers 1 and 2 respectively. Further disagreement was also evident between observers 1 and 3 in the mixed brown and red algae (MBMR) category where only 45% of class attributions were the same. Comparisons between years (observers 1 and 2) showed less variation in attribution agreement; however, similar patterns across classes were discernible. At the 6 class level (Figure 2.7), lowest agreement occurred in the MBMR class, while at the 4 class level both observers misclassified the ALG/INV class most commonly.

### 2.3.4 Tests of Marginal Homogeneity

McNemar tests were used to assess homogeneity of row and column marginals for each class (Table 2.5) irrespective of classification agreement. Homogeneous marginal values indicate that misclassifications between observers for each class are similar, that is, similar errors of omission and commission are occurring for each class within each observer comparison.

At the 4 category level, marginal values were found to differ significantly between observers 1 and 3 for the MB class (p < 0.01) and the ALG/INV class.
(p < 0.01). Row and column marginals for the comparison of observers 2 and 3 were significantly different in the MB (p < 0.01) and MRI (p < 0.01) classes, while in the inter-year comparison of observer 1 a significant difference was found in the ALG/INV class (p < 0.01) between years.

At the 6 category level of the classification scheme, similar significant differences were observed between observers and years in the ALG/INV class which was equivalent at both levels of the scheme. Significant differences were also found between observers 1 and 3, and observer 1 between years for the ALDOM class (p < 0.01 in both cases). No significant differences were detected between observers 1 and 2, or between times for observer 2 at either the 4 class or the 6 class level of the classification scheme.
Table 2.5. Summary of overall agreement, Kappa coefficient (\(K\)) of agreement and McNemar tests (p-value) for proportional homogeneity.

<table>
<thead>
<tr>
<th>Habitat sub-type (6-class)</th>
<th>Overall agreement</th>
<th>(K)</th>
<th>MB</th>
<th>MBMR</th>
<th>ALG/INV</th>
<th>MRI</th>
<th>INV</th>
<th>SED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1(v2) (2010)</td>
<td>82</td>
<td>0.78</td>
<td>0.65</td>
<td>0.13</td>
<td>0.02</td>
<td>0.04</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 1(v3) (2010)</td>
<td>72</td>
<td>0.65</td>
<td>&lt;0.01*</td>
<td>0.59</td>
<td>&lt;0.01*</td>
<td>0.06</td>
<td>0.07</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 2(v3) (2010)</td>
<td>65</td>
<td>0.57</td>
<td>&lt;0.01*</td>
<td>0.67</td>
<td>0.041</td>
<td>&lt;0.01*</td>
<td>0.04</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 1 (2008 v 2010)</td>
<td>78</td>
<td>0.73</td>
<td>0.13</td>
<td>0.02</td>
<td>&lt;0.01*</td>
<td>0.69</td>
<td>0.38</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 2 (2008 v 2010)</td>
<td>78</td>
<td>0.74</td>
<td>1.0</td>
<td>0.109</td>
<td>0.63</td>
<td>0.38</td>
<td>0.06</td>
<td>0.25</td>
</tr>
</tbody>
</table>

* \(p < \) Bonferroni-adjusted significance criterion of 0.01

<table>
<thead>
<tr>
<th>Habitat type (4-class)</th>
<th>Overall agreement</th>
<th>(K)</th>
<th>ALGDOM</th>
<th>ALG/INV</th>
<th>INVDOM</th>
<th>SED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1(v2) (2010)</td>
<td>82</td>
<td>0.85</td>
<td>0.35</td>
<td>0.02</td>
<td>0.07</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 1(v3) (2010)</td>
<td>78</td>
<td>0.71</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>0.59</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 2(v3) (2010)</td>
<td>76</td>
<td>0.67</td>
<td>0.72</td>
<td>0.41</td>
<td>0.65</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 1 (2008 v 2010)</td>
<td>86</td>
<td>0.81</td>
<td>&lt;0.01*</td>
<td>&lt;0.01*</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Observer 2 (2008 v 2010)</td>
<td>90</td>
<td>0.87</td>
<td>0.03</td>
<td>0.63</td>
<td>0.69</td>
<td>0.25</td>
</tr>
</tbody>
</table>

* \(p < \) Bonferroni-adjusted significance criterion of 0.017
2.4 Discussion

This study assessed error inherent in positioning and interpreting in situ towed video information used to inform and assess acoustically derived habitat maps. While it is recognised that the effects of spatial and thematic error are not independent of one another, as a positioning error has the potential to create a thematic error, for reasons of scope and clarity they are discussed separately. Our results indicate that potential propagated error in the positioning system described is highly correlated with depth of operation, and varies from 1.5m near the surface to 5.7m in 100m of water. In order of decreasing classification hierarchy, mean overall observer agreement was found to be 98% (range 6%), 82% (range 12%) and 75% (range 17%) for the 2, 4 and 6 class levels of the scheme respectively. For the observers that repeated the classification between times, agreement was somewhat higher and less variable at 98% (range 0%), 88% (range 4%) and 78% (range 0%). Patterns in between observer variation discovered in this study demonstrate that levels of agreement between observers are related to the level of detail imposed by each hierarchy of the classification scheme, the feature of interest and also potentially to the level of observer experience.

2.4.1 Spatial Error Budget

In land based environmental mapping activities, misregistration of reference information with remotely sensed data sets has the recognised potential to significantly affect the spatial arrangement of derived habitat distribution
models (Czaplewski, 1992, Canters, 1997), and limit the confidence with which estimates of thematic error can be applied and interpreted (Foody 2002a, Foody 2002b). In marine systems, where reference information is collected beyond the direct capabilities of terrestrial GPS systems, geodetic positioning of reference samples is considerably more involved.

The error budget presented here reveals the potential limitations in the spatial accuracy of geodetically located video frames used to inform and validate marine habitat mapping studies. It also demonstrates that the spatial error term is not static, but instead propagates through the various components of the relative positioning system, and is in turn amplified with increasing depth of operation. For a positioning system comprising a number of different technologies, it is important to understand how the contribution of each part affects the whole. By partitioning the theoretical spatial error into its components, each related to an associated interval of measurement error, a greater understanding of the sources and magnitude of such error can be gained. This logic allows appropriate steps to be taken in order to mitigate the effects of spatial error throughout the mapping process.

Critically, *a priori* knowledge of accuracy limitations of a system used to establish the position of reference data points can be used at the planning stage of a project in order to define a conservative or appropriate goal for mapping resolution based on system configuration and depth of operation. In this manner subsequent measures to establish metrics of map reliability may be viewed as already incorporating the potential for misregistration of map products and reference data by reducing the model’s sensitivity to spatial error.
Previous studies have demonstrated that pre-classification aggregation of data (data degradation) has the potential to reduce the effects of map to reference data misregistration (Dai & Khorram 1998, Van Rompaey et al. 1999), thereby increasing map accuracy albeit at the cost of reduction in spatial resolution and thematic information. Contrary to this view, Carmel (2004) suggests that by aggregating data post-classification that thematic accuracy can be better balanced with the requirements of optimum spatial resolution, resulting in more efficient data usage. A clear requirement of both of these approaches, however, is an understanding of the magnitude of spatial uncertainty in the reference data.

From an operational perspective, partitioning the spatial error provides the means to establish survey guidelines that account for limiting factors such as the effect of sea conditions on system accuracy. We found that combined pitch and roll error (19° and 24° RMS amplitude respectively) became more important with depth and contributed to 30% of total propagated error at a depth of 100m. Incorporating this type of analysis into survey design permits either restriction of survey operations to an optimal range of conditions for a given system, consequent rejection of data collected beyond those optima, or by better informed error reporting to the end user of subsequent mapping products. Further, by better understanding how each component contributes to total propagated error, informed decisions can be made regarding the adequacy of each element of the relative positioning system for a given spatial accuracy target. In some instances the benefits of increased spatial accuracy may be outweighed by the costs and operational requirements of implementing a
dedicated positioning system. For example, where expected spatial error is potentially less important than the resolution mismatch between MBES and \emph{in situ} data (Cutter et al. 2003, Dolan et al. 2008).

Effects of spatial error are likely to be higher in areas of locally patchy or heterogeneous class distribution than in areas of homogenous habitat. Osborne and Leitao (2009) report that spatial errors in bird occurrence data are relatively unimportant in modelling core habitat but can be misleading in more marginal or fragmented areas. By definition, areas of the seafloor which display fine-scale heterogeneity in habitat type will contain a greater number of inter-class boundaries and are therefore likely to be more prone to the effects of spatial misregistration of datasets. Some studies have sought to account for this issue by omitting reference samples occurring at class boundaries (Foody 2002), although there is evidence to suggest that excluding such boundary information limits the usefulness of measures of map accuracy to homogeneous areas of habitat (Muller et al. 1998). Powell et al., (2004) noted that while exclusion of edge and mixed pixels from a validation dataset resulted in higher overall map accuracy, validation samples available in rarer classes were reduced significantly, resulting in lower and less rigorous class specific accuracies, or in classes that were eliminated entirely from the validation process. Although no analysis of patch size or spatial arrangement of errors was incorporated into this study it is undoubtedly an important concern when considering the effects of spatial accuracy and is an important direction for future research.
2.4.2 Thematic Error

Our results suggest that a more generalised classification scheme reduces both inter and intra-observer variation in classification results. This is ostensibly a factor of reducing the number of density related choices made by interpreters that directly affect the classification outcome. While this finding is somewhat self-evident, in that a reduction in the number of choices faced by an interpreter results in less misclassification, it is important to quantify at what level of a given classification scheme the effects of declining classification repeatability lead to an unacceptable level of uncertainty in deriving measures of map accuracy.

For a single date census of biotopes, where delineating areas in which rarer habitat categories occur is given added importance, a higher level of classification uncertainty may be acceptable. Where multi-temporal change studies are concerned however, classification accuracy is often seen as more important than classification resolution as errors may be propagated through time-series making meaningful interpretation of results difficult (Coppin et al. 2004, Foody 2010). In a simulation study, Foody (2009) found that class allocation inaccuracies as small as 5% in a ground data set lead to systematic underestimation of perceived change detection accuracy and overestimation of change extent. He further notes that the magnitude of bias in these estimates was positively related to the amount of inaccuracy present in ground reference datasets. Acceptable levels of error therefore, are determined by the specific goals of a mapping study. Where the goals of a mapping study are to provide a census of biotopes which can also be used for change comparison with future
models, as is often the case (Harris & Baker 2011), a classification scheme that allows end mapping products to be easily generalised to a simpler hierarchical level may be more appropriate than one that does not.

Of the 6 classes in the most detailed hierarchy of the classification scheme, interpretation of the mixed brown and red algae class (MBMR) resulted in the lowest rates of observer agreement. The MBMR class was most often confused with either the mixed brown algae class (MB) or the mixed algae and invertebrates class (ALG/INV). Confusion with MB occurred in cases where qualitative density estimates of the red algal understory were either overestimated as medium (25-75%) or underestimated as sparse (<25%). Non-significant comparisons of marginal values (Table 2.5) indicate that disagreement over class attribution was similar, that is, interpreters experienced similar difficulties in separating MB from MBMR. Systematic errors such as these are characteristic of the classification scheme imposing lines between habitats that change along a continuum and are therefore not easily resolved. Resolution of ecotones between similar habitats, especially at superficial boundaries is a problem common to all interpreters in this study and suggests that a hard classification scheme may not adequately encompass the finer nuances of shared boundaries between habitats. Similar difficulties in differentiating thematic classes that share gradational boundaries have been reported in classification of forest types (Powell et al. 2004, Wulder et al. 2007), plant communities (Cherrill & McClean 1995, Cherrill & McClean 1999, Hearn et al. 2011) and urban landscapes (Zhou et al. 2010). The imposition of discrete boundaries partitioning benthic communities into
mutually exclusive categories is in itself a fundamental complexity faced by 
seafloor habitat mapping studies. Benthic habitat types that share a 
discontinuous boundary such as that between a reef edge and adjacent soft 
sediments are generally less common than those that display changes in benthic 
community structure over a continuum (Lucieer & Lucieer 2009, Brown et al. 
2011). While habitat categories that share a gradational boundary may be easily 
separable at their extremes, demarcation between them becomes increasingly 
difficult and potentially more arbitrary closer to the shared boundary (Gopal & 
Woodcock 1994). Study of between-observer agreements in quantitative 
density estimates rather than in habitat classes may elucidate the cause of the 
difference (such as a persistent observer bias).

Confusion between MBMR and ALG/INV occurred at a more fundamental 
level where class distinctions were based on the presence or absence of sessile 
invertebrates within the red understory of the kelp canopy. Unlike the sessile 
invertebrates occurring on deeper, kelp free reefs which tended not only to be 
more numerous but also to exhibit more instances of erect and branching 
morphologies, the sponges occurring in the kelp understory were commonly 
found to be display massive or encrusting growth habits. This makes 
identification of invertebrates on algal dominated reefs potentially more 
challenging than where they occur without macroalgae. Disagreement in this 
respect may be attributed to operator expertise as the two observers with the 
most experience agreed on this habitat more consistently both between 
operators and between times, although further work is needed to state this with 
any certainty.
A number of strategies have been proposed to mitigate observer bias in terrestrial image-based class allocation. Gopal and Woodcock (1994) suggest that classification (scheme) ambiguity and interpretive uncertainty can be incorporated in a modified overall error term by having interpreters assign an index of confidence to each class allocation. A similar ‘soft’ classification approach is set out by Zhang and Foody (1998) whereby classes are defined according to both a primary and secondary label. This approach is adopted by Wulder et al., (2007) who also incorporate spatial uncertainty into their error assessment analysis by defining map/reference data agreement as the modal class occurring in a 3 by 3 pixel neighbourhood surrounding a map pixel of interest. Reducing bias by having multiple (≥ 3) interpreters review the same images and allocating class information by consensus has also been proposed (Powell et al. 2004, Wulder et al. 2007). Although recent efforts to automate image analysis due to institutional backlogs of video imagery (Schoening et al. 2012) may indicate a wider issue stemming from the disparity between the capacity to collect imagery and the time and costs involved in interpretation. Quantitative evaluation of sources of classification uncertainty, as described in this study, provides a basis for directed interpreter calibration thereby minimizing costs acquired through duplication of effort.

### 2.5 Conclusions

This study demonstrated the potential for error in the spatial and thematic attributes of video derived reference data and to our knowledge is the first to comprehensively address these issues in relation to marine habitat mapping. While uncertainty was evaluated using data from a single survey our results
provide a cautionary tale for the collection and interpretation of underwater video data in general. We found that spatial uncertainty surrounding the geodetic positioning of a towed video system was linearly correlated with depth of operation, and that the contribution of each element of the positioning system to the total error term was also depth dependent. Combined error from the VRU and gyro-compass was found to be greater than USBL acoustic error across all depths. This highlights the importance of incorporating measurement errors from all components of a positioning system in developing an overall repeatable error term. The error propagation methodology presented here is of value to workers in the field especially where the resolution of swath acoustic imagery approaches that of in situ video observations, or where habitat facies of interest are known to be fragmented or patchy. The effects of spatial error may be mitigated through careful planning, field design and reporting.

Increasing complexity of a benthic classification scheme was found to increase variability in class allocation both between and within observers. Disagreement between observers was more likely to occur between classes that shared similar attributes, or between classes where the class defining attribute was potentially cryptic. Furthermore, rates of intra-observer variation, while marginally better overall, were found to be similar when defining potentially ambiguous classes to those found between observers. Quantifying uncertainty in reference data used to inform and validate biological habitat maps is important considering the inherent assumption that reference data are error free. This work identifies spatial and thematic error associated with video reference data but does not extend to its effects on the accuracy or interpretation of subsequent
applications. It does however provide a benchmark to better understand errors associated with both tools and methodological processes used to inform and validate benthic habitat distribution models.
3. Linking bathymetry and backscatter for benthic habitat characterisation.
3.1 Introduction

Habitat classification and mapping is a prerequisite for assessment of the status of marine ecosystems allowing improved management and protection. Technological developments in the past two decades have provided new tools for scientists and managers to better understand the marine environment (Hughes Clarke et al. 1996, Kenny et al. 2003). There has also been an increase in the spatial coverage and resolution of data used to monitor and assess the marine environment, with a broader seascape approach to environmental management complementing the species-by-species, small spatial scale management approach (Anderson et al. 2002).

