UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Ocean and Earth Science

National Oceanography Centre Southampton

Evaluation of local- and medium-scale habitat heterogeneity as proxy for biodiversity in deep-sea habitats

by

Katleen Robert

B.Sc., McGill University, 2008

M.Sc., University of Victoria, 2011

Thesis for the degree of Doctor of Philosophy

November 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Ocean and Earth Science

Thesis for the degree of Doctor of Philosophy

EVALUATION OF LOCAL- AND MEDIUM-SCALE HABITAT HETEROGENEITY AS PROXY FOR BIODIVERSITY IN DEEP-SEA HABITATS

by Katleen Robert

The deep sea represents the largest biome on earth, and for most of it, no maps of resolutions comparable to terrestrial environments are available. As new species continue to be discovered, it is clear that our knowledge of species spatial patterns is insufficient to properly inform marine spatial planning, and for complex habitats, high-resolution surveys are crucial for understanding species-environment relationships. This thesis examined two deep-sea areas of the NE Atlantic, Rockall Bank and Whittard Canyon. By linking acoustic maps to benthic imagery datasets, environmental variables describing the spatial arrangement of different substratum types and topographic variability were found to be good predictors of species composition and biodiversity. Employing an ensemble of statistical techniques provided a more robust approach for the creation of biological full-coverage predictive maps and allowed for the identification of areas with high biodiversity. With these maps, it was possible to demonstrate that biological spatial patterns in Whittard Canyon required mapping resolutions of 20-50m while the more heterogeneous Rockall Bank area needed to be mapped at <5m. The continued sparse availability of biological datasets in the deep-sea remains a significant limiting factor in informing conservation needs, but the work carried out shows improvements over previous approaches, and can be applied to identify biodiversity hotpots and assess habitat suitability for vulnerable marine ecosystems, such as cold-water corals. Through such hierarchical multi-disciplinary studies, the currently available biological information can be employed to increase our understanding of the relationships between habitat heterogeneity and biodiversity as well as help establish the baseline state of these ecosystems in order to effectively monitor potential impacts.

Table of Contents

List of Tables

Supporting Information

List of Figures

Supporting Information

Appendices

Declaration of Authorship

I, Katleen Robert, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Evaluation of local- and medium-scale habitat heterogeneity as proxy for biodiversity in deep-sea habitats

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as:

Robert K, Jones DOB, Huvenne VAI (2014) Megafaunal distribution and biodiversity in a heterogeneous landscape: The iceberg scoured Rockall Bank, NE Atlantic. Marine Ecology Progress Series 501:67-88

Robert K, Jones DOB, Tyler PA, Van Rooij D, Huvenne VAI (*in press*) Finding the hot-spots within a biodiversity hotspot: fine-scale biological predictions within a submarine canyon using high-resolution acoustic mapping techniques. Marine Ecology.

Signed:

Date: November 3rd, 2014

Acknowledgements

Dedication

A mon Amour et ma famille, votre support, même de loin,
me fut indispensable

Chapter 1:
Introduction

"In all deep-sea investigation it is of course of the first importance to have a means of determining the depth to the last degree of accuracy, and this is not so easy a matter as might be at first supposed" (Thomson 1874).

1.1   Background and Mapping Evolution

With an increasing interest for exploration and possible exploitation of resources in the deeper areas of our oceans, it becomes increasingly important to quantify the spatial variability of the seafloor in order to understand the diversity of habitats present as well as the species they harbour. Although the spatial heterogeneity (the uneven distribution of environmental variables over space) of the seafloor is now well recognized (Thistle 1983, Grassle 1991, Levin & Sibuet 2012), it had previously been thought that the seafloor was a mostly homogeneous environment exhibiting long-term stability (Sanders 1968). This belief changed with our greater ability to map the deep sea at increasingly higher resolutions.