Central to this approach is the discrimination of areas of uniform environmental conditions providing specific physical environmental niches for biotic assemblages. It is recognised that taxa are not randomly distributed between varying habitats; rather, they show associations with the physical properties of the surrounding environment such as geology, light attenuation in the water column and wave/current exposure (in the broader context of geography, chemistry and biological interactions) (Guisan & Zimmermann 2000). By delineating boundaries between specific benthic assemblages at sampled locations and relating these differences directly to patterns in full coverage acoustic datasets representing depth, bottom morphology and acoustic scattering, inferences can be made about distributions beyond sampled locations alone (Ierodiaconou et al. 2007, Post 2007, Holmes et al. 2008).
The application of multibeam echosounder (MBES) technology allows for full coverage of the seafloor with co-registered bathymetry (depth) and backscatter (intensity of acoustic returns) information at a fine resolution (1’s-10’s m² pixel size) over broad geographic regions (10’s–100’s km²). Coupled with concurrent advances in computer technology, vehicle motion sensors, underwater videography and navigation, unprecedented datasets are available with which to map seabed habitats at multiple spatial scales (Wilson et al. 2007). With increasing volume and complexity of available datasets, automated image classification techniques (Pal & Mather 2003) are becoming more important in providing repeatable and efficient methods for improving benthic characterisation (Diaz et al. 2004). Much of the marine benthic mapping literature has focused on the classification of non-biogenic structure such as substrate, grain size and geomorphic features (Lathrop et al. 2006, Fonseca & Mayer 2007). More recently MBES data has been shown to provide useful information about the nature and distribution of biological communities found at the seafloor (Ierodiaconou et al. 2007, Fonseca et al. 2008, Holmes et al. 2008).

The nature of the benthic structure has a profound effect on the communities that can establish (Kostylev et al. 2001). It is widely recognised that species are not randomly distributed between varying habitats. Rather, species show associations with the topographic variability of the benthic terrain (Beaman et al. 2005). For example bathymetry (depth) is an important modifier of distributions of biological communities through (1) light attenuation, which influences growth in photosynthetic plants and (2) exposure to surface wave
action and surge in high-energy environments with strong associations found between depth and biotic community structure (Iampietro et al. 2005).

Environmental complexity can be investigated through analysing seafloor information from hydroacoustic sources through the application of landscape metrics typically used in terrestrial studies (see Wilson et al. 2007). For example Gratwicke and Speight (2005) found rugosity to be an overriding factor in determining fish species richness, while Iampietro et al. (2005) used Topographic Position Index to estimate the distribution of 8 species of rock fish. Toohey (2007) found topographic variation to be important in determining the diversity of macroalgal dominated assemblages in temperate waters.

Acoustic backscatter from multibeam datasets has also been shown to be a valuable source of data in defining geomorphological features (Gardner et al. 2003), yet their application for defining benthic biological communities has been limited.

The combination of bathymetry and backscatter have been found to be useful in the characterisation of seafloor substrates (Dartnell & Gardner 2004, Rooper & Zimmermann 2007). However, testing the contribution of each of these data types in discriminating biological communities has been limited (Ierodiaconou et al. 2007). Further assessment of the relative importance of these data sources in discriminating benthic communities is required.

In this study we present the application of an automated image processing technique using decision tree (DT) classifiers to provide detailed baseline characterisation of the broad biological communities of the Cape Nelson site in southwest Victoria, Australia. MBES data (bathymetry, backscatter and
derivative products) are integrated with biological observation data collected using a towed video system for model training and assessment. This study provides critical baseline information regarding the distribution of sessile biological communities on the shallow coastal shelf, within the Otway bioregion in temperate south-eastern Australia. It also provides insight into the relative importance of bathymetry and backscatter data sources derived from multibeam echosounder data in benthic habitat characterisation.

3.2 Methods

3.2.1 Study Area

The study area covers an area of 32 km² (Figure 3.1), between Lawrence Rocks and Cape Nelson. Benthic biological community distribution information in this region prior to this study was non-existent. The location was identified as a priority as part of the Victorian Marine Habitat Mapping Project due to its interesting geological history, location within an upwelling zone, and proximity to the expanding bulk carrier port of Portland. The Otway bioregion is an area of high species diversity with a considerable (>90%) indigenous component (Phillips 2001). The location is characterised by high, deep water wave energy attenuated by a steep offshore gradient. Waters are cold temperate, with mean temperatures ranging from 11°C in winter to 18°C in summer, and is particularly productive due to the large, predictable, upwelling events from the nearby Bonney upwelling zone. The area between Cape Sir William Grant and Cape Nelson (Nelson Bay) is formed by a large circular volcanic depression, the northern rim of which largely constitutes the modern coastline.
Figure 3.1 Hill shaded MBES bathymetric coverage of Cape Nelson (Victoria, Australia) study area with towed video transects overlaid.

3.2.2 Modelling overview

The habitat mapping process presented here comprises 4 major elements (Figure 3.2). MBES acoustic data were collected, cleaned and gridded to 2.5 metre resolution, comparable to the positional accuracy accomplished for video observation data, before derivation of secondary geophysical predictor datasets. Major biological habitat types for the site were compiled from visual examination of georeferenced video transect data and overlaid with geophysical predictor datasets for determination of conditional relationships between the 2 data types using DT classifiers. Finally, relationship rules established between video training data and predictor variables were used to assign class labels to pixels where no training data were available. Model
validation was accomplished by comparing observed and predicted habitat classes at sample locations that were not used in the model training process.

![Conceptual diagram of data analysis procedures used for the development of classified habitat maps.](image)

**Figure 3.2** Conceptual diagram of data analysis procedures used for the development of classified habitat maps.

### 3.2.3 Sonar Data Acquisition

Multibeam echosounder data were collected on the 5\textsuperscript{th} and 6\textsuperscript{th} November 2005 using a hull-mounted Reson SEABAT 8101 (240 kHz) MBES yielding a swath width of approximately 4 times water depth. Prior to survey, calibration offsets for pitch, roll, yaw and latency were applied after conducting a detailed patch test. Daily sound velocity profiles were collected at the deepest (~ 60m) area of the site during survey operations to correct for local variations in sound velocity through the water column during processing.

Positioning was accomplished using a real-time Differential GPS (DGPS), integrated with a Positioning and Orientation System for Marine Vessels (POS MV) for heave, pitch, roll and yaw corrections (± 0.1° accuracy). Navigation, data logging, real-time quality control and display were carried out using
Starfix suite 7.1 (Fugro proprietary software). The sounding data were edited on board ship to hydrographic standards by trained surveyors using Starfix Suite 7.1 for data cleaning and application of corrections due to tides, sound velocity, vessel draught, settlement, squat and relative position of the transducer head. The XYZ data were then used to produce a bathymetric grid at 1m horizontal resolution and a range resolution of ±12.5mm which was subsequently regridded to 2.5m horizontal resolution (Figure 3.3). Backscatter values were corrected for gain and time varied gain using the University of New Brunswick (UNB1) algorithm (Starfix suite 7.1). Backscatter processing also included the correction for transmission loss, the actual area of ensonification on the bathymetric surface, source level, and transmit and receive beam patterns (see Fonseca et al. 2008). Additionally backscatter was corrected for seafloor bathymetric slope from the multibeam bathymetry dataset. This resulted in normalised corrected grid (2.5 m resolution) representing relative backscatter intensity (dB) across the study site (Figure 3.3). Striping was evident in the backscatter dataset particularly in transitional zones between reef and sediment. Whilst artefacts in the backscatter intensity were observed they were included in further analysis to assess their capacity to differentiate biological groups defined.
To further characterise local variation within the MBES imagery and delineate analogous regions of morphology and signal scattering, a suite of products were derived from the bathymetry and backscatter datasets (Table 3.1). These derivatives were selected for their expected influence over distribution of biological assemblages in terms of exposure to wave energy and benthic currents (aspect, BPI), susceptibility to sediment accumulation (slope, BPI), complexity and surface area of reef structure (complexity, rugosity, maximum
curvature) and variations in high and low frequency signal scattering properties of the substratum (HSI) (Figure 3.4).

The HSI transformation was developed by Daily (1983) to filter noise from SeaSat RADAR returns and has been alternatively referred to as IHS (Nasr Hamed Nasr et al. 2002, Hong et al. 2009) and HIS (Pu et al. 2011, Jiang et al. 2012). Common usages of the transformation are merging multi-sensor data (Fonseca & Manjunath 1996), spectral image enhancement (Richards 2005) and feature segmentation (Nasr Hamed Nasr et al. 2002). In this study a synthetic colour image was derived by applying high and low pass filters to the backscatter image in order to separate high and low frequency information then mapped to hue (chromatic) and intensity (achromatic) respectively with a fixed saturation value. The HSI components were then transformed into the red, green, blue (RGB) colour space to derive 3 variables which represent low to high intensity backscatter returns. These variables were expected to provide ancillary information to the classifier by respectively reducing noise-related artefacts present in the backscatter intensity layer. While commonly applied to airborne spectral datasets, the HSI transformation process has more recently been applied to acoustic implementations of benthic habitat mapping (Ierodiaconou et al. 2007, Rattray et al. 2009, Ierodiaconou et al. 2011, Rattray et al. 2013) and fish habitat suitability modelling (Monk et al. 2010, Monk et al. 2011, Monk et al. 2012).
**Table 3.1** Derivative products from MBES bathymetry and backscatter.

<table>
<thead>
<tr>
<th>Derivative</th>
<th>Source</th>
<th>Description</th>
<th>Analysis window</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rugosity</td>
<td>Bathymetry</td>
<td>The ratio of planar area to actual surface area across a neighbourhood surrounding a central pixel. (ArcGIS)</td>
<td>3x3</td>
<td>Jenness 2004</td>
</tr>
<tr>
<td>Slope</td>
<td>Bathymetry</td>
<td>Slope denotes the maximum change in elevation between each cell and cells in its analysis neighbourhood. Calculated in degrees from horizontal. (ENVI)</td>
<td>3x3</td>
<td>Wilson et al. 2007</td>
</tr>
<tr>
<td>Aspect:</td>
<td>Bathymetry</td>
<td>Aspect (azimuthal bearing of steepest slope) was separated into 2 components representing ‘northness’ sin(aspect) and ‘eastness’ cos(aspect) of steepest slope to account for inherent circularity in the azimuthal aspect variable. (ENVI)</td>
<td>3x3</td>
<td>Wilson et al. 2007</td>
</tr>
<tr>
<td>Complexity</td>
<td>Bathymetry</td>
<td>Complexity refers to the second derivative or rate of change in the slope (slope of slope) and is a measure of local variability in benthic terrain. (ENVI)</td>
<td>3x3</td>
<td>See above for slope calculation.</td>
</tr>
<tr>
<td>Benthic Position Index</td>
<td>Bathymetry</td>
<td>The BPI value was derived as a categorical measure of elevation differences between a focal group of cells and the mean elevation of the surrounding cells within a user-defined annulus. (BTM/ArcGIS)</td>
<td></td>
<td>Lundblad et al. 2006</td>
</tr>
<tr>
<td>Maximum Curvature</td>
<td>Bathymetry</td>
<td>Maximum curvature describes the steepest curve of either plan or profile convexity through a defined cell neighbourhood. (ENVI)</td>
<td>3x3</td>
<td>Schmidt et al. 2003</td>
</tr>
<tr>
<td>HSI</td>
<td>Backscatter</td>
<td>A synthetic colour image was derived by applying high and low pass filters to the backscatter image in order to separate frequency information. High and low frequency information is then mapped to hue (chromatic) and intensity (achromatic) respectively with a fixed saturation value. (ENVI)</td>
<td>3x3</td>
<td>Daily 1983</td>
</tr>
</tbody>
</table>
Figure 3.4 Multibeam derivatives: (a) Complexity; (b) Maximum curvature; (c) HSI (R); (d) Northness; (e) Slope; (f) HSI (G); (g) Benthic Position Index; (h) Rugosity; (i) HSI (B).
3.2.4 Video Data Acquisition

Following visual analysis of the preliminary multibeam datasets, 7 video survey transects were run perpendicular to the coast and selected to capture the gradient of biological communities with increasing depth. A VideoRay Remote Operated Vehicle (ROV) fitted with a hydro-wing was adapted as a towed camera platform for video survey. Survey speeds were generally between 0.5 – 1.0 ms\(^{-1}\) (1-2 knots). The towed video platform was ‘flown’ at approximately 1m above the seafloor by a winch operator observing real-time video on board the survey vessel in order to ensure a consistent field of view and maintain continuous spatial resolution for video classification. The camera was acoustically positioned using a Tracklink Ultra Short Baseline (USBL) tracking system. Vessel motion (pitch, roll and yaw) was corrected using a KVH motion sensor. A real-time DGPS unit was used to fix vessel location and apply corrections for the acoustically fixed camera position (±3 metres). Navigation and data logging were carried out using Starfix suite 7.1. In total, 17.5 linear kilometres of video survey was collected over 4 days in March 2006 incorporating depths from 10 to 57m.

Each video frame was linked to a georeferenced position via an overlaid GPS time stamp allowing classified frames to be accurately integrated with MBES data sets for model development. The video data were classified using the Victorian Towed Video Classification Program (available on request). The program was designed in Microsoft Access to allow the direct import of camera positioning information, classification of video information, and direct
export to GIS/remote sensing packages for further spatial analysis. A video
data library with samples of defined categories was created to ensure
consistency in the video classification process. Video data was classified and
 appended to spatial coordinates (1 second interval), incorporating a layback
distance of one metre to account for spatial offset between field of view and
camera position. Dominant biological groups and percentage cover were
recorded frame by frame using decision rules presented in Chapter 2 (2.2.9).
Since the towed video observations provided a continuous measure of benthic
habitats, coding were only changed if the above class types encompassed at
least 3s of video (approximately 3m). Observation data were transformed to a
classified raster data layer based on a nearest neighbour function. Classified
pixels were randomly assigned to a training dataset (70%) used for model
propagation, and a validation dataset (30%) which was precluded from
thematic classification for model validation.

3.2.5 DT model training and evaluation

The Quick, Unbiased, Efficient Statistical Tree (QUEST) DT classifier (Loh &
Shih 1997b) using the RULGEN extension v.1.02 in ENVI v. 4.2 was used to
determine conditional relationships between video observation data and MBES
derived predictor variables. These relationships were then used to assign class
labels to pixels where no video observation data were available. The DT
classifier recursively splits predictor variables by user defined class based on a
chain of binary decision rules (nodes) until a terminal decision (leaf) which
represents the class label is reached. The advantages of using DT classifiers
over traditional statistical methods such as maximum likelihood classifiers are
that there are no underlying assumptions around the frequency distributions of classes or independence of the predictor variables, input data are able to be measured at different scales, and categorical predictor variables can be incorporated with continuous variables in the modelling process (Pal & Mather 2003). The QUEST classifier was employed for the present study in preference to other DT algorithms such as CART (Breiman et al. 1984) as it uses a non-exhaustive search routine thereby reducing the likelihood of bias towards variables that produce more splits and avoiding overfitting of the classification tree to noise in the training data. It is acknowledged that the effects of overfitting may be reduced by post pruning outputs from exhaustive search algorithms but this has been shown to substantially increase computation and processing times (Pal and Mather 2003). The QUEST classifier has been shown to be useful for habitat classification in terrestrial landscapes (Pal & Mather 2003, Sesnie et al. 2008), yet are relatively unexplored for marine mapping applications.

To compare relative contribution of bathymetry and backscatter derived variables to model accuracy, 3 QUEST decision trees were generated using identical training and error assessment data; backscatter and its HSI derivatives, bathymetry and derivatives (rugosity, complexity, maximum curvature, slope, aspect, BPI) and finally backscatter, bathymetry and all derivatives combined. Model accuracies were assessed using cross validation error matrices, summarising overall, User’s and Producer’s accuracies. User’s and Producer’s accuracies (Congalton & Green 2009) are reported here as a measure of accuracy for individual classes within the overall classification
accuracy. Producer’s accuracy represents the probability of a reference pixel being correctly classified (omission error) while user’s accuracy is representative of errors of commission, or the probability that a classified map pixel actually represents that category on the ground. Kappa analysis (Congalton & Green 2009) was performed to provide an estimate of model error by incorporating omission and commission errors and an estimate of the agreement of the classification attempt relative to the reference data. The Kappa statistic makes use of both the overall accuracy of the model and the accuracies within each class, both in terms of the predictive model and the validation sample points, to correct for chance agreement between categories. Models were compared using averaged User’s and Producer’s accuracies for each habitat class to determine optimum model performance for the 3 sets of MBES derived geo-physical layers used in model development.

3.3 Results

3.3.1 Biological assemblages

The ground truth information collected using a towed video system produced a total of 13880 classified training and validation pixels, comprising 5 primary biotic classes at the Cape Nelson site (Table 3.2). The general trend shows a distribution of benthos along a depth gradient from shallow algal dominated reefs to deeper, patchy and often sand swept reef systems dominated by a range of filter feeders with a predominance of sponges which are referred to collectively as invertebrates. Benthic assemblages were found to be almost wholly restricted to areas of reef. Areas of unconsolidated sediment (SED) were universally devoid of visible epifauna, with the exception being some
areas of fringing reef in deeper sections (<40m), potentially where low lying reef is periodically scoured by sand. The MB class consists predominantly of the kelp species *Ecklonia radiata* with occasional small patches of *Phyllospora comosa*, *Cystophora sp.* and a discrete patch of *Durvillea potatorum* on a shallow, high profile peak. *E. radiata* forms dense canopies with a sparse understorey of coralline and encrusting red algae in the shallow reef systems of the survey area from 13 to 40m. The largest algal class, MBMR is composed of a thinning canopy of *E. radiata* with an emergent understorey of medium to dense assemblages of thallose, and fleshy red algae. Sessile invertebrates were identified in approximately 50% of classified video pixels. A diverse range of sponges including encrusting, massive, papillate, erect and branching morphologies dominate these communities. Ascidians, gorgonians, bryozoans, octocorals and soft corals were also identified in the video analysis, but were often sparse or difficult to identify consistently due to insufficient video resolution and water clarity in some areas.
Table 3.2 Biotic classes extracted from towed video data collected at the Cape Nelson site, enumerating corresponding training/validation pixels and mean depth distributions (m) (± s.d.). Habitat classes are mixed brown algae and invertebrates (ALG/INV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED).

<table>
<thead>
<tr>
<th>Habitat class</th>
<th>% Cover</th>
<th>Ground truth pixels (2.5m)</th>
<th>Mean depth (m) (± s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Red algae</td>
<td>Brown algae</td>
<td>Invertebrates</td>
</tr>
<tr>
<td>MB</td>
<td>≤ 25%</td>
<td>≥ 75%</td>
<td>Absent</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>Any*</td>
<td>≥ 25%</td>
<td>≥ 25%</td>
</tr>
<tr>
<td>MBMR</td>
<td>≥50%</td>
<td>≤ 50%</td>
<td>Absent</td>
</tr>
<tr>
<td>MRI</td>
<td>Present</td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>SED</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
</tbody>
</table>
Invertebrate assemblages were classified into two groups defined by accompanying algal type. A narrow depth band of invertebrates was found to co-exist with mixed brown algae, mainly *E. radiata* at the lower limits of its depth range. Invertebrates in this class (ALG/INV) were often observed to be globe shaped, massive, or encrusting sponges, while erect, or branching forms were rare. Sponges were predominantly found in fissures, cracks and under ledges in high relief reef systems, where thallus attachment by larger kelp species was rare due to near vertical or inverted surfaces.

Sponge dominated invertebrate communities interspersed with encrusting and coralline red algae (MRI) characterise reef and mixed reef/sand substrates below 40m. In the deeper sections of the study site, more prone to sand inundation, differences were observed in the proportion of sponge morphological types. Sponges inhabiting low relief horizontal surfaces were observed to be erect and branching forms, extending higher into the water column. Sponge density and size was observed to be greatest in higher relief areas or where reef crests and peaks protruded through sandy sediment. Low relief areas, subject to heavy sand inundation support sparse communities of small sponges.

### 3.3.2 Decision tree models

Decision trees created using training data from video sources were applied to full coverage stacks of multibeam derived geophysical predictors to create 3 habitat distribution models of the Cape Nelson site. Decision trees for bathymetry and backscatter combined, bathymetry and derivatives only and
backscatter and derivatives resulted in 259, 309 and 85 nodes respectively. Test set accuracies of averaged User’s and Producer’s accuracies for each of 5 habitat classes were then compared for models created from decision trees using bathymetry derived, backscatter derived, and combined predictor variables (Figure 3.5).

**Figure 3.5** Comparison of averaged User’s and Producer’s accuracy values obtained from decision trees incorporating; backscatter and derivatives, bathymetry and derivatives and combined backscatter/bathymetry and derivatives. Habitat classes are mixed brown algae and invertebrates (ALG/INV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED).

The QUEST run for backscatter and derivatives resulted in an acceptable overall accuracy of 70 % but a low *kappa* statistic of 0.52 reflecting the inability of the model to distinguish MB and ALG/INV from other classes. The
overall accuracy of the model may be inflated due to good prediction (90 %
accuracy) of the large (1415 validation pixels) unconsolidated sandy sediment
class SED. The model created using bathymetry and derivatives performed
better, with an overall accuracy of 73 % ($kappa = 0.63$), corresponding with
comparatively good separability between all classes. The QUEST run
combining bathymetry and backscatter predictor variables was found to greatly
improve overall accuracy at 87 %, and the $kappa$ statistic at 0.83 (Figure 3.6).
Higher rates of accuracy from the decision tree run using combined variables
suggest that both bathymetry and backscatter variables are important predictors
in resolving habitat classes using a supervised decision tree approach.

Class confusion for the model incorporating all available geophysical data (i.e.
backscatter, bathymetry and all derivatives) are presented in Table 3.3. Results
in the main diagonal represent correctly classified data, while values in the off-
diagonal represent errors of omission and commission. Most relative
confusion was observed between the MB and MBMR algal class with 43% of
MB validation pixels incorrectly assigned to the MBMR class by the decision
tree classifier. Misclassification error (29 %) is also evident between the mixed
invertebrate classes ALG/INV and MRI. Both major sources of classification
error correspond with depth sequential classes and also with misclassification
of smaller class test sets to larger class test sets, indicating that confusion may
lie with class to class disparity in training dataset size, or confusion in
boundaries where class intergradations occur along a cline of habitat change.
Table 3.3 Confusion matrix for biotic classification of Cape Nelson site, including User’s and Producer’s (underlined italics) accuracies (%). Habitat classes are mixed brown algae and invertebrates (ALGINV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED).

<table>
<thead>
<tr>
<th>Video Class</th>
<th>MB</th>
<th>ALGINV</th>
<th>MBMR</th>
<th>MRI</th>
<th>SED</th>
<th>Total</th>
<th>User (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MB</td>
<td>40.9</td>
<td>6.0</td>
<td>3.4</td>
<td>0.0</td>
<td>0.1</td>
<td>5.0</td>
<td>77.6</td>
</tr>
<tr>
<td>ALGINV</td>
<td>4.5</td>
<td><strong>56.0</strong></td>
<td>2.4</td>
<td>0.8</td>
<td>0.4</td>
<td>3.9</td>
<td>64.8</td>
</tr>
<tr>
<td>MBMR</td>
<td>43.4</td>
<td>7.6</td>
<td><strong>93.2</strong></td>
<td>0.3</td>
<td>1.6</td>
<td>27.3</td>
<td>81.3</td>
</tr>
<tr>
<td>MRI</td>
<td>9.5</td>
<td>29.4</td>
<td>0.0</td>
<td><strong>97.5</strong></td>
<td>5.8</td>
<td>31.8</td>
<td>86.9</td>
</tr>
<tr>
<td>SED</td>
<td>1.8</td>
<td>1.1</td>
<td>0.9</td>
<td>1.4</td>
<td><strong>92.1</strong></td>
<td>32.1</td>
<td>97.4</td>
</tr>
</tbody>
</table>
Figure 3.6 Biotic classification map created using combined bathymetry, backscatter and derived products. Habitat classes are mixed brown algae and invertebrates (ALG/INV), mixed brown algae (MB), mixed red algae and invertebrates (MRI), mixed brown and mixed red algae (MBMR) and unconsolidated sediments (SED).
3.4 Discussion

The nature and extent of benthic biological communities at the Cape Nelson site were mapped using full coverage acoustic datasets and georeferenced video validation information with a decision tree classification approach. The decision tree classification using both bathymetry and backscatter derived variables was most effective in discriminating between dominant biological communities. The automated image classification techniques adopted in this study represent a substantial advantage over traditional methods such as manually digitizing community types, and combined with error assessment techniques provides important information for the iterative progression of future marine mapping studies.

The classified habitat maps of the Cape Nelson site is a generalised model, providing a snapshot of broad biological distribution. The usefulness of these models is determined by the degree of accuracy with which they can be used. Estimating model error not only provides users of the model with a quantified degree of confidence, but is also invaluable as a tool for improving future models. The error matrix and associated kappa analysis approach to assessment of classification accuracy has been commonly accepted, and is in widespread use in terrestrial land-cover and land-use mapping applications (Foody 2002, Jensen 2005). The advantage of using this technique is that it makes some compensation for chance agreement and a variance term is able to be calculated. Benthic biological communities in coastal seas are not static in
space or time (Foster-Smith et al. 2004). The capture of time-series hydroacoustics and video observations in future mapping studies may provide important insights into the nature of changes in benthic biological communities. The assessment of map accuracy is critical to determine confidence in change detection results.

The QUEST run combining bathymetry and backscatter predictor variables had an overall thematic accuracy of 87%, with a kappa coefficient of agreement value of 0.83, indicating that the derived mapping products are suitable for management applications (Congalton & Green 2009). Almost half the error (45%) can be attributed to the misclassification of the mixed brown algal class to the mixed brown and red algal class and, vice versa. This is not surprising, as both algal classes share similar species composition, depth ranges, and are associated with reef habitats which generally higher slope and rugosity values compared to the surrounding seascape. In biological terms it could be argued that these classes are of similar composition, separated only by increasing ascendancy of under-story red algal species over canopy forming brown algal species with increasing depth. This change in community composition does not exhibit a definite boundary, but changes gradually over a depth gradient.

Jensen (2005) suggests that a hybrid classification system, comprising elements of neural networks and decision tree classifiers could prove useful in optimising classification results. Hard classification techniques, used in the present study, are practical in that they precisely delineate the boundaries of marine benthic habitats. It is clear, however, that such distinct boundaries between habitats are rare with difficulty allocating transitional zones to habitat
classes in the video observation can be contributing to confusion in classification results (e.g. MB and MBMR). Habitats generally intergrade with one another along clines of some physical variable such as depth (Aburto-Oropeza & Balart 2001), or exposure (Goldberg & Kendrick 2004). ‘Fuzzy’ classification logic allows for this within-class heterogeneity by allowing each pixel to have partial membership in a number of classes, thus introducing shades of grey to an otherwise black and white classification (Jensen 2005). It is unknown whether a ‘fuzzy’ classification algorithm or neural network/hybrid approach would improve the results of the present study, but the extensive ground information collected during this study will certainly prove invaluable in future works directed at improving classification techniques to distinguish between marine habitats.

Underwater video provides a non-destructive approach which can be utilised in areas beyond the range of traditional diver methods. Improved technology and miniaturisation of equipment now allows for cost effective collection of precisely located data from relatively small vessels to “ground-truth” acoustic datasets (Kenny et al. 2003). The towed video transect approach employed here maximised the spatial area covered during field operations, increasing the likelihood of obtaining adequate representation of habitat classes within the site. Although the taxonomic resolution afforded by the video data in this study is limited, it was found to be acceptable for identifying major components of the broad habitat classes defined for the study. In towed video studies this may also be affected by camera resolution, lighting and sea conditions. Video also provides a permanent record for quality control and further analysis. For
example, video observations of demersal fish species may provide insights into species-habitat relationships when integrated with sonar imagery (Monk et al. 2011).

While towed video methods have the advantage of capturing continuous information on seafloor habitats allowing for the identification of transitional zones, they are limited by issues associated with spatial dependence of within transect data points. Non-independence of training and validation data has the potential to over-inflate accuracy measures leading to false confidence in model capability (Congalton & Green 2009). It is assumed here that by using identical training and validation points that the effects of spatial autocorrelation will be similar, allowing for comparison between the 3 models developed. Revision of survey and sampling design and/or statistical methods (Foster et al. 2009) are key requirements for future research in order to assess and mitigate the effects of spatial autocorrelation within video transect data.

Multibeam backscatter and bathymetry were found to provide complementary information for the discrimination of biological communities in this study. Geomorphic proxies in mapping the distribution of benthic biological communities are reliant on identification of predictor variables which correlate with differences in biological heterogeneity. Through the application of landscape metrics typically used in terrestrial studies (Wilson et al. 2007), we can objectively assess environmental complexity in seascapes. Further research is required to determine whether greater predictive power can be achieved using other potential derivative products from MBES data as well as
the relative importance of predictor variables in class differentiation and application at lower hierarchical levels.

The present study provides an automated classification method for mapping sessile benthic communities in a way that is meaningful, both to detailed scientific application and to the much broader requirements for management of the coastal marine environment. This study shows that bathymetry and backscatter variables from remotely sensed sources provide complimentary information for the differentiation of biological habitats. The ability to determine biotic characteristics through acoustic mapping is significant in that it provides a new dimension of information for managing our coastal seas. Map products and metadata produced in this study are a critical baseline assessment of the current state and extent of benthic habitats at the Cape Nelson site. This baseline is the starting point for the evaluation and modelling of future change in marine landscapes in this area, and provides a quantitative means of comparison to other marine communities mapped in this fashion. This study provides a framework to better understand our marine coastal environment by presenting a more coherent picture of benthic coastal ecosystems and therefore will assist to manage human impacts therein.
4.1 Introduction

Marine systems are at risk worldwide from a litany of human induced pressures. Threats range from local to global scales, are often synergistic in nature and effects may lead to fundamental shifts in ecosystem dynamics expressed as reduction of biodiversity, loss of production and disruption of ecological services (Worm et al. 2006). Information is required from across a range of disciplines in order to understand how marine systems are changing in the face of these stressors so that mitigating management practices can be applied. An understanding of the temporal dynamics of spatial distribution and composition of benthic habitats is a key component to their effective management through prioritising, monitoring and evaluating conservation targets.

The coastal seas of southern Australia are recognised as a temperate marine biodiversity hotspot (Wernberg et al. 2011). Characteristically high energy rocky reefs along Australia’s east-west running southern coastline support 30-40% of the world’s species of macroalgae of which some 50% are considered endemic. Invertebrate communities are similarly highly speciose and display significant (>90%) rates of endemism. Ocean temperatures in south eastern Australia have warmed at an average rate of 0.023 °C year⁻¹ since the early 1950’s, a rate that is 4 times the global ocean warming average for that period (Ridgway 2007). Benthic species in this region are deemed to be particularly vulnerable to increase in water temperatures associated with global and meso-scale influences of climate change and changes in ocean circulation patterns.
because the lack of suitable hard substratum to the south of the Australian continent precludes any poleward migration. Because of the length of the temperate southern coastline (>3000km) and the relatively narrow latitudinal band that it occupies (Wernberg et al. 2011), even small changes in distribution of key habitat forming species such as kelps have the potential to translate into significant habitat shifts across thousands of square kilometres.

Kelps are a foundation species that control communities and ecological processes by modifying the physical environment and resources of other organisms (Wernberg et al. 2010). Deleterious effects of human induced stressors on kelp habitats, particularly those associated with increasing ocean temperatures are therefore likely to have widespread cascading effects across other temperate marine habitats and ecosystems (Johnson et al. 2011). Surface canopies of the giant kelp *Macrocystis pyrifera* in the south eastern Australian region have been in decline since as far back as 1944 (Edyvane 2003) with some areas experiencing as much as 90% reduction in visible surface canopy. This trend has been related to incursions of warm, nutrient poor water during strong El niño southern oscillation events (Johnson et al. 2011) and southward range expansion of a barren forming urchin species (Ling 2008, Ling et al. 2009). Estimates of changes in distribution of *M. pyrifera* have been derived largely through the use of historical aerial imagery, facilitated by the fact that the species often forms dense canopies at the surface. Most temperate reef habitats however, are beyond the range of optical sensors due to the attenuation of light in the water column with increasing depth, compounded by factors such as turbidity and wave action at the surface (Kutser et al. 2006).

Concerted efforts are underway to improve capacities for spatial demarcation and management of marine benthic habitats beyond the range of optical sensors. The relatively recent proliferation of full coverage high resolution acoustic datasets and subsequent development of methods to relate acoustic returns with distinct geological and biological facies of the seafloor (see Brown et al. 2011 for review) have engendered interest in establishing acoustic remote sensing based approaches for continuous monitoring of benthic species (Paul et al. 2011, Wall et al. 2011) and communities (Nelson et al. 2011, Van Rein et al. 2011). In a recent synthesis of 57 case studies, 24 authors reported that habitat mapping activities were intended to be a part of an ongoing monitoring program while of the remainder, 24 reported that habitat maps would form the baseline for monitoring future changes (Harris & Baker 2011). Despite the clear intent of practitioners in the field to assess temporal variability in marine
habitats using acoustic remote sensing methods there is a notable dearth of published studies exploring methods to achieve this end.