1.1.1   Seafloor Mapping

At the onset, seafloor depths were measured using a lead line (Thomson 1874). Obviously the time required to send a weight to the seabed, particularly in deeper regions, greatly reduced the number of measurements that could be taken and hence the resolution of the maps created (Figure 1.1). The ability to take acoustic measurements, by sending an acoustic signal to the seabed and determining depth based on the time taken to receive the echo, greatly enhanced mapping possibilities (Slee 1932, Colman 1933). As implied by the name, a single signal was sent to the seafloor as the vessel moved along its survey line. This technique increased speed of acquisition, but still resulted in disjointed lines of measurements for which interpolation was still required. In any case, single beam echo sounders proved useful in studying geological

features at greater resolutions than previously possible (Menard & Dietz 1951, Laughton et al. 1960). From the late 1970s (Farr 1980), multibeam sonars started appearing in seafloor mapping studies. Their major improvement involved a fan-shaped beam array which allowed for the acquisition of a wide swath of data across the path taken by the vessel (Courtney & Shaw 2000). In addition to measuring the time taken for an echo to return, it became also possible to measure the intensity of the return signal, or its backscatter (de Moustier 1986). In areas of hard seabed a strong signal is returned, while soft sediments cause the signal to be absorbed. Before multibeam backscatter was commonly processed into useful products, sidescan sonars had already been used to characterise seafloor sediment facies using the intensity of the acoustic return (Fader & King 1981, Able et al. 1987). They are typically characterised by higher frequencies and as such, provide higher resolutions and have remained in use.

Resolution varies based on the frequency of the signal and this in turn is constrained by the distance between sonar and seafloor. Higher frequency signals provide more resolution, but are attenuated faster in water and as such, cannot be used to map the deep seabed from a vessel (Kenny et al. 2003). Either a lower frequency resulting in coarser resolutions must be used, or the sonar must be mounted on an instrument platform that can be positioned closer to the seabed. Towed platforms were first used, but navigation was problematic; now remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) are increasingly being employed (Wynn et al. 2014). Their location as well as movements (roll, pitch and yaw) are recorded and used in post-processing to ensure accurate spatial positioning and removal of acquisition artefacts. AUVs are particularly well-suited to seabed mapping owing to their constant speed and elevation. In addition, as they do not require continuous ship support, additional work can be carried out concurrently. However, ROVs are still of great value when more complex environments are encountered (Huvenne et al. 2011) for their real-time interactive abilities and capacity to take samples.
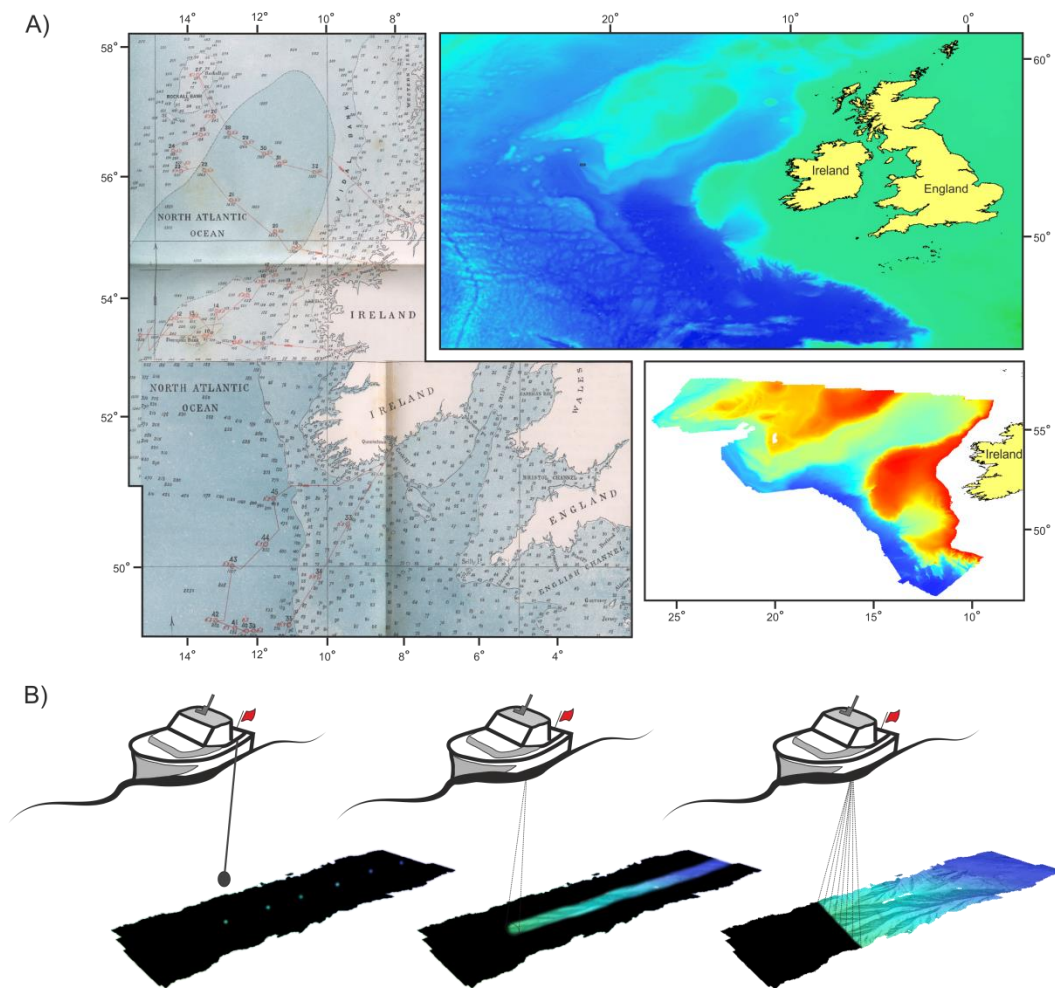
Figure 1.1: Evolution of mapping products A) an image out of Thomson's expedition of 1869 based on lead-line measurements (left) to GEBCO's (General Bathymetric Chart of the Oceans, http://www.gebco.net/) current best available global data (30 arc-second) mostly acquired from amalgamation of ship depth soundings (top right) and the economic exclusion zone of Ireland mapped at 111m resolution through the INFOMAR (Integrated Mapping for the Sustainable Development of Ireland's Marine Resources, http://www.infomar.ie/) project using multibeam sonar (bottom right). B) Pictorial differentiation between lead line, singe-beam and multibeam acquisition techniques and the relative seabed mapped.