A number of studies describe temporal morphological dynamics of the seafloor using serial acoustic bathymetry, (Duffy & Hughes-Clarke 2005, Smith et al. 2005, Smith et al. 2007, Chiocci et al. 2008, Yoshikawa & Nemoto 2010) but until recently assessments of biological change beyond the range of optical sensors have been based primarily on ground sampling methods that are logistically limited to safe, repeatable diving depths (ca. <20m) and are generally confined to relatively small areas of the seafloor due to safety and cost constraints.

This paper investigates change detection approaches to quantify temporal change in biological benthic habitats from a spatially explicit seascape-scale perspective using acoustic means. Our objectives are to:

1. Quantify change in terms of gains and losses in the extents of benthic habitats at a site on the temperate south-east Australian continental shelf between the years 2006 and 2007;

2. Distinguish between systematic and random patterns of habitat change;

3. Assess the applicability of supervised acoustic remote sensing methods for broad-scale habitat change assessment.
4.2 Methods

4.2.1 Study site

The study area (Fig 4.1) comprised of an 18km$^2$ overlap between two MBES surveys conducted in 2006 and 2007 as part of the Victorian Marine Habitat Mapping Program (VMHMP). The 2006 survey targeted five Marine National Parks and 13 locations of perceived significance within Victorian state waters resulting in 655km$^2$ of MBES bathymetry and backscatter data. The 2007 survey resulted in a further 607 km$^2$ of MBES data and was aimed at linking 3 discrete survey sites in the central western sector of the state to provide continuous coverage between these areas within the state’s three nautical mile marine jurisdiction. A total of 1262km$^2$ of MBES soundings and associated intensity data were collected over both surveys representing approximately 13% of Victorian state waters.

The study area was primarily dominated by morphologically flat sandy sediments although some areas of sediment starvation were apparent where underlying calcarenite was exposed between dune structures. Three distinct areas of contiguous reef characterise the site. Two highly complex basalt reef systems extend seaward (SE) from Point Hawdon (38° 40’ 19” S, 143° 52’ 10” E) and Cape Patton (38° 41’ 25” S, 143° 50’ 27” E) to depths of 40m and a well-bedded comparatively low profile shore parallel reef between 35 and 50m depth at the site’s south-western margin. The site is exposed to prevailing south-westerly wind and sea conditions and subject to a maximum tidal range of ca. 2m. Coastal sediment transport occurs via prevailing west to east long-shore drift conditions. Sediment movement was visually assessed by
comparing time-series bathymetric profiles and was evident between the two multibeam surveys, particularly in a channel between the 2 reef systems in the south-west of the site.

Figure 4.1 Site location in Bass strait, Australia showing bathymetric zones overlaying artificially illuminated multibeam bathymetry. Video transect locations for each year are displayed as black (2006) and white (2007) lines.
Detail shows bathymetric profile of advancing barchans dune waves between the years 2006 and 2007.

4.2.2 Data acquisition and processing

MBES data were acquired as part of the ongoing Victorian Marine Habitat Mapping Program. Surveys were carried out with a hull-mounted Reson Seabat 8101 echosounder (December 2006 and December 2007) at an operational frequency of 240 kHz and 150° angular sector coverage using the same vessel and equipment. Ping rate was range dependent between 40 pings sec\(^{-1}\) at the shallow margins of the site, to 15 pings sec\(^{-1}\) at depths approaching 50m. Positioning was accomplished using Starfix HP differential Global Positioning System (dGPS) (±0.30m), integrated with a POS MV (Positioning and Orientating System for Marine Vessels) for pitch, roll heave and yaw corrections (±0.02° accuracy). Data logging, real-time quality control, display, navigation and post-processing were carried out using the Starfix Suite 7.1 (Fugro proprietary software). Prior to commencing the sonar survey a tide gauge was deployed, with data corrected to Lowest Astronomical Tide (LAT). Daily sound velocity profiles were collected at the deepest (ca. 50m) point of the site during survey operations to correct for local variations in sound velocity through the water column during processing. Track spacing was designed to provide 50% overlap between adjacent lines over all depths (i.e. increasing distance between survey tracks with depth) allowing for rejection of data at acute grazing angles.
4.2.3 MBES data processing

The sounding data were post-processed on board the vessel by trained surveyors to hydrographic standard weighted by IHO order 1 uncertainty. Post-processing involved data cleaning, corrections for tides, sound velocity, vessel draught, settlement, squat and relative position of the transducer head. The XYZ data were then used to produce a bathymetric grid at 3m horizontal resolution and a range resolution of ±12.5mm. Backscatter values were corrected for gain and time-varied gain using the University of New Brunswick (UNB1) algorithm (Starfix suite 7.1). Backscatter processing incorporated the correction for transmission loss, the actual area of ensonification on the bathymetric surface, source level, and transmit and receive beam patterns (see Fonseca et al. 2009). Additionally, backscatter was corrected for seafloor bathymetric slope from the MBES bathymetric dataset to normalise the signal to the value at nadir. This resulted in a normalised corrected grid (1m resolution) representing relative backscatter intensity (dB) across the study site. In order to standardise 32-bit backscatter intensity images between surveys, they were additionally rescaled to 8-bit images. Full resolution 8-bit backscatter intensity mosaics (1m) were resampled to the optimum grid-cell resolution of the bathymetry (3m) to facilitate combined processing of both data types. This resolution accounts for the majority of spatial error in the co-location of video transect points (±3m propagated horizontal error at 50m depth).

To further characterise local variation within the MBES imagery and to aid in delineating analogous regions of morphology and signal scattering, a suite of
secondary products were derived from the bathymetry and backscatter datasets (Table 3.1 and Figure 3.4) These derivatives were selected for their expected influence in the distribution of biological assemblages as found in previous investigations. They represent variation in seafloor characteristics in terms of exposure to wave energy and benthic currents (aspect, Benthic Position Index (BPI)), susceptibility to sediment accumulation (slope, BPI), complexity and surface area of reef structure (complexity, rugosity, maximum curvature) and variations in high and low frequency signal scattering properties of the substratum (Hue-Saturation-Intensity).

**4.2.4 Towed video reference data**

Within the context of classification based analyses, *in situ* reference data are used for calibration (i.e. training) and validation (i.e. testing). The reference data used here were acquired using an acoustically positioned towed video sled collected as soon as was operationally possible after each MBES survey (in both cases within one month). After consulting all available datasets, five shore-perpendicular transects were selected to capture the range of morphological, bathymetric and acoustic scattering variation within the site. Surveys in both years were carried out along the same transects defined using an Omnistar dGPS receiver, although due to differences in localised wind and sea conditions this did not always translate to exact co-positioning of the camera system at the seabed between surveys (Figure 4.1). Detailed descriptions of methods used to acquire and classify video imagery are available in Rattray *et al.* (2009) and Ierodiaconou *et al.* (2011) and will therefore only be briefly summarised here for reasons of clarity.
The towed video platform was maintained at approximately 1m from the seabed by a shipboard operator viewing a real-time video feed via an umbilical control and data cable. An Ultra Short BaseLine (USBL) transponder attached to the video unit allowed 3-dimensional positioning of the video unit relative to the vessel’s dGPS antenna which was located directly above the pole mount housing the USBL transceiver. Angular rates of roll, pitch and azimuth (±0.1°) at the dGPS antenna were measured and corrected using a KVH motion sensor.

4.2.5 Classification scheme

Video data were reviewed and frames deemed unsuitable for classification purposes due to sub-optimal visibility caused by turbidity or proximity to the seafloor were rejected. Video frames were classified by the same operator in both years according to a simplified four class schema (adapted from Rattray et al. 2009) describing the major biological and physical components of the site (Table 4.1)
### Table 4.1 Summary of the four category classification scheme employed with associated geo-physical (MBES) descriptive statistics for each category.

<table>
<thead>
<tr>
<th>Habitat</th>
<th>Qualitative description of class end-members</th>
<th>Descriptive statistic Bathymetry (m)</th>
<th>Backscatter (relative 8-bit)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALGDOM</strong></td>
<td>Dominant canopy forming macroalgal species - small patches of the kelp <em>Phyllospora comosa</em> and occasional <em>Sargassum</em> and <em>Cystophora spp.</em> in the bathymetric highs giving way to dense canopies of the common kelp <em>Ecklonia radiata</em> on deeper reefs with sparse to medium understorey of mixed red algae.</td>
<td>Maximum 2006: -31.6</td>
<td>2007: -28.7 2006: 45.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range 2006: 4.5</td>
<td>2007: 3.2 2006: 70.6</td>
</tr>
</tbody>
</table>
The classification scheme comprises a sediment class (SED) where no epibiota were visible in the video footage, and three biological reef classes; a kelp dominated class (ALGDOM), an invertebrate dominated class (INVDOM) and a mixed class at the interface of algal and invertebrate dominated reef habitats (ALG/INV). The mixed class was included in the classification scheme as it is deemed to represent an important transition in both community composition and physical habitat structure between algal dominated shallow reefs and deeper invertebrate dominated reef communities at the site (Rattray et al. 2009). Classified point video data were resampled using a nearest neighbor function and gridded to 3m cell size to facilitate integration with co-located physical datasets.

4.2.6 Supervised image classification

The relationships between MBES derived data products and classified video frames were examined using a supervised QUEST (Loh & Shih 1997a) decision tree (DT) approach. This approach has been found useful in applications of remotely sensed data which do not generally follow a Gaussian frequency distribution (Sesnie et al. 2008, Rattray et al. 2009, Ierodiaconou et al. 2011). The univariate DT classification approach derives hierarchical non-linear relationships within the data by recursively partitioning dataset feature spaces into increasingly homogeneous categories based on a splitting criterion. The particular split threshold of a given variable that produces the largest deviance measure is chosen to recursively partition the dependent data and has the potential to be re-examined and used again as input to the tree structure. Ten subsets of the training data were taken and used to prune
resultant trees based on cross validation to an optimum size. This resulted in parsimonious tree models that have less chance of being over fitted to noise in the training data (Loh & Shih 1997a).

4.2.7 Model validation

The two resulting habitat distribution maps were assessed for accuracy using contingency matrices (Congalton & Green 2009). The Kappa (κ) statistic, used here, considers all cells in a contingency matrix, providing a correction for the proportion of chance agreement between the training sites and test data sets. The κ-statistic is a standard statistic to evaluate overall classification accuracy, providing a more conservative estimation than simple percent agreement value (Congalton 1991). Other accuracy statistics such as Producer's Accuracy or omission error (indicating how well training set pixels were classified), User's Accuracy or commission error (indicating the probability that a classified pixel actually represents that category in reality), and Overall Accuracy (the total number of correctly classified pixels divided by the total number of reference pixels) were also evaluated.

4.2.8 Change assessment

A 5×5 majority filter was applied to classified habitat maps from each time in order to reduce noise and eliminate single pixels in any class while preserving class boundaries. A transition matrix was created for the time series comparison detailing ‘from to’ habitat transitions of each class (Pontius et al. 2004, Braimoh 2006). Potential patterns in habitat change were analysed in terms of swap and net change and gross gains and losses following the methodology of Pontius et al. (2004). Net change is defined as the difference in
area of a habitat category between two times. Swap refers to the change in location of a habitat, whilst the relative quantity remains the same over time. Gross gain refers to the increase in area of a habitat category, whilst gross loss refers to a decrease in area of a habitat category between two time periods. This method allows the identification of patterns of habitat change separately from a given level of persistence in a landscape (Pontius et al. 2004, Ierodiaconou et al. 2005, Alo & Pontius 2008).

4.2.9 Habitat Persistence

Persistence characteristics of each habitat category in relation to gain, loss and net change were assessed using persistence indices from Braimoh (2006). The gain to persistence ratio was calculated as $g_p = \frac{\text{gross gain}}{\text{ persistence}}$, the loss to persistence ratio was calculated as $l_p = \frac{\text{gross loss}}{\text{ persistence}}$ and the net change to persistence ratio was calculated as $n_p = g_p - l_p$. Values of $g_p$ exceeding 1 indicate a higher tendency of habitat categories to gain from other habitat categories than to persist, while habitat categories displaying values of $l_p$ exceeding 1 indicate a higher tendency to transition to another category than to persist.

4.2.10 Random and systematic transitions

Gains and losses of habitat types over time can be attributed to random or systematic patterns of habitat change (Pontius et al. 2004). This approach distinguishes important patterns of landscape change distinct from the high level of persistence common to most habitat change investigations. Pontius et al. (2004) detail methods to compute expected habitat gains under a random process of gain by distributing observed gains among categories according to
their relative proportions at \( t_1 \). Similarly, expected losses can be calculated by distributing observed losses among categories relative to their proportions at \( t_2 \); see Pontius (2004) and Braimoh (2006) for formulae. The relative differences between observed and expected gains and observed and expected losses are then compared to derive a measure of the nature of transitions between habitat categories.

Large positive and negative deviations of this measure from zero indicate that systematic inter-class transitions, rather than random transitions, occurred between two habitat types (Braimoh 2006). Positive values indicate the inclination of one class to gain or lose from another, while negative values indicate a disinclination of one class to gain or lose from another. Expected gains and losses are calculated independently of one another with a total of 48 possible gain/loss signals between classes using a four class schema. It is possible for systematic gain or loss of one habitat category from another category to occur independent of any reciprocal systematic gain or loss. Therefore, reciprocal systematic gain/loss relationships between two classes are needed in order provide conclusive evidence of a dominant signal of habitat transition (Alo & Pontius 2008).
4.3 Results

4.3.1 Temporal habitat classifications

Independently derived spatial models representing distribution of the four representative habitat classes at the site were developed using an automated decision tree approach. Distribution of benthic habitats displayed similar patterns for each temporal classification (Figure 4.2) with the *E. radiata* dominated kelp class (ALGDOM) occupying areas of contiguous reef in the bathymetric highs changing to a narrow transitional band of mixed kelp and sponges (ALG/INV) with increasing depth. Sponge dominated invertebrate communities (INVDOM) occupied deeper hard substrata and also heterogeneous reef/sand complexes in areas of low morphological complexity associated with transitions from reef to unconsolidated sediments.
Figure 4.2. Habitat maps of the study area from a) 2006 and b) 2007 derived using an automated decision tree approach. (See figure 4.1 for analysis extents)
Confusion matrices for each temporal classification are presented in Table 4.2. Overall classification accuracy was estimated to be 92.5% and 91.9% respectively for the 2007 and 2008 classifications, while Kappa coefficient of agreement was estimated at 0.83 for both classifications (95% confidence intervals for $K$ for the 2007 and 2008 classifications are 0.80 – 0.86 and 0.79 – 0.85 respectively). Between class confusion was found to exhibit strong similarities for both classification attempts with producer’s accuracy of all classes other than ALG/INV displaying greater than 85% agreement with the reference data.

### 4.3.2 Summary of habitat transitions

A pattern of overall persistence is evident with greater than 90% of the study area remaining static between the 2006 and 2007 classifications, largely driven by the unconsolidated sediment (SED) class (Table 4.3). Around 78% of the study area that was classified as unconsolidated sediments in the 2006 classification remained unchanged in 2007. Of the biological classes, INVDOM experienced the highest persistence at 8% and ALG/INV the lowest at 0.39% of the study area.
Table 4.2. Confusion matrices for the classified images, 2006 and 2007. Each column corresponds to the ground reference pixels used for accuracy assessment for a single class. The values in the column indicate the number of those ground observation pixels classified into each class, while the values on the main diagonal (italicised) indicate agreement between ground points and classified maps.