1.1.2   Biological Mapping

Niche theory suggests that each species is suited to a select range of environmental conditions (fundamental niche), but that through density dependent effects, such as competition or predation, species may occupy only a portion of their potential habitat (realized niche), resulting in a specialized resource usage distinct from other species (Hutchinson 1959, Vandermeer 1972). In spatially heterogeneous environments, a larger number of

environmental conditions can be encountered, allowing for greater resource partitioning and a greater number of niches to be created, reducing competitive exclusion between species (MacArthur & MacArthur 1961, Pianka 1966). As a result of this, more species can co-occur. More diverse ecosystems are suggested to provide enhanced ecosystem functions and services (the biogeochemical processes needed for maintaining ecosystems, such as providing structural complexity or recycling nutrients) leading to higher productivity (Diversity-Productivity hypothesis). They may also exhibit higher stability and resilience to disturbances (Diversity-Stability hypothesis) (MacArthur 1955, Tilman 2001). Biodiversity can enhance ecosystem functioning by ensuring that the available resources are utilized more thoroughly as well as by allowing a broader range of resources to be employed (Complementarity effect) or by increasing the likelihood that a particularly important trait will be present within the community (Selection effect) (Loreau 2000). Biodiversity and ecosystem functioning is of particular importance in a changing environment, as in a more diverse community where many links within the ecosystem are present, it is more likely that other species can continue to carry out important functions despite changes in conditions (Insurance hypothesis) (Yachi & Loreau 1999). Spatial heterogeneity also helps to ensure that populations in less favourable areas ('sinks') are maintained through immigration from more suitable habitats ('sources') (Levins & Culver 1971).

As such, by finding areas exhibiting high habitat heterogeneity, one can potentially use this information as an abiotic proxy or surrogate for more difficult to measure biological diversity (Noss 1990, Ward et al. 1999). There exist many measures of diversity, the simplest being 'Species Richness' or the number of species present within an area, giving equal weight regardless of relative abundance, while other indices, such as the Shannon's Diversity Index (Shannon 1948) or Simpson's Diversity Index (Simpson 1949), give more (Shannon) or less weight (Simpson) to rarer species. The latter quantifies the likelihood that two randomly sampled individuals are from different species, while the former represents the uncertainty in predicting a randomly sampled individual's species (Schmitz 2007). When evenness is high (abundance of all species within the community is similar), a random sample could contain any of the species present with equal probability (high uncertainty). On the other

hand, if a single species represents the majority of observations, a random sample is more likely to contain this common species as opposed to rarer ones, implying a less diverse community.