<table>
<thead>
<tr>
<th>Reference</th>
<th>SED</th>
<th>ALGDOM</th>
<th>ALG/INV</th>
<th>INVDOM</th>
<th>Total</th>
<th>%Producers</th>
<th>%Users</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2006</strong> (overall accuracy = 93%; ( K = 0.83 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>1020</td>
<td>1</td>
<td>12</td>
<td>37</td>
<td>1070</td>
<td>96</td>
<td>95</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>11</td>
<td>111</td>
<td>2</td>
<td>-</td>
<td>124</td>
<td>97</td>
<td>89</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>6</td>
<td>2</td>
<td>24</td>
<td>3</td>
<td>35</td>
<td>56</td>
<td>69</td>
</tr>
<tr>
<td>INVDOM</td>
<td>31</td>
<td>1</td>
<td>5</td>
<td>219</td>
<td>256</td>
<td>85</td>
<td>86</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1068</td>
<td>115</td>
<td>43</td>
<td>259</td>
<td>1485</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2007</strong> (overall accuracy = 92%; ( K = 0.83 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>1070</td>
<td>8</td>
<td>1</td>
<td>36</td>
<td>1115</td>
<td>94</td>
<td>96</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>11</td>
<td>116</td>
<td>5</td>
<td>3</td>
<td>135</td>
<td>91</td>
<td>86</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>4</td>
<td>4</td>
<td>21</td>
<td>-</td>
<td>29</td>
<td>62</td>
<td>72</td>
</tr>
<tr>
<td>INVDOM</td>
<td>52</td>
<td>-</td>
<td>7</td>
<td>277</td>
<td>336</td>
<td>88</td>
<td>82</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1137</td>
<td>128</td>
<td>34</td>
<td>316</td>
<td>1615</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3 Basic change transition matrix (% of study area) for the years 2006 and 2007. The major diagonal of matrix (italicised) represents persistence (no change) of classes between years while other values represent ‘from to’ changes between habitat categories.

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>SED</td>
<td>78.39</td>
<td>81.13</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>0.39</td>
<td>5.08</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>0.39</td>
<td>1.40</td>
</tr>
<tr>
<td>INVDOM</td>
<td>4.36</td>
<td>12.40</td>
</tr>
<tr>
<td>Total 2008</td>
<td>83.92</td>
<td>4.24</td>
</tr>
</tbody>
</table>

Landscape changes were summarised in terms of swap and net change, and gain and loss (Table 4.4) The two largest classes, SED and INVDOM, experienced the largest gross gain, 5.53% and 2.85% respectively. INVDOM experienced the largest gross loss of 4.58% while SED showed a gross loss of 2.74% of the study area, these two classes therefore underwent the greatest total (sum) gross change between the two date classifications (Figure 4.3). All classes experienced a net (quantity) loss in percentage cover between the two classifications apart from SED which experienced a net gain of 2.79%.

The difference between total change and net change is the amount of swapping or location change. Swap location change accounts for the greatest change component of all classes. Of the total change over all classes 71% is classed as swap type change. That is, losses in a given class are replaced by gains from another class. Proportionally, the ALG/INV class experienced the greatest swap type change with 87% of total gross change attributed to location change and only 13% attributed to change in quantity (net change).
Figure 4.3 Spatial representation of gross gains, gross losses and persistence between the years 2006 and 2007 for classes (a) ALGDOM, (b) ALG/INV, (c) INVDOM and (d) SED. Gains are represented in yellow, losses in red and persistence in blue.
4.3.3 Persistence of habitats

Gain \((g_p)\) and loss \((l_p)\) to persistence ratios (Table 4.4) are used to assess the tendency of habitat classes to transition, ratios exceeding 1 indicate that a habitat category is more likely to gain or lose to other categories than to persist between classifications (Braimoh 2006). The only class to exceed this threshold in terms of either gains or losses was the ALG/INV class with a \(g_p\) ratio of 2.0 and \(l_p\) ratio of 2.57. These results suggest that this class has both a higher tendency to lose to other classes than to persist and also a higher tendency to gain from other classes than to persist. This is notable in light of the high proportion of swapping change and the low proportion of net change exhibited by this class.
Table 4.4 Summary of classification changes between the years 2006 and 2007 expressed as percentage of the study area. Note: $g_p =$ gain/persistence, $l_p =$ loss/persistence, $n_p =$ net change/persistence.

<table>
<thead>
<tr>
<th>Location</th>
<th>Total 2006</th>
<th>Total 2007</th>
<th>Gross Gain</th>
<th>Gross Loss</th>
<th>Sum</th>
<th>Swap Location</th>
<th>Net (quantity)</th>
<th>$g_p$</th>
<th>$l_p$</th>
<th>$n_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SED</td>
<td>81.13</td>
<td>83.92</td>
<td>5.53</td>
<td>2.74</td>
<td>8.26</td>
<td>5.47</td>
<td>2.79</td>
<td>0.07</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>5.08</td>
<td>4.24</td>
<td>0.50</td>
<td>1.34</td>
<td>1.84</td>
<td>1.01</td>
<td>0.84</td>
<td>0.13</td>
<td>0.36</td>
<td>-0.22</td>
</tr>
<tr>
<td>ALGINV</td>
<td>1.40</td>
<td>1.17</td>
<td>0.78</td>
<td>1.01</td>
<td>1.79</td>
<td>1.56</td>
<td>0.22</td>
<td>2.0</td>
<td>2.57</td>
<td>-0.57</td>
</tr>
<tr>
<td>INVDOM</td>
<td>12.40</td>
<td>10.66</td>
<td>2.85</td>
<td>4.58</td>
<td>7.43</td>
<td>5.70</td>
<td>1.73</td>
<td>0.36</td>
<td>0.59</td>
<td>-0.22</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>9.66</td>
<td>9.66</td>
<td>19.32</td>
<td>13.74</td>
<td>5.58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.3.4 Systematic and random transitions

The expected gains under a random process of gain are presented in Table 4.5a. In order to examine between class transitions represented by the off diagonal values in the matrix, values in the major diagonal are identical to proportional persistence values presented in Table 4.3. Holding persistence constant allows signals of change to be differentiated from the dominant signal of persistence across the study area. Expected gains under a random process of gain were generated by distributing the gain of each class across other classes relative to their proportions in 2006 assuming that the gain of each class and the proportion of each class in 2007 are fixed (see Pontius et al. 2004). The relative difference between observed and expected gains (Table 4.5b) expressed as (observed value – expected value)/expected value indicates the difference between the observed value and the expected value relative to the magnitude of the expected value. Values closer to zero indicate a random or expected process of gain, while numbers further away from zero signify a more systematic process of gain.

Expected losses under a random process of loss and relative differences between observed and expected losses (Table 4.6a and 4.6b) are analogous in principle to the results presented in Table 4.5 (a and b) however the roles of rows and columns are reversed. That is, losses are distributed according to the relative proportions of other classes in 2007 assuming that the losses for each class and the proportion of each class in 2006 are fixed. Relative differences between observed and expected losses may be interpreted in a similar manner.
to their converse gain relationships with values close to zero indicating losses
are occurring randomly while values further away from zero represent
increasingly systematic type change.

**Table 4.5 Inter-category gains 2006 to 2007.**

<table>
<thead>
<tr>
<th>2006</th>
<th>2007</th>
<th>SED</th>
<th>ALGDOM</th>
<th>ALGINV</th>
<th>INVDOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Expected gains under a random process of gain (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>78.39</td>
<td>0.43</td>
<td>0.64</td>
<td>2.64</td>
<td></td>
</tr>
<tr>
<td>ALGDOM</td>
<td>1.49</td>
<td>3.74</td>
<td>0.04</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>ALGINV</td>
<td>0.41</td>
<td>0.01</td>
<td>0.39</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>INVDOM</td>
<td>3.63</td>
<td>0.07</td>
<td>0.10</td>
<td>7.82</td>
<td></td>
</tr>
<tr>
<td>(b) Relative differences between observed inter-category transitions and expected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>0.00</td>
<td>-0.09</td>
<td>-0.74</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>ALGDOM</td>
<td>-0.47</td>
<td>0.00</td>
<td>8.70</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>ALGINV</td>
<td>-0.04</td>
<td>14.11</td>
<td>0.00</td>
<td>10.08</td>
<td></td>
</tr>
<tr>
<td>INVDOM</td>
<td>0.20</td>
<td>-1.00</td>
<td>1.27</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.6 Inter-category losses 2006 to 2007.**

<table>
<thead>
<tr>
<th>2006</th>
<th>2007</th>
<th>SED</th>
<th>ALGDOM</th>
<th>ALGINV</th>
<th>INVDOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Expected losses under a random process of loss (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>78.39</td>
<td>0.72</td>
<td>0.20</td>
<td>1.81</td>
<td></td>
</tr>
<tr>
<td>ALGDOM</td>
<td>1.17</td>
<td>3.74</td>
<td>0.02</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>ALGINV</td>
<td>0.85</td>
<td>0.05</td>
<td>0.39</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>INVDOM</td>
<td>4.30</td>
<td>0.19</td>
<td>0.05</td>
<td>7.82</td>
<td></td>
</tr>
<tr>
<td>(b) Relative differences between observed inter-category transitions and expected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SED</td>
<td>0.00</td>
<td>-0.46</td>
<td>-0.16</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>ALGDOM</td>
<td>-0.33</td>
<td>0.00</td>
<td>22.82</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>ALGINV</td>
<td>-0.54</td>
<td>1.34</td>
<td>0.00</td>
<td>3.69</td>
<td></td>
</tr>
<tr>
<td>INVDOM</td>
<td>0.01</td>
<td>-1.00</td>
<td>3.16</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>
The most distinct systematic gains between the 2007 and 2008 classifications occur between the ALGDOM, ALG/INV and INVDOM classes (Table 4.5b). The relative difference ratio between expected gains under a random process of gain and observed gains between the ALGDOM and ALG/INV indicates that when the ALGDOM class gains it systematically replaces the ALG/INV class (14.11). These 2 classes also exhibit a reciprocal gain relationship in that when the ALG/INV class gains it systematically targets the ALGDOM class (8.70) (Figure 4.4.).

A strong systematic gain signal is evident between the ALG/INV and INVDOM classes (10.08) and suggests that when INVDOM gains it systematically gains from the ALG/INV class. Negative two way systematic relationships occurred between the ALGDOM and INVDOM classes. When the ALGDOM class loses to another habitat class it systematically avoids losing to INVDOM and when INVDOM gains it systematically avoids gaining from ALGDOM. This is not surprising given that these two categories are generally separated by the transitional mixed algae and invertebrates class (ALG/INV).
Differences in observed losses compared to expected losses indicate that when the algal dominated class (ALGDOM) loses, it is systematically replaced by the mixed algae and invertebrate class (ALG/INV) (22.82). Similarly, when the ALG/INV class loses it is systematically replaced by the invertebrate dominated class (INVDOM) (3.69).

Figure 4.4 Reciprocal systematic gain/loss relationships occurring at the study site between the years 2006 and 2007.
4.4 Discussion

Monitoring the spatial dynamics of sublittoral benthic habitats beyond the reach of optical sensors and diver surveys is important for marine conservation and planning. This study has demonstrated the applicability of supervised acoustic methods to achieve this end by summarising change in terms of gains and losses, swap (location) change and absolute (net) change of 4 habitats. Further analysis allowed differentiation of systematic from random patterns of change between classes, based on ratios between observed and expected change.

Quantifying change based solely on areal extents would indicate that the most important transitions were occurring between the sediment-dominated (SED) and the invertebrate-dominated (INVDOM) classes, which comprise >85% of the site. Even relatively small changes in these classes have the potential to obscure transitions in smaller classes (Alo & Pontius 2008). The largest change component was found to be random, with ~68% of the total change (~6.5% of the study area) between SED and INVDOM, mostly where reef margins meet unconsolidated sediments in areas of low morphological complexity (Fig. 3). This may be a result of scouring processes caused by sediment movement along reef fringes, but may also reflect the inability of the classification approach to distinguish between bare sand and sparsely populated sponge habitat in the flat and non-morphologically complex reef–sand interface due to similar
acoustic response patterns. The latter interpretation is borne out by overlap between these 2 classes evident in both single date error matrices (Table 2). An amended classification scheme where prolific invertebrate-dominated habitats evident on higher complexity reefs are differentiated from sparse invertebrate communities at reef margins may aid in resolving ambiguity between these classes. A dynamic relationship between biological classes was observed in this study. Concurrent incidences of systematic gains and losses between classes show a clear transition from ALGDOM to ALG/INV and ALG/INV to INVDOM in a positive depth shift of the transition zone between algal-dominated and invertebrate-dominated habitats. *Ecklonia radiata* beds are known to exhibit high variability in biomass over short (seasonal and annual) time scales. At similar latitudes, Wernberg & Goldberg (2008) recorded seasonal variation of ~400% in standing *E. radiata* biomass between spring senescence and summer peaks over a 2 year, bi-monthly sampling regime, observing a 2 month difference in timing of peak biomass between summers. Variation in biomass was related to short-term fluctuation in environmental variables such as water temperature, day length, seasonal quantum light dose and wave height. Systematic habitat transitions observed in this study show patterns consistent with inter-annual variation of canopy density resulting in retraction of kelp cover at the deeper end of its range and subsequent replacement by adjacent habitats.

Central to ecologically meaningful change detection is the accuracy of the single date classified images. Quantitative accuracy assessment of single date classifications is important for understanding developed change detection
results and also in applying those results for decision making. While single date accuracy assessment of classified images has been well discussed in the literature, methods for assessing accuracy of change detection results are less well explored (Foody 2010). As with single date classifications, change detection error can be categorised by errors associated with spatial (location) and classification (thematic) error (Congalton & Green 2009). Spatial error in this case has been largely mitigated by the use of dGPS and accurate co-registration of both serial acoustic layers, and ground reference data with acoustic layers using a USBL positioning system. Classification error is more difficult to control and can arise from inconsistently classified reference data, repeatability of the classification process, radiometric differences between serial acoustic datasets, and acoustic similarity of different habitat classes.

An accepted method of providing a general measure of change detection validation is to multiply overall classification accuracies for each single date classification (Serra et al. 2003, Coppin et al. 2004). Applying this method to the current study yields a change detection accuracy of 86% \((0.93 \times 0.92 \times 100)\), although this measure fails to account for the bias contributed by larger, better defined classes (e.g. SED) to overall classification accuracy. A more conservative error term may be derived by multiplying the 2-date Kappa coefficient of agreement terms which are independent of bias due to class size and provide change detection accuracy of 69% \((0.83 \times 0.83 \times 100; Prenzel & Treitz 2006)\). A number of approaches to establishing generalised confidence measures for change detection have been proposed. An extension of the single date error matrix for assessment of change detection accuracy was outlined by
Congalton & Green (2009) but requires representative sampling of each possible ‘from–to’ change class. Pontius & Lippitt (2006) proposed methods to assess the sensitivity of 2 error matrices in order to determine levels of change above classification error.

Issues inherent in controlling geometric and radiometric distortion between serial acoustic surveys are inextricably linked to the quality of subsequent applications of the data for detecting change. Disparities between time series acoustic datasets may arise as a result of sea conditions at the time of acquisition (Hughes-Clarke 2003), changeability in onboard acquisition parameters (Fonseca & Mayer 2007) and speed and direction of survey (McGonigle et al. 2010). Backscatter intensity values for each habitat class showed considerable variation between surveys (Table 1). Use of a post-classification approach mitigates this issue by comparing independently derived habitat distribution models (Coppin et al.2004). However, this raises concerns for future studies, especially those relying on direct comparison of properties of the acoustic return, e.g. algebraic or image differencing approaches (Singh 1989, Coppin et al. 2004). Calibration of time series backscatter remains an issue that needs to be addressed if these data are to be used effectively for conducting change analysis.

The ability to quantify short-term seasonal changes in distribution of kelp-dominated habitats using an acoustic remote sensing approach holds promise for longer-term monitoring efforts, particularly in view of the geographic range and importance to overall ecosystem functioning associated with these species.
(Wernberg et al. 2011). The dynamic nature of the habitats investigated underscores the requirement to collect contemporary reference data in order to properly represent habitat distribution at a given point in time. To adequately define the extents of temporally variable habitats, further time series data that reflect the variability of controlling environmental drivers are required (Coppin et al. 2004). Consequently, these results should be interpreted within the context of the temporal scale within which they have been measured and the taxonomic resolution of the classification scheme which has been employed. It is, however, important to understand variation over small spatio-temporal scales in order to assess potential future patterns of directional change (Habeeb et al. 2007).