Depending on the spatial scale considered, the main environmental factors responsible for the mosaic structure of the seafloor contribute differently to spatial heterogeneity (Jumars 1976). Early research found that at very broad spatial scales (~100km - 1,000km) depth, or more likely covarying factors such as water mass properties (e.g. oxygen, temperature, current velocity and flux of organic material), create broad regions of similar environmental conditions (Menzies et al. 1973, Rex 1981, Gage 1986). At medium scales (~km - 10km), the presence of large topographic features such as seamounts, submarine canyons or hydrothermal vents contribute to seabed heterogeneity (Haedrich et al. 1980, Grassle 1985, Wilson Jr et al. 1985). The importance of sediment properties (e.g. hardness, grain size and organic content) and disturbance rate appears only at finer spatial scales (~m - ~km) (Hecker 1982, Rhoads & Boyer 1982, Billett et al. 1983, Smith et al. 1986). However, for most of these environmental factors, changes occur across gradients and hard boundaries may not exist or be difficult to determine.

By establishing species-environment relationships and determining which environmental factors are useful in describing a species realized niche, it becomes possible to predict where else a given species is likely to occur (Guisan & Thuiller 2005). As seafloor characteristics such as topography and sediment hardness are known to be important environmental variables controlling benthic species distribution, acoustic mapping can quickly provide full-coverage information. Although suggestions of the use of side-scan sonar for identifying seafloor structures of biological interest (dredge marks and *Posidonia* sp. patches) appeared earlier (Newton & Stefanon 1975), its use in deeper waters for studying fauna-sediment relationships appears to arise from the use of acoustic techniques in fisheries research (Dunton et al. 1983) or oil and gas exploration (Gettleson et al. 1982). Some of the first studies used side-scan sonar maps to link areas exhibiting specific acoustic signals to suitable fish habitats (Able et al. 1987, Edsall et al. 1989). Manual interpretations were also soon used to produce sediment maps which were then related to visual observations of communities (Schneider et al. 1987,

Phillips et al. 1990). With the understanding that acoustic mapping may provide a cost-efficient survey method for sediment characterisation, automated classification methods of acoustic signals from echo sounders were developed through commercial programs (known as 'acoustic ground discrimination systems') such as RoxAnn™ (Stenmar Ltd, Aberdeen) (Schiagintweit 1993, Magorrian et al. 1995) and QTC™ (Questor Tangent Corporation, Sidney BC, Canada)(Galloway & Collins 1998, Morrison et al. 2001), which was later extended to other types of backscatter (Preston et al. 2001).

Yoklavich et al. (1993, 1995) and Greene et al. (1995) showed some of the first attempts at developing a geographical information system (GIS) composed of sidescan and multibeam layers as well as sub-bottom profiles linked to underwater observations in order to examine rockfish and their associated habitats in a conceptual framework which is still employed today (Figure 1.2). The usefulness of side-scan sonar in describing the spatial arrangement of habitat patches to study the potential influence of landscape characteristics on species distributions was also recognized (Auster et al. 1995). Subsequently, one of the first standardized deep seabed classification schemes that helped define the terminology to improve communication between geologists, geophysicists and biologists was created by Greene et al. (1999). This ensured that habitat mapping evolved into a multi-disciplinary field whose aims are still valid in nowadays research: "to characterize the seabed in terms of texture and morphology, sediment movement, effects of physical disturbance by storms, trawling and dredging, distribution of benthic species, and dependence of species on particular habitats for survival" (Valentine & Schmuck 1995).

The use of multiple environmental layers, including second-order derivatives from backscatter and bathymetry, also soon started to appear (Diaz 2000) and detailed community-based analysis of species assemblages and their association with physical variables quickly became a significant part of habitat mapping studies (Riegl & Piller 2000, Brown et al. 2001, Kostylev et al. 2001). Building on this early research, in 2001, the GeoHab (marine GEOlogical and Biological HABitat Mapping, http://geohab.org/) organisation was established and since then has promoted the formation of an international

association of marine scientists through its annual conference. Interest for this field has been increasing and biological mapping is becoming more important.
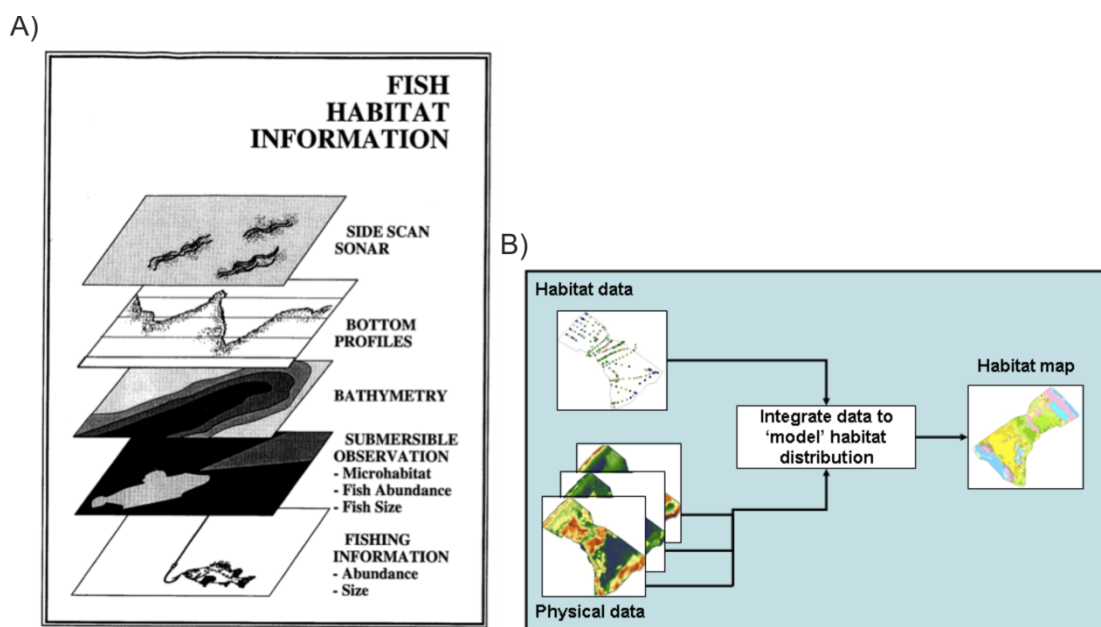
Figure 1.2 Conceptual idea of the habitat mapping process A) as presented by Yoklavic et al. (1995) and B) its continuous application and evolution as illustrated by an example drawn from the MESH program (Mapping European Seabed Habitats http://www.searchmesh.net, 2010).

Nonetheless, even now, only a few countries (Norway; MAREANO, http://www.mareano.no/, Ireland; INFOMAR, http://www.infomar.ie/, UK; MAREMAP, http://www.maremap.ac.uk/; USA, http://coast.noaa.gov/ digitalcoast/) are currently endeavouring to produce broad-scale bathymetric maps covering large extents of their exclusive economic zone (EEZ). However, the biological sampling needed to convert these topographic seabed maps into habitat maps is still lagging behind. In Europe, large scale efforts in collaborative projects such as EUNIS (European Nature Information System, http://eunis.eea.europa.eu/, Figure 1.3), MESH (Mapping European Seabed Habitats, http://www.searchmesh.net/) and EMODnet (European Marine Observation and Data Network, http://www.emodnet.eu/) have started to inventory and map broad-scale biotopes, making information freely available online as part of their deliverables. Similarly, Australia has created seascape maps, delineating areas of similar physical properties for the extent of its