In this assessment of change in patterns of habitat distribution between two discrete classifications, no assumptions regarding habitat condition or potential drivers of habitat change can be made. The study does identify areas of the seafloor susceptible to change and quantifies the significance of inter category transitions within the constraints of the time series analysed and the classification scheme employed. In describing patterns of variation between broadly defined biological communities at the seafloor in a spatially explicit manner, a better understanding of the temporal dynamics of these communities can be obtained.
5. Wave driven exposure as a surrogate for benthic habitat distribution.
5.1 Introduction

Acoustic-based habitat mapping utilises sonar-derived physical variables as proxies to describe the range of abiotic conditions (e.g. substrate type) and processes (e.g. light availability) that define the realised niche and subsequent distribution of benthic species and assemblages. In this thesis features used to predict the distribution of benthic species assemblages have thus far been derived directly from the bathymetry and acoustic backscatter response provided by MBES. In previous chapters, the role of wave exposure on habitat distribution was only indirectly considered through postulated associations with water depth and seafloor orientation (aspect). Wave energy, however, varies spatially and temporally, and is locally modified by factors such as coastline geometry and bottom topography. It is therefore unlikely in shallow coastal zones that depth and orientation of an area of the seafloor are fully indicative of structuring effects of exposure on the biological community that exists there, especially in areas which are known to experience pronounced gradients in wave activity. This study investigates the effectiveness of incorporating a proxy for wave exposure for benthic habitat characterisation that may better represent the spatial distribution and degree of wave energy than proxies based solely on acoustic data.

The southern Australian coastline is one of the highest energy coastlines in the world (Hemer et al. 2008, Hughes & Heap 2010). As a result, wave energy is arguably one of the primary variables influencing the morphology, community
structure and spatial organisation of benthic taxa in the region (Wernberg & Goldberg 2008, Wernberg & Vanderklift 2010). The effects of wave energy on
the composition, functional morphology and distribution of species and
assemblages have been documented in most areas of the shallow marine
environment across a wide range of taxonomic groups. The hydrodynamic
energy regime has been demonstrated as an important factor controlling the
spatial distribution of macroalgae (Pedersen et al. 2012, Thomson et al. 2012),
 sessile invertebrates (Bell & Barnes 2000, Chollett & Mumby 2012),
 seagrasses (Fonseca & Bell 1998, Turner et al. 1999), molluscs (Boulding et
al. 1999, Pfaff et al. 2011) and fishes (Letourneur 1996, Friedlander et al.
2003), and has been identified as a key indicator of species abundance and
diversity (Denny 2006).

Wave energy determines benthic habitat availability through a number of direct
and indirect processes which can exert both positive and negative effects on
benthic organisms (Denny 2006). Sessile benthic taxa are reliant on water
circulation for delivery of nutrients and oxygen, timing and dispersal of larvae
and propagules, and removal of waste. Hydrodynamic exposure is also an
important agent of stress and disturbance through sediment flux processes,
specifically abrasion, burial and limitation of light availability (Airoldi 2003),
or mechanical tearing or removal of sessile species from their places of
attachment (Thomsen et al. 2004). On shallow rocky reefs dominated by
 canopy forming kelps, wave energy may also determine canopy size,
morphology and spatial patchiness, influencing understory community
composition through altering light availability, water motion and direct physical abrasion (Toohey et al. 2004).

Of primary interest to marine habitat mapping studies is the role of wave exposure on the distribution of benthic taxa. Despite the strong associations between marine biota and their hydrodynamic environment there are relatively few reported studies that use a direct proxy of exposure as a variable for predictive mapping (Kostylev & Hannah 2007, Hill et al. 2010, Harris & Hughes 2012). Quantitative estimation by cartographic fetch models or more complex mathematical simulations of sea state have been used to derive exposure/organism relationships and also to predict their distributional patterns (Bekkby et al. 2008). At the local scale, cartographic fetch models based on the distance from a given location over which wind waves are able to generate (i.e. distance to barrier) have commonly been used to quantify a metric of exposure often assigned to a fixed number of ordinal categories (Lindegarth & Gamfeldt 2005). Fetch-based exposure models have been demonstrated to respond well in enclosed or semi-enclosed areas where coastal perturbations, inlets or islands are the principle mediators of local wave energy (Ekebom et al. 2003, Greenlaw et al. 2011), but are potentially less applicable to open coasts where submarine topography such as offshore banks or reefs are often the significant factors mediating fully-developed wave conditions originating from remote synoptic events (Chollett & Mumby 2012). Numerical wave modelling approaches are commonly used in coastal engineering applications and are capable of incorporating the combined effects of complex seabed topography and coastlines as well as spatial variation in wave energy caused by shallow
water processes such as refraction, diffraction, wave on wave interactions and energy dissipation due to white-capping and wave breaking. Their use in local-scale ecological studies however has not been widely reported (England et al. 2008). This is potentially due to the computational complexity and expert knowledge required for their implementation (Hill et al. 2010).

The specific aims of this research chapter are to:

1. Generate a depth-attenuated spectral wave model for a site on the Victorian coastline that experiences a range of hydrodynamic conditions.

2. Compare the classification accuracy and interpretation of a habitat map derived using acoustic variables only to that of a habitat classification derived with the addition of modelled wave exposure.
5.2 Methods

5.2.1 Study area

The study was conducted on the Otway coast of Victoria, South Eastern Australia. The site extends approximately 95 km from east to west around Cape Otway, the prominent coastal feature of western Victoria (Figure 5.1). MBES data for the site were acquired in four survey blocks of approximately equal area using a Reson Seabat 101 MBES operating at a frequency of 240kHz aboard the Australian Maritime College vessel R.V. Bluefin. Block 1 was surveyed in November 2005 and blocks 2 to 4 in November 2007. Together, the four survey blocks encompass 624 km² of seafloor ranging in depth from 8m to 79m. Large sandy embayments characterise the site with topographically complex rocky reef systems extending offshore from major headlands. Areas of shallow reef (10-30m) are populated by diverse assemblages of macroalgae which are characterised by the canopy forming kelps *Phyllospora comosa* and *Ecklonia radiata*, while deeper reefs are dominated by communities of sponges and other sessile invertebrates.

The wave climate at the site, like much of the continental margin of southern Australia, is largely dominated by swell waves with little tidal influence. Waves are propagated from west to east moving low pressure systems in the Southern Ocean (Hemer et al. 2008). The majority of Australia’s southern shelf is subject to persistent high energy swells of above 3.5 m 30-50% of the time (Porter-Smith et al. 2004) and annual return significant wave heights of up to 8.7m (Harris & Hughes 2012). The orientation of Cape Otway to prevailing swells originating from the south west quadrant causes a gradient of wave energy across the site from highly exposed on the western side to moderately exposed in the east.
Figure 5.1 Location map of the study site at Cape Otway south-eastern Australia superimposed with hill-shaded bathymetry. Numerals represent each of the four MBES survey blocks undertaken at the site. Letters are associated with detail insets represent major reef systems that are referred to throughout the chapter.
5.2.2 MBES data acquisition and processing

Data acquisition parameters and processing of raw sounding data are detailed in chapters 3 and 4 of this thesis. Processed bathymetry and backscatter grids from each of the 4 survey blocks were combined using the mosaic function in ENVI 4.8 at their highest resolution of 2.5m. Edges between each of the survey blocks were normalised whereby overlapping values at a distance of 50 pixels (250m on ground) from the edge of each block were averaged using a linear ramping technique. In order to minimise misregistration error between MBES products and in situ video observations (Chapter 2) mosaiced bathymetry and backscatter images for the entire site were resampled to a resolution of 5m cell size before further processing.

A suite of environmental (explanatory) data was derived from high resolution MBES datasets using a variety of neighbourhood based topographic and spectral methods. Derivation of these data is detailed in Chapter 3 (3.2.3). MBES bathymetry, backscatter and their derivatives were geographically overlayed to form an image stack of 12 predictor variables analogous to those presented in chapters 3 and 4 (Table 5.1). Further to the previously described set of MBES and MBES derivative predictor variables, a dataset representing energy exposure at the seabed was developed.
Table 5.1 Variables used in the Random Forest models describing scale of analysis and spatial grain.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Analysis Scale</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bathymetry (m)</td>
<td>---</td>
<td>5m</td>
</tr>
<tr>
<td>2. Backscatter intensity (dB)</td>
<td>---</td>
<td>5m</td>
</tr>
<tr>
<td>3. Slope (degrees)</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>4. Complexity</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>5. Maximum Curvature</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>6. Aspect (Eastness)</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>7. Aspect (Northness)</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>8. Rugosity</td>
<td>3x3</td>
<td>5m</td>
</tr>
<tr>
<td>9. HSI$_{R}$ (Red band)</td>
<td>High pass 11x11</td>
<td>5m</td>
</tr>
<tr>
<td></td>
<td>Low pass 5x5</td>
<td></td>
</tr>
<tr>
<td>10. HSI$_{G}$ (Green band)</td>
<td>High pass 11x11</td>
<td>5m</td>
</tr>
<tr>
<td></td>
<td>Low pass 5x5</td>
<td></td>
</tr>
<tr>
<td>11. HSI$_{B}$ (Blue band)</td>
<td>High pass 11x11</td>
<td>5m</td>
</tr>
<tr>
<td></td>
<td>Low pass 5x5</td>
<td></td>
</tr>
<tr>
<td>12. Benthic Position Index</td>
<td>Inner radius 10</td>
<td>5m</td>
</tr>
<tr>
<td></td>
<td>Outer radius 50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scale factor 133</td>
<td></td>
</tr>
<tr>
<td>13. $U_{max}$ (m s$^{-1}$)</td>
<td>---</td>
<td>60m</td>
</tr>
</tbody>
</table>

(Maximum Orbital Velocity)
5.2.3 Exposure model

A fine-scale (60m cell size) estimation of wave induced orbital velocities at the seabed, used here as a surrogate for wave induced exposure, was created using a 4 step process:

1. Results of a global wave hindcast model were downscaled to a regional scale (Victorian coastline) to provide a longer term assessment of the prevailing annual wave climate in this region.

2. A site specific model was created by incorporating local bathymetric variation (MBES and LIDAR derived) within the regional-scale model. Detailed spectral wave models of the Otway coastline (study area) based on one year of representative annual wave conditions were derived from the longer term wave climate assessment.

3. Wave induced orbital velocities transferred to the seabed were estimated by applying linear wave theory to surface spectral wave conditions.

5.2.4 Regional scale model parameterisation

Numerical wave modelling was accomplished using the MIKE 21 spectral wave (SW) model (DHI 2012) applied to a mesh generated from the Geoscience Australia 2009 bathymetry grid (0.025°). MIKE 21 SW is a 3rd generation spectral wind-wave model capable of simulating wave growth by action of wind, non-linear wave-wave interaction, dissipation by white-capping, dissipation by wave breaking, dissipation due to bottom friction, refraction due to depth variations, and wave-current interaction. The model domain incorporated the western and eastern coastlines of Victoria, Tasmania.
and adjacent areas of continental shelf including Bass Strait (Figure 5.2).

Global hindcast results (10m $u$ and $v$ wind velocity) from the National Oceanic and Atmospheric Administration (NOAA) Wave Watch III model were extracted and linearly interpolated (0.25° spatial, 3 hourly temporal) to provide boundary inputs for the regional scale spectral wave model. Spatially and temporally varying open wave results from the NOAA model provided wave boundary conditions along the western, southern and eastern model boundaries.

**5.2.5 Model Calibration**

The regional spectral wave model was calibrated and validated against measured wave buoy data from Cape Sorrel on the west coast of Tasmania (42° 7.2S 145° 0.0E) and Point Lonsdale, south west of Melbourne (38° 18.2S 144° 34.2E) (Figure 5.3) for the year 2000. Comparative agreement of hindcast
wave conditions (significant wave height and peak period) to measured data was considered appropriate to use this model to assess longer term wave climate along the Victorian coastline.

Figure 5.3. Results of the Cape Otway coast spectral wave model hindcast for the year 2000. Summaries of significant wave height (H_{sig}) and direction are shown for waters a) west and b) east of the southernmost point of the cape under typical prevailing south-westerly swell conditions. Extents of study area (Figure 5.1) are detailed in box.
5.2.6 Local scale model parameterisation

A site-specific spectral wave model (Figure 5.3) was generated for the waters around Cape Otway with western, southern and eastern boundary conditions provided by the regional scale model. Of the four years (2000-2004) hindcast using the regional-scale model, the year 2000 was chosen as a year broadly representative of the long term annual wave climate for the region. A spectral wave hindcast was generated for the year 2000 using a combination of the 0.0025° bathymetry grid, local MBES bathymetry (5m) and bathymetric LIDAR (5m) to provide depth attenuation inputs in order to propagate the model up to the coast.

5.2.7 Bottom orbital velocity

Modelled wave conditions corresponding to significant wave height and spectral peak period for the year were used to calculate a spatially explicit estimate of maximum instantaneous bottom orbital velocity ($U_{\text{max}}$), used here as a surrogate for exposure to wave induced energy. Linear wave theory was then used to predict the horizontal component of the wave orbital velocity ($u_o$) at a particular area on the seabed for small-amplitude, monochromatic waves as follows:

\[
u_o = \left[\frac{H \pi}{T \sinh(kd)}\right] \cdot \cos(kx - wt)
\]

Where:

H = wave height (m)

T = wave period (s)
d = water depth (m)

k = wave number

w = radian frequency

x = position

t = time

As \( u_o \) varies sinusoidally through a wave period, the maximum velocity \( u_b \) occurs when \( \cos(kx - wt) = 1 \). Instantaneous maximum seabed orbital velocity was calculated for the entire study site at a resolution of 60m (cell size) and subsequently resampled to match the resolution as the other physical predictor layers. While this did not increase the resolution of the dataset it rendered it compatible with the remaining grids for further processing (i.e. satisfies software requirements).

5.2.8 Observation Data

In situ observational data were collected using an acoustically geolocated towed video sled in February 2006 (MBES survey block 1) and February 2008 (MBES survey blocks 2-4). Video system configuration and methods of deployment are outlined in Chapter 2 (2.2.2). A total of 35 video transects were used to capture the range of depths, topographic and textural diversity at the site determined by visual examination of the MBES bathymetry and backscatter intensity products. Video footage was processed following the methods outlined in Chapter 2 (2.2.9) and organised into 4 broad habitat classes analogous to those defined in Chapter 2 (2.2.9) and Chapter 4 (4.2.5).
5.2.9 Habitat classification

The Random Forests (RF) classification algorithm (Breiman 2001) was used to quantify relationships between environmental data layers and video observations. This procedure is similar to the decision tree approach outlined in chapter 3 (3.2.5). Instead of defining species-environment relationships based on the output of a single classification tree however, the RF procedure trains an ensemble of classification trees and aggregates their results through a majority voting process.

The RF algorithm uses bootstrap samples of the training data and randomly selected subsets of available predictor variables to grow multiple classification trees. For each tree approximately two-thirds of the original training data is sampled with replacement (bagging) and a user-defined number of predictors are drawn randomly from the total pool of candidate variables at each node to create each tree from its bootstrap training set. Trees are left unpruned (i.e. fully fitted to the training data) in order to diminish potential bias introduced by any stopping rules. Limiting the number of variables used to determine splitting at each node serves to reduce computational load in comparison to similar methods such as boosting (Gislason et al. 2006), and creates trees that display low correlation to other trees in the ensemble (Prasad et al. 2006). The algorithm yields an ensemble that can achieve both low bias and low variance (from averaging over a large ensemble of low-bias, high-variance but low correlation trees).
At each bootstrap iteration of the RF process the resultant tree is used to predict those data not included in the training process (‘out of bag’ or OOB observations) and calculate a misclassification rate. Probabilities of membership for the various classes are estimated by the proportions of OOB predictions in each class (Cutler et al. 2007). Each tree gives a unit vote for the most popular class at each input instance and the final classification label is determined by a majority vote of all trees in the ensemble.

5.2.10 Variable importance measures

An advantage of using RF ensemble methods over a single classification tree approach is that OOB samples for each tree can be used to derive measures of variable importance. The importance of a given feature is evaluated based on the difference between the misclassification rate of the OOB data and the misclassification rate if values of a given variable are randomly permuted for the OOB observations and passed down the tree to create new predictions. The differences (averaged over all trees) between the predictions of the modified and original OOB data are normalised by the standard error and provide a measure of how much influence the feature exerts on predictive accuracy. In this way the magnitude of the mean decrease in accuracy for a given feature between the original and modified OOB observations can be used to gauge its importance to the classification process.

In this study the RF procedure was applied using a MATLAB implementation (Jaiantilal 2009) of the code by Breiman and Cutler (available online at http://www.stat.berkeley.edu/users/breiman/). A further routine was developed in the MATLAB environment to enable pixel-based classification using the
native ENVI data format (i.e. image data and training data) (Hasan et al. 2012). Prior to implementation of the RF routine the video observation dataset was split using a stratified (by class) random sampling method to a training dataset (70%) used for model propagation and a validation dataset (30%) withheld for model validation.

Classification rules and importance measures were obtained from two separate implementations of the RF procedure. The first model run included 12 predictor variables derived from and including the primary bathymetry and backscatter products. The second model was run with the addition of a grid layer representing annual maximum orbital velocity at the seabed. The performance of the RF models was evaluated by comparing each one against a subset (30%) of video observation data that was withheld from the modelling process. Global accuracy of each model was established using confusion matrices (Overall accuracy and kappa statistic) (2.2.10), similarly class specific accuracy was derived using metrics of users and producers accuracies (3.2.5)
5.3 Results

5.3.1 Exposure Model

A model representing maximum bottom orbital velocity ($U_{\text{max}}$) for a single year was created using inputs from a global wave model attenuated by a bathymetric surface composed of coarse-scale (~270m) regional bathymetry and fine-scale (5m) local bathymetry. Values of $U_{\text{max}}$ range from 0.5 m/s to 1.36 m/s (Figure 5.4) and are highly variable across the site. The spatial pattern of bottom orbital velocities reflects the bathymetry and orientation to surface wave conditions which arrive predominantly from the south-west quadrant. As a result, highly energetic hydrodynamic conditions at the seabed are evident in the western half of the site reducing to moderate conditions in the eastern portion of the site which is largely sheltered from prevailing wave conditions by Cape Otway.
Figure 5.4. Distribution of modelled maximum orbital velocity ($U_{\text{max}}$) values across the Cape Otway study site. Boxes delineate major reef systems at the site and are analogous to those presented in Figure 1.
5.3.2 Classification comparisons

The RF algorithm was applied to both sets of variables using the same training samples. The RF routine was first run with the number of decision trees \( n_{\text{tree}} \) specified at 200 and 500 trees, but as no discernible decrease in the error rate was observed beyond 200 trees these are the only models that will be presented here. The number of variables selected from the pool of predictor variables for splitting at each node \( m \) was the square root of the number of available predictors (i.e. 3 randomly selected input features at each node), a value which has been commonly used in other implementations of RF (Breiman 2001, Cutler et al. 2007).

Comparison of error rates derived from internal cross validation using the OOB data for each set of variables showed similar reduction of errors as the number of trees permuted increased (Figure 5.5). The model run inclusive of the exposure variable (referred to hereafter as the exposure model) performed slightly better than the model run only with acoustic variables (hereafter referred to as the acoustic model) exhibiting a difference of approximately 10% in internally cross-validated error rates. No overfitting of the data was evident (within 200 iterations) as the error rates approach their asymptotic values. It appears as though the error rate would continue to decline marginally beyond the 200 iterations although reduction of error occurred largely in the first 50 trees.
**Figure 5.5** Out of Bag (OOB) cross-validation error rates plotted against the number of trees permuted for each model used in the study (exposure and acoustic variables in red; acoustic variables only in blue)

Cross-validated classification accuracy metrics obtained from the 30% of observations withheld from the modelling process correspond well with those obtained from internal validation using the OOB data (Table 5.5). Overall accuracy of the exposure model was found to be somewhat higher (93%) than the acoustic model (88%) although it is clear that the notably greater number of pixels available for accuracy assessment resulted in a bias in the estimate of overall accuracy towards the class results for SED. Accuracy as defined by $K$ was higher for the exposure classification (0.87) than the acoustic classification (0.77). A pairwise test for significance of the $K$ statistic for each error matrix (Congalton & Green 2009) revealed a significant difference between the two error matrices ($Z = 13.3$) indicating that the exposure model performed
significantly better than the acoustic model. User’s and producer’s accuracies for each habitat class were found to be universally higher for the exposure model. Increase in accuracy was especially evident for the ALG/INV class which was commonly misclassified as either ALGDOM or INVDOM in the acoustic classification. Producer’s accuracy increased from 47% to 76% and user’s accuracy increased from 68% to 82% in this class with the addition of the exposure layer to the classification.
Table 5.2 Confusion matrices for the classified images derived from the acoustic model (top) and the exposure model. Each column corresponds to the ground reference pixels used for accuracy assessment for a single class. The values in the column indicate the number of those ground observation pixels classified into each class, while the values on the main diagonal (italicised) indicate agreement between observations and classified maps.

<table>
<thead>
<tr>
<th>Reference</th>
<th>%Producer’s Accuracy</th>
<th>%User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALGDOM</td>
<td>ALG/INV</td>
</tr>
<tr>
<td>Acoustic Only (Overall accuracy = 88%; $K = 0.77$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALGDOM</td>
<td>596</td>
<td>113</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>58</td>
<td>251</td>
</tr>
<tr>
<td>INVDOM</td>
<td>43</td>
<td>156</td>
</tr>
<tr>
<td>SED</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>709</td>
<td>531</td>
</tr>
</tbody>
</table>

Acoustic and Exposure (Overall accuracy = 93%; $K = 0.87$)

<table>
<thead>
<tr>
<th></th>
<th>%Producer’s Accuracy</th>
<th>%User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALGDOM</td>
<td>ALG/INV</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>647</td>
<td>50</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>34</td>
<td>404</td>
</tr>
<tr>
<td>INVDOM</td>
<td>16</td>
<td>64</td>
</tr>
<tr>
<td>SED</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>709</td>
<td>531</td>
</tr>
</tbody>
</table>
5.3.3 Variable importance measures

Importance indices from each implementation of RF were obtained by randomly permuting the values for each input variable in the classification in the OOB samples for each tree. Decrease in accuracy caused by effectively removing a particular feature from a tree denotes its relevance to the classification accuracy of that tree. Changes in accuracy as a result of permutation is averaged across all trees in the forest and used to calculate a relative measure of variable importance (permutation importance measure) based on mean decrease in accuracy for each feature used in the classification across all classes.

Variables identified as most important over all classes for the acoustic classification in order of decreasing importance (Figure 5.7) were bathymetry (1), rugosity (8), the backscatter derivative HSIR (9) and backscatter intensity (2). These predictors were also found to be most important in varying degrees to the discrimination of individual habitat classes except for the ALG/INV class which was not well resolved by backscatter intensity. Maximum curvature, the variables representing aspect (northness and eastness) and Benthic Position Index (BPI) were identified as the least important predictors across all habitat classes.

The introduction of the exposure variable to the classification did not appreciably change the relative patterns of contribution of the acoustic variables to the classification. The exposure proxy $U_{max}$ was identified as an important feature (second only to bathymetry) across all habitat classes except...
for the ALG/INV class where it was of primary importance to the
discrimination of that class from all others. The relationship between depth,
exposure and habitat categories east and west of Cape Otway is evident in
Figure 5.6. Habitat classes are particularly well partitioned along the exposure
axis into observations made west and east of Cape Otway. Observations
occurring west of Cape Otway display higher separability between classes,
again along the exposure axis, than those east of Cape Otway which overlap
along the depth axis.

![Graph showing depth vs. maximum orbital velocity](image)

**Figure 5.6.** A sample of 400 randomly selected *in situ* video data points were
used to illustrate how values from the exposure model ($U_{\text{max}}$) (x-axis) varies
with depth (y-axis) and habitat class.
Figure 5.7 Variable importance by mean decrease accuracy of the acoustic model (top) and exposure model (bottom). Variable order is detailed in Table 5.1.
Interpretation of habitat classifications

Decision rules derived from the two RF classifications were executed over the full extents of their respective sets of predictor variables to create full coverage habitat maps of the site. Class coverage in the less accurate acoustic classification were overestimated for the ALGDOM class (8.7%) and the SED class (3.5%) (Table 5.3). Areal coverage was underestimated for the INVDOM class (8.8%) and most notably for the ALG/INV class which was estimated to cover an area 33.8% greater in the exposure classification than the acoustic classification.

Table 5.3 Class area estimations derived from the model containing only acoustic variables (Acoustic) and the model containing both acoustic and exposure variables (Exposure).

<table>
<thead>
<tr>
<th>Habitat Area Areal Difference</th>
<th>Areal Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habitat Area</td>
<td>Acoustic (km²)</td>
</tr>
<tr>
<td>ALGDOM</td>
<td>22.1</td>
</tr>
<tr>
<td>ALG/INV</td>
<td>12.6</td>
</tr>
<tr>
<td>INVDOM</td>
<td>115.4</td>
</tr>
<tr>
<td>SED</td>
<td>473.6</td>
</tr>
<tr>
<td>Total</td>
<td>623.7</td>
</tr>
</tbody>
</table>

Insets detailed in Figures 5.1 and 5.4 representing the major reef systems from the exposed western end of the site (A) (Figure 5.8) through to its more sheltered eastern extents (E) (Figure 5.9) showed a clear trend in the zonation of benthic habitat types achieved by each of the models in the study. At the most exposed area of the site (Inset A) (Figure 5.8) the acoustic classification
predicted a zone of change (ALG/INV) between the ALGDOM and INVDOM classes in accordance with other areas of reef at similar depths, although there are no records of that habitat in the ground observations. The exposure classification however, showed an obvious delineation between macroalgal dominated reef and invertebrate dominated reef with the ALGDOM class extending to approximately 34m, notably deeper than at any other area of the site.

Reef systems depicted in insets B and C showed an opposing trend. While reef coverage of the ALGDOM class appear very similar, predictions of the ALG/INV class by the exposure model showed it to cover a considerably more extensive area and extend to greater depths (ca.40m) than the acoustic model (ca.32m). It also appears that these areas of the site may be the principal source of differences in estimated area of the ALG/INV class between the two classifications.

By contrast, there appear to be only relatively small differences between the two habitat classifications on the sheltered eastern side of the site. Both classifications visually show a similar pattern of habitat distributions ( Insets D and E).
Figure 5.8 Habitat characterisations of representative reef systems west of Cape Otway overlaid with 10m isobaths and classified observation data from towed video transects. Insets shown are identified in Figures 5.1 and 5.44. Classifications on the left hand side of the figure were derived using acoustic only variables, while those on the right were derived using an additional exposure dataset.
Figure 5.9 Habitat characterisations of representative reef systems east of Cape Otway overlaid with 10m isobaths and classified observation data from towed video transects. Insets shown are identified in Figures 5.1 and 5.4. Classifications on the left hand side of the figure were derived using acoustic only variables, while those on the right were derived using an additional exposure dataset.
5.4 Discussion

This study has demonstrated that a model incorporating a direct measure of wave exposure performs significantly better than one incorporating indirect acoustic proxies alone in describing patterns of benthic habitat distribution at a wave exposed site in temperate Australia. Improvement in the model was largely due to increased separability of shallow reef habitat types along an exposure gradient allowing their distribution to be more accurately predicted. This pattern was observed in cross-validation measures which showed an improvement in classification accuracy of all habitat categories used in the study and also in measures of variable importance derived using a Random Forests classification approach. Map interpretation showed that estimates of habitat area differed between classifications particularly that of a transitional habitat between shallow algal dominated and deeper reef characterised by sponge dominated sessile invertebrates.

The variables of primary importance to classification accuracy of the best performing model were bathymetry, exposure, rugosity, backscatter intensity and HSI backscatter derivatives. This pattern is also reported by Bekkby et al., (2009) in their distribution modelling study of the kelp *Laminaria hyperborea* where depth, terrain curvature, wave and light exposure were found to be the most important geophysical factors explaining the distribution of the species. Similarly depth, slope, wave and light exposure were found to best explain the
potential distribution of the fucoid kelp *Saccharina latissima* in Norwegian waters (Bekkby & Moy 2011).

Differentiation of benthic habitats in this study was also largely determined by proxies of light availability (bathymetry), availability of suitable substrate for attachment (rugosity) and hydrodynamic energy (seabed orbital velocity). While the contribution of backscatter intensity and its HSI derivatives is more difficult to interpret, it is surmised that these products are important to the classifier in distinguishing textural differences between inhomogeneous substrate types that are indicative of suitable areas for attachment of sessile species. It may be that by filtering the noise inherent in the backscatter intensity feature and optimising various ranges within the overall dataset (nominally R, G, and B) that the HSI transformed backscatter products allow the classifier to partition the data more effectively.

The high relative contribution of the depth and exposure variables in explaining habitat distribution, in particular for the two classes defined by the presence of canopy forming kelp (ALG/INV and ALGDOM), is well supported by the ecological relevance of these features in explaining distribution of the kelp *Ecklonia radiata*, which accounted for greater than 95% of canopy forming kelp observations. Bathymetry acts as an indirect mediator for light availability, an essential prerequisite for marine algae, and one that limits the depth at which they are able to meet the basic requirements for photosynthesis. The influence of wave energy is also attenuated by depth although it is evident that bathymetry did not capture the full influence of exposure as a variable structuring the distribution of habitats in this study.
The four habitat classes used in the study displayed distinct patterns of distribution across the site that seem likely to be mediated by wave energy. In the moderately exposed eastern side of the site there were only minor differences between classifications produced with and without the inclusion of the exposure variable which was limited in its range from 0.54 ms$^{-1}$ to 0.74 ms$^{-1}$. In both models the kelp dominated (ALGDOM) class transitions to invertebrate dominated reef (INVDOM) through a narrow depth band (5-7m) of the mixed algae and invertebrates class (ALG/INV) (Figure 5.9 insets D and E). The kelp _E. radiata_ which is a defining feature of both the ALGDOM and ALG/INV classes is restricted in vertical distribution to depths less than 30m beyond which invertebrate dominated reef becomes the primary reef habitat type.

West of Cape Otway a different pattern emerges, with observations of _E. radiata_ extending to depths of 49m and the transition zone between algal and invertebrate dominated reef types spanning a greater depth range (~25m). This east-west variation in depth distributions is captured to some extent by the model incorporating only acoustic variables which predicts the ALG/INV class occurring marginally deeper (33m) than in the east but is better described by the exposure model which predicts the ALG/INV class to occur both deeper (42m) and across a greater depth range (~20m).

Differences in patterns of distribution of the ALGDOM class between reef A, and reefs B and C on the western side of the site are potentially caused by a temporal mismatch in collection of observation data which in MBES survey block 1 (Figure 5.1) was collected 2 years prior to the remainder of the study.
site. Although observation data were collected in February of each year there is evidence to suggest that canopy density is temporally variable and largely dependent on the timing of optimum environmental conditions conducive to growth, for example temperature, nutrient and light availability (Wernberg & Goldberg 2008). It is therefore conceivable that the observation data associated with survey block 1 represents a different stage on the annual senescence to peak biomass cycle of the species and in that respect is not consistent with observational data from the rest of the site.

Alternatively, these differences may reflect the interaction of incoming wave energy with local reef geometry which is noticeably different between reefs. The reef shown in inset A displays a relatively steep and regular offshore gradient with little topographic diversity descending to depths of 60m close to the coast. Reefs depicted in insets B and C by comparison have a shallower offshore gradient and are characterised by rugged terrain composed of medium to high-profile crests (<1-2m), troughs and ridges extending farther offshore. While modelled maximum orbital velocity is of similar magnitude for all of these areas, the complexity of the reef surface at areas B and C may provide a wider range of hydrodynamic conditions caused by localised topographic diversity allowing the establishment of invertebrate communities in a mosaic of lower flow areas within the reef. This theory is corroborated by the work of Toohey and Kendrick (2008) who linked greater species richness on reefs with complex topography to a reduction in the structuring effects of *E. radiata* canopy on understorey communities.
It would seem that the role depth plays as an indirect proxy for light availability is strongly mediated by variation in hydrodynamic energy. If water column conditions (i.e. turbidity) are similar across the study site this may indicate that light is not the limiting factor in the vertical distribution of *E. radiata* in the east of the site. The limited depths attained by the species compared to the west of the site are potentially the result of competitive interactions with sessile invertebrates for limited hard substrate suitable for attachment. Under this assumption it can also be suggested that stronger exposure conditions in the west of the site afford some measure of competitive advantage to *E. radiata* allowing the species to successfully occupy space to a greater depth. This contention is supported by the known ecology of the species which exhibits a plastic morphology in response to hydrodynamic stress. Individuals at exposed sites have been reported to display drag reducing morphological characters such as smaller size, narrower laterals and blades as well as thicker holdfast and stipes (Wernberg & Thomsen 2005, Wernberg & Vanderklift 2010). Higher energy conditions may also control the influence of *E. radiata* on understorey communities through increased effects of direct physical abrasion by fronds (Toohey et al. 2004, Fowler-Walker et al. 2005). There is also evidence to suggest that some kelp species achieve a higher rate of primary productivity, increasing both individual density and canopy biomass in high versus low exposure environments (Hurd 2000). Pederson et al., (2012) relate this pattern to higher epiphytic load and self-shading in low exposure sites and speculate that higher exposure conditions may increase light availability to the canopy through continuous and frequent movement.
The results presented have increased our knowledge of the structuring effects of wave exposure on subtidal habitats and demonstrated its relevance to benthic habitat mapping. There are however a number of limitations concerning derivation of the exposure model that should be considered when interpreting these results. Foremost, the temporal resolution of the spectral wave model used to calculate orbital velocity at the seabed is restricted to a single year which may not have fully captured the upper range of extreme wave conditions experienced at the site. Significant wave heights modelled in this study did not exceed 6.2m for the year 2000 although Hemer et al., (2008) estimate a centennial return significant wave height of 15.51m for Cape Sorell (Figure 5.2) and cite a 13.2m event measured by the wave buoy in 1985. Therefore habitat structuring by wave energy at the site could well be the result of larger wave events occurring outside the temporal resolution of the study.