continental margin (Whiteway et al. 2007, Lucieer & Lucieer 2009, Last et al. 2010), while a first classification for South African waters has also been recently released (Sink et al. 2011). A large-scale inventory of the species present in Finnish waters was carried out as part of VELMU (http://www.ymparisto.fi/en-us/VELMU). While its mapping component, FINMARINET, targeted specific marine protected areas and focused on key species and habitats. Although providing crucial information, many areas, particularly in deeper regions, are strongly limited in resolution and only a few biotopes are described in regions strongly suspected to be more biologically complex (Figure 1.3).
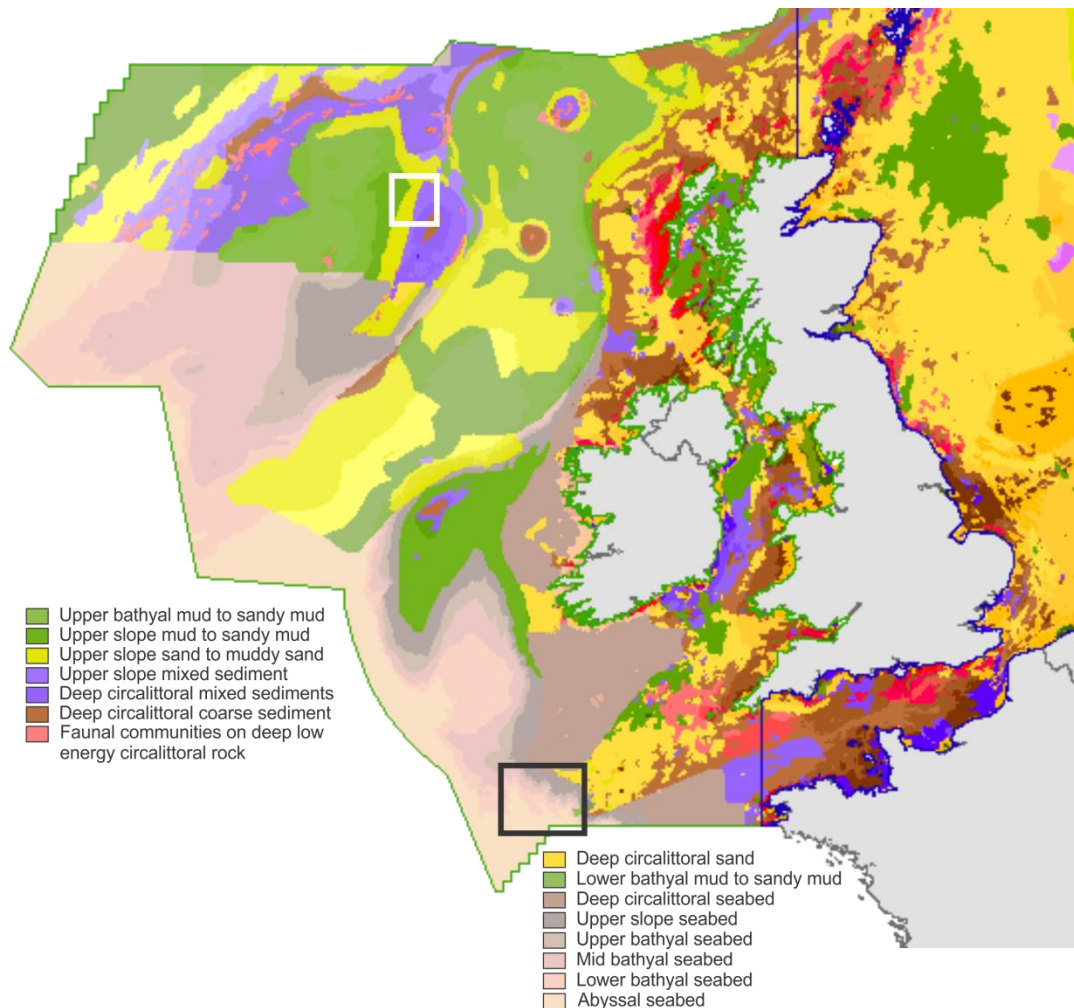
Figure 1.3: Map of the EUNIS classification for the Celtic and North Seas from the EUSeaMap interactive mapping portal (http://jncc.defra.gov.uk/page-5040). For simplicity, only habitats described as present within the two sites investigated are labelled. Rockall Bank is identified in white and the Whittard Canyon in black.

1.2   Thesis Rationale and Objectives

While the foundations for wide-spread habitat mapping have been laid in the past decades, and the first countries are implementing comprehensive mapping programmes (as discussed above), the next challenges in the field are associated with improving classification accuracy and predictive ability by creating automated objective methods able to handle large datasets over multiple scales.

1.2.1   Scale Issue

Multi-scale studies are needed to describe how patterns in biodiversity and spatial heterogeneity vary across scales in order to identify driving ecological processes, particularly since multiple concurrent processes may generate patterns at a different scale (Levin 1992, Chave 2013). Environmental variables useful in explaining species distributions at one scale may lose their predictive ability when considered at another scale and mismatches between the scales at which environmental and biological variables are measured can compound this issue. Choice of resolution, spatial extent and sampling scheme are all factors likely to effect the ecological inferences which can be drawn (Dungan et al. 2002). Extent examined can have a particularly strong effect on species diversity as the greater the area surveyed, the more species are likely to be sampled, at least until all the species within a specific system have been sampled (Species-Area relationship) (Cain 1938). In a larger area, higher species richness is expected as immigration rates increase and extinction rates decrease (Theory of island biogeography), while spatial heterogeneity is also likely to increase (MacArthur & Wilson 1963, Williams 1964). On the other hand resolution will influence the level of spatial heterogeneity than can be captured, with small dispersed features being lost fastest (Turner et al. 1989b), while distance between samples will affect the ability of a study to capture and reconstruct existing spatial structures (Fortin et al. 1989). Moreover, if an incorrect extent is used resulting in environmental gradients being truncated, the curve of the species-environment relationship may be misleading (e.g. appearing as linearly increasing as opposed to unimodal) and cause significant issues when attempting to predict outside the originally sampled area (Austin 2007). Similarly, at too coarse a