Secondly, the spatial grain of the exposure model (60m) was observed in a small number of cases to cause block artefacts in the habitat classification, predominantly in the areas classified as ALG/INV. This is presumably a function of the value of the exposure dataset in defining these areas and is of consequence as it potentially masks fine-scale variation in habitat boundaries important in analysis of patch metrics (e.g. Ierodiaconou et al. 2011).

Exposure to hydrodynamic energy is one of the fundamental variables of the coastal environment (Nishihara & Terada 2010) and has been well demonstrated to play in integral role in the life histories and evolutionary biology of the organisms that live there (Hurd 2000). There is a wealth of evidence to suggest that the degree of adaptation to varying levels of exposure
strongly influences the available niche of many species. Despite the evidence linking the distributional ecology of marine taxa to the physical aspects of their hydrodynamic environment there are few studies that explore the application of these variables for local-scale (10’s -100’s km²) predictive distribution modelling. In this study benthic habitats at a wave exposed site were characterised according to environmental variables obtained from acoustic methods only, and compared with a characterisation based on the addition of a fine-scale exposure model. Measures of classification accuracy obtained with the addition of the exposure variable to the model were significantly higher overall and contributed to greater resolvability between habitat classes than sonar-derived variables alone. Furthermore, an insight was gained into the interaction between the structuring effects of depth (a proxy for light availability) and exposure to wave energy over the full depth range of a foundation kelp species that affects biodiversity and ecological functioning on shallow reefs across temperate Australasia. This study highlights the suitability of exposure measures for predictive habitat modelling on wave exposed coastlines and provides a basis for continuing work relating patterns of biological distribution to measurable aspects of the physical environment.
6. Summary and Key findings
An explicit knowledge of the spatial arrangement of key benthic community
types (habitats) is an essential prerequisite for their management and protection
under the core tenets of an ecosystem based management paradigm. Creating
suitable habitat information for natural resource planning and prioritising
conservation activities requires methods to integrate data from a variety of
mapping and sampling systems, as well as the ability to incorporate existing
environmental datasets. Acceptable approaches for benthic habitat
characterisation must ultimately be ecologically valid and operationally
feasible. This thesis has identified and examined several important issues
central to the advancement of our knowledge concerning the way in which the
distribution and status of subtidal biological habitats is assessed. The rationale
motivating the work presented here is the provision of accurate and
comprehensive models of habitat distribution at a scale and resolution
applicable to management.

The foundation of the ‘bottom up’ approaches to marine habitat classification
used in this thesis is to characterise organisms’ response to measurable
characteristics of the physical environment and to use those relationships to
predict their distributions. While analysis of species-environment relationships
is not new to ecology, it is the new found ability to measure physical attributes
of the environment at high resolution across broad spatial scales that has driven
the rapid evolution of benthic habitat mapping as a field in its own right.
Improvement of the resolution and ecological validity of models derived from
the new generation of survey and sampling tools (i.e. how closely they
represent real patterns of biological distribution) is unarguably the next major goal of the evolving discipline.

Based on the recognition that the effectiveness of current approaches to management of coastal marine areas is limited by a paucity of spatially explicit information on benthic habitats, and the requirement that habitat distribution models are ecologically valid, four core research objectives were identified in the introductory chapter:

1. *To quantify and critically evaluate potential sources of uncertainty in the interpretation and integration of underwater video used to inform supervised benthic habitat classifications.*

2. *To test the influence of MBES backscatter and bathymetry on the accuracy of predictive benthic habitat distribution models.*

3. To test the influence of MBES backscatter and bathymetry on the accuracy of predictive benthic habitat distribution models.

4. *To assess the impact of hydrodynamic energy on the accuracy of benthic habitat classification.*

These objectives were fulfilled through the accomplishment of four independent research chapters. As a whole they represent a novel contribution to our understanding of both the nature of spatial variation in near shore benthic habitats, and the tools and approaches used to predict their distributions.

The quality of data input to a model is of key importance to the reliability of subsequent predictions and can adversely affect its suitability for management
applications. In addition to acoustic data which provide abiotic environmental measures of habitat surrogacy, supervised (bottom up) mapping approaches require information to provide biological context to environmental patterns.

The relevance of underwater video methods to provide this information is reflected in their widespread application for benthic habitat mapping. In Chapter 2 the first research objective was met by addressing the implicit assumptions that video observation data provide a measure of the biological component of habitat that is free from positional or thematic (attribution) error. Significant differences were found in the way that observers interpreted still video images, which became more evident with increasing complexity of a relatively simple benthic habitat classification scheme. Differences in class attribution between observers were related to differences in qualitative estimates of density and misidentification of potentially cryptic taxa that were a defining feature of a particular habitat class. A spatial error budget created for a commercially available acoustic positioning system revealed that error propagated through the various components of the system was substantially higher than the ‘percentage of slant range’ error term often quoted by manufacturers and reported in the literature.

These findings are applicable to video-based marine sampling methods in general. In the first instance as some level of interpretive subjectivity is unavoidable in video classification, especially for large or complex classification schemes, where habitat differentiation is based on ever finer differences between images. The work raises a range of methodological implications for video applications including identification of a repeatable
taxonomic resolution for benthic habitat classification schemes, and mitigation of the potential for video/acoustic data misregistration effects through informed project planning and practice. Similar issues surrounding spatial and thematic uncertainty in reference data are raised in the terrestrial remote sensing literature (e.g. Powell et al. 2004, Wulder et al. 2007) but have not yet been fully investigated in relation to marine studies. Additional work is needed to further investigate the potential for observer bias in interpretation of underwater video and also to assess the effects of imperfect observation data on subsequent predictive models of habitat distribution.

Chapter 3 provided an implementation of a decision tree approach to establishing relationships between benthic habitat observations and a suite of environmental variables derived from MBES survey data. The QUEST decision classifiers used were found to be well suited for integrating MBES acoustic data to produce detailed and accurate maps of benthic habitat distribution. Differences in distribution models derived using a combination of backscatter and bathymetry derived variables showed that the inclusion of all acoustic variables (bathymetry and backscatter) provided the most accurate classification.

Key findings support the notion that bathymetry and backscatter, assuming that it is of acceptable quality, act as surrogates describing a range of different environmental processes that determine habitat availability (McArthur et al. 2010). Bathymetry derived variables better described habitat classes occurring on hard reef where complexity of terrain features may be a good indicator of hard substratum suitable for attachment. Bathymetry is an indirect proxy for
light availability (Bekkby et al. 2009, Bekkby & Moy 2011) which influences
distribution of habitats containing macroalgae and also mediates wave
exposure conditions.

While alluded to in numerous studies (Harris & Baker 2011), change
assessment using acoustic surrogates of habitat distribution have seldom been
reported in the marine literature (e.g. Grove et al. 2002, Collier & Humber
2007, Van Rein et al. 2011). This is undoubtedly related to the relatively short
period of time that the technology has been available. Comparison of time-
series habitat classifications presented in Chapter 4 demonstrated that
systematic changes in seasonably variable shallow reef habitats were
discernible above random patterns of change.

Information that describes change in biological distributions is useful for
management as it provides a temporal dimension to otherwise static
characterisations of biological habitats. The ability to characterise change in a
spatially explicit manner over wide areas is especially pertinent in light of the
ubiquitous and sobering predictions of changing global environmental
conditions (Worm et al. 2006). In cool-temperate Australian waters, southern
range expansion of a number of warm-temperate species has been linked to
stronger and more frequent incursions of the south flowing Eastern Australian
current brought about by broad-scale changes to circulation patterns in the
South Pacific Gyre (Johnson et al. 2011). These changes may not only have the
effect of altering the distribution of existing habitats (Wernberg et al. 2011) but
may fundamentally alter the structure (and function) of cool-temperate reef
communities (Poloczanska et al. 2007) leading to a range of environmental,
social and economic ‘knock on’ effects. For example, establishment of the barren-forming urchin *Centrostephanus rodgersii* on the east coast of Tasmania has already led to a measurable decline in reef biodiversity and poses a significant threat to the State’s rock lobster and abalone fisheries (Ling 2008).

Change detection seems an inevitable evolution of temporally static benthic habitat mapping approaches and will no doubt continue to evolve (Anderson et al. 2008). The research presented here provides a repeatable and sound methodological basis for subsequent investigations into benthic habitat change assessment both in the temperate Australian context and farther afield. Further research is required to identify natural scales of variability within and between habitats where distributions of key species (e.g. kelps) share close associations with short-term (seasonal to annual) variation in light availability and physico-chemical attributes of the water column. By establishing a baseline for this variability, changes occurring as a result of longer-term environmental perturbations may be more readily distinguishable from those arising from short-term environmental cyclicity.

Extant models representing hydrodynamic energy are generally only available at broad spatial scales such as continental shelves (e.g. Kostylev & Hannah 2007, Harris & Hughes 2012). The grain of these models (100’s m to km) makes them unsuitable for local-scale coastal benthic habitat mapping as they are not representative of shallow water processes such as shoaling, refraction or breaking (Hill et al. 2010). As a result there are few reported applications of predictive habitat modelling using variables representing the hydrodynamic...
regime beyond acoustic surrogates such as water depth (bathymetry) and orientation to prevailing hydrodynamic conditions (aspect).

Research objective 4 was addressed in the final research chapter of the thesis (Chapter 5) in a study that incorporated a fine-scale numerical wave model, extended to the seabed using linear wave theory. Comparison of classifications implemented using the Random Forests algorithm, an ensemble variant of the decision tree approach used in Chapters 3 and 4, established that significantly more accurate characterisations of habitat were obtained using the exposure model than MBES acoustic data alone. Variable importance measures and map interpretation indicated that the exposure model was most influential in discriminating habitat classes containing the canopy forming kelp *Ecklonia radiata* in highly exposed areas.

It is important to note that variable importance measures derived for acoustic features in Chapter 5 must be interpreted with respect to the scale with which the features were derived (3 x 3 analysis windows – 15m²). Wilson et al., (2007) highlight the importance of multi-scale analysis of acoustic data to define species-environment relationships, as species’ responses to their environment may occur over a range of spatial scales. Their findings are supported by a number of subsequent studies demonstrating that distributions of temperate corals (Guinan et al. 2009), fish species (Monk et al. 2011) and benthic communities (Holmes et al. 2008) are best predicted using acoustic variables described at a range of analysis scales. Broad-scale orientation of the coastline to prevailing wave conditions is a factor immediately apparent in contributing to the distribution of wave energy at the site.
6.1.1 Conclusion

The detail of our knowledge of the seafloor beyond safe repeatable diving depths (ca.30m) is generally very sparse (Kostylev & Hannah 2007) and the cost of collecting observational data beyond these limits has historically been very high (Foster et al. 2009). As continuing technological developments improve our ability to quantitatively map and observe physical and biological components of the seafloor, so too does our ability to quantify ecologically relevant patterns of habitat distribution.

The current lack of knowledge on seascape patterns and their ecological consequences represents both a major void in our understanding of marine and coastal ecology and is recognised as an exciting new frontier for research (Pittman et al. 2011). There is much to discover regarding the environmental factors that drive the distribution of marine communities. The identification of representative indicators of the geo-physical factors that influence biotic distributions is critical information that can be used as a surrogate to define patterns in biological habitats over broad scales.

Making sense of these patterns will require a seascape ecology approach for developing and testing novel ecological hypotheses, with substantial potential for providing ecological information at spatial scales that are operationally relevant to management (Pittman et al. 2011). Landscape ecology in the terrestrial realm is comparatively advanced to that in marine systems, where a paucity of spatially explicit data still exists. Rapid advances in marine remote sensing data are beginning to close this knowledge gap and provide an opportunity to gain an understanding of the mechanisms by which key
ecological processes influence spatial patterns across multiple spatial scales.
7. References cited
Bekkby T, Moy FE (2011) Developing spatial models of sugar kelp (Saccharina latissima) potential distribution under natural conditions and areas of its disappearance in Skagerrak. Estuarine Coastal and Shelf Science 95:477-483

170


Fonseca MS, Bell SS (1998) Influence of physical setting on seagrass landscapes near Beaufort, North Carolina, USA. Marine Ecology-Progress Series 171:109-121


(eds) Seafloor Geomorphology as Benthic Habitat: GeoHab Atlas of
seafloor geomorphic features and benthic habitats. Elsevier, Amsterdam
Harris PT, Hughes MG (2012) Predicted benthic disturbance regimes on the
Australian continental shelf: a modelling approach. Marine Ecology-
Progress Series 449:13-25
Methods for Benthic Habitat Mapping Using Backscatter from Multi-Beam
Sonar. Remote Sens 4:3427-3443
repeatability of vegetation classification and mapping. J Environ Manage
92:1174-1184
characteristics during large wave events on the Southern Australian
margin. Continental Shelf Research 28:634-652
Hill NA, Pepper AR, Puotinen ML, Hughes MG, Edgar GJ, Barrett NS, Stuart-Smith
RD, Leaper R (2010) Quantifying wave exposure in shallow temperate
reef systems: Applicability of fetch models for predicting algal
biodiversity. Marine Ecology Progress Series 417:83-95
bottom-types worldwide and implications for coral reef remote sensing.
Remote Sensing of Environment 85:159-173
distribution of marine benthos from hydroacoustics and underwater
video. Continental Shelf Research 28:1800-1810
Hong G, Zhang Y, Mercer B (2009) A Wavelet and IHS Integration Method to Fuse
High Resolution SAR with Moderate Resolution Multispectral Images.
Photogrammetric Engineering and Remote Sensing 75:1213-1223
Howell KL, Holt R, Endrino IP, Stewart H (2011) When the species is also a
habitat: Comparing the predictively modelled distributions of Lophelia
pertusa and the reef habitat it forms. Biological Conservation 144:2656-
2665
sonars: a new tool for investigating seafloor processes in the coastal zone
and on the continental shelf. Marine Geophysical Research 18:607-629
Hughes MG, Heap AD (2010) National-scale wave energy resource assessment for
Australia. Renewable Energy 35:1783-1791
Journal of Phycology 36:453-472
Iampietro PJ, Kvitek RG, Morris E (2005) Recent advances in automated genus-
specific marine habitat mapping enabled by high-resolution multibeam
mapping using multibeam data, georeferenced video and image
classification techniques in Victoria, Australia. Journal of Spatial Science
52
Ierodiaconou D, Laurenson L, Leblanc M, Stagnitti F, Duff G, Salzman S, Versace V
(2005) The consequences of land use change on nutrient exports: a
regional scale assessment in south-west Victoria, Australia. J Environ
Manage 74:305-316
of automated classification techniques for predicting benthic biological
communities using hydroacoustics and video observations. Continental
Shelf Research 31:S28-S38


Monk J, Ierodiaconou D, Harvey E, Rattray A, Versace VL (2012) Are We Predicting the Actual or Apparent Distribution of Temperate Marine Fishes? PLoS ONE 7:e34558

only modelling and multibeam sonar. Marine Ecology-Progress Series 420:157-174


Thomson DP, Babcock RC, Vanderklift MA, Symonds G, Gunson JR (2012) Evidence for persistent patch structure on temperate reefs and multiple hypotheses for their creation and maintenance. Estuarine, Coastal and Shelf Science 96:105-113

associated foliose algal assemblage: the importance of photoacclimation. Marine Biology 144:1019-1027
Van Rein H, Brown CJ, Quinn R, Breen J, Schauenburg W (2011) An evaluation of acoustic seabed classification techniques for marine biotope monitoring over broad-scales (> 1 km(2)) and meso-scales (10 m(2)-1 km(2)). Estuarine Coastal and Shelf Science 93:336-349
Wall CC, Donahue BT, Naar DF, Mann DA (2011) Spatial and temporal variability of red grouper holes within Steamboat Lumps Marine Reserve, Gulf of Mexico. Marine Ecology-Progress Series 431:243-254