resolution a pixel may appear to exhibit suitable environmental conditions, but if the variation captured is not at a similar scale to that experienced by the organism, no habitat may be available if all suitable conditions are not encountered at a single location (Guisan & Thuiller 2005).

As our current knowledge of deep-sea biodiversity has remained geographically limited, with the majority of the ocean remaining to be examined, the most appropriate resolution and spatial extent for individual management or conservation decisions are rarely known *a priori*. Conservation measures aiming at protecting ecosystems and conserving biodiversity are regularly designed and implemented at ~km - 100km scales (Halpern 2003, Leslie 2005). On the other hand, conservation measures targeting single species may require habitat descriptions over finer scales. Overall the need is to identify a spatial scale fine enough to retain useful biological information while remaining coarse enough to be economically feasible.

Environmental descriptors obtained from high-resolution acoustic surveys have the potential to increase considerably the information available for marine spatial planning as they are able to show spatial variation in seabed characteristics over a range of scales (~m to ~100km) (Malatesta & Auster 1999). However, despite significant recent advances in acoustic techniques allowing larger areas of the seafloor to be mapped more quickly, 200 ship-years were estimated to still be required in order to produce a complete coverage map of our ocean at resolutions comparable to those available in terrestrial environments (Sandwell et al. 2006). With this estimate only regarding the mapping of bathymetric features, associated biological information on species distribution remains even more limited (Przeslawski et al. 2011).

Biological information on species spatial patterns continues to be lacking for many areas, especially in the deep sea, owing to the high time investment needed for specimen collection and the reliance on taxonomic expertise for species identification. As such, hierarchical approaches involving nested survey designs are often employed. They involve a combination of broader-scale cost-effective geological map creation and detailed biological studies

covering smaller spatial extents (Brown et al. 2011). Larger areas are mapped using instruments such as ship-board multibeam sonars, while smaller seafloor areas are targeted for higher resolution mapping using AUV or ROV mounted systems. In even smaller areas, sediment and biological sampling can be carried out and video surveys used to describe species assemblages. Although the concept of nested surveys is intuitive and clear, there still is a need to develop quantitative approaches capable of bridging the gap between the broad-scale environmental information and the detailed localized biological data in order to capture accurately habitat complexity and create reliable habitat maps. Tools are also needed to make use of broader-scale maps in identifying specific areas for which finer-scale mapping or additional sampling would be particularly valuable.

1.2.2   Predictive Habitat Mapping

Nowadays, two main approaches to building full-coverage predictive habitat maps have emerged, top-down or bottom-up, but a wide variety of statistical techniques and models can be applied in either cases. In top-down approaches, the acoustic maps are separated into acoustically similar regions using automated classifications based on environmental descriptors such as bathymetry and backscatter (Brown et al. 2011). Secondary descriptors such as slope, aspect, curvature, rugosity or bathymetric position index can be derived from the bathymetry, while the backscatter can be made into sediment interpretation maps (Wilson et al. 2007). More and more often, an object-based approach is being employed, using computer software such as eCognition to segment the acoustic maps into 'objects' (continuous groups of pixels showing similarities) which are subsequently classified using sets of rules based on a large suite of characteristics (Lucieer 2008). Biological samples can then be taken from each of the classes identified, in order to characterise the actual habitats and species. Bottom-up methods, on the other hand, involve the fine-scale acquisition of biological information (often via imagery transects), the selection of the optimal acoustically-derived environmental variables useful in explaining the species distributions and the application of these species-environment relationships to extrapolate biological information across the full extent of the acoustic maps.

Few studies have simultaneously compared both approaches, but a common issue with the top-down approach is that multiple species assemblages can often be found within each of the resulting acoustic classes (Hewitt et al. 2004). Conversely, the high spatial heterogeneity and complex species-environment relationships identified using the bottom-up approach may cause misleading interpolations if variability in significant environmental factors is not adequately captured (Kostylev 2002). As both techniques suffer from distinct limitations, it is clear that both are valuable and the most appropriate approach will depend of the objectives as well as the habitats and scales considered.

1.2.3   Objectives and Hypotheses

Based on the challenges in habitat mapping as described above, the aim of this PhD thesis is to increase the amount of biologically relevant information which can be extracted from acoustic maps in order to produce reliable full-coverage biological maps at the most appropriate scale for a given ecological question or conservation goal. In order to do so the following research objectives will be addressed:

- Evaluate the potential of landscape indices in increasing the amount of biological variation which can be explained by sediment interpretation maps

- Examine the importance of vertical structures and habitat complexity measures in structuring spatial patterns in benthic megafaunal communities of a submarine canyon

- Find the most appropriate modelling approach to extrapolate localised biological information to the full extent covered by acoustic maps, using externally acquired datasets to assess classification and prediction accuracy

- Quantify the change in information contained in datasets of varying resolutions and its impact on model performances and resulting predictive maps

- Address these findings with respect to current conservation needs and future information requirements

This PhD thesis will focus on two specific areas of the Northeast Atlantic, Rockall Bank and Whittard Canyon (Figure 1.3), for which more detailed biological maps were needed in order to identify and further examine species and specific habitats of ecological importance. The first and second objectives will be addressed using the Rockall Bank and Whittard Canyon datasets respectively, while the last three objectives will be considered with respect to both study areas. This PhD is part of the CODEMAP project (COmplex Deep-sea Environments: Mapping habitat heterogeneity As Proxy for biodiversity, http://www.codemap.eu/), whose aim is to "develop an integrated, robust and fully 3-dimensional methodology to map complex deep-sea habitats, to quantify their heterogeneity, and to test if the heterogeneity measures derived at different scales reflect epibenthic megafauna biodiversity (Huvenne, ERC Start Grant no. 258482)".

These two geographic areas were selected for their habitat complexity and heterogeneity over differing scales, for the availability of integrated multi-disciplinary datasets from multiple institutions as well as for their ecological importance and conservation relevance. For both areas, nested datasets of broader-scale environmental variables derived from acoustic mapping techniques (multibeam bathymetry and sidescan backscatter) and fine-scale information on megabenthic invertebrate distributions extracted from ROV imagery transects were available at multiple resolutions. On Rockall Bank, focus was given on quantifying fine-scale heterogeneity in substratum composition in a relatively morphologically simple context, while in the case of Whittard Canyon, the effects of broader-scale topographic features were investigated.

Rockall bank is an important site for cold-water coral conservation and a 4,180km^2 area was closed to fisheries in 2007. In 2010, the Joint Nature Conservation Committee (JNCC) nominated a nearly overlapping area (4,300km^2) for a more permanent status as a candidate 'Special Conservation Area' (JNCC 2010a). However, which boundaries were most appropriate for cold-water coral protection remained unclear. This provided an opportunity to apply habitat mapping techniques to address a current conservation issue. In the case of Whittard Canyon, the high spatial variation usually associated with submarine canyons (Tyler et al. 2009, Levin & Sibuet 2012) required that a

wide range of spatial scales were explored. As previously identified as areas of particularly high biodiversity within canyons of the Bay of Biscay (Van Rooij et al. 2010a, Huvenne et al. 2011, Johnson et al. 2013), vertical structures were given particular attention. Although infrequent, they can be inhabited by large colonies of cold-water corals (*Lophelia pertusa)*, limid bivalves (*Acesta excavata)*, and deep-water oysters (*Neopycnodonte zibrowii)*, and may provide refugia against the impacts of trawling (Huvenne et al. 2011). Identifying biodiversity patterns as well as the driving processes and the scales over which they act are fundamental ecological questions whose better understanding are needed in order to implement effective management decisions at an ecosystem level.

In the first data chapter of this thesis, class and landscape metrics were used to explore the effects of the broader landscape structure on megabenthic species distributions. These metrics are used to describe the size, complexity, distribution and diversity of habitat patches composing the seafloor mosaic (Turner & Gardner 1991, McGarigal et al. 2012). The inclusion of these metrics into modelling techniques allows the environmental conditions of the broader surroundings to be described and included when attempting to explain species distribution patterns. For example, the presence of dead wood or a drop stone in a large area of soft sediment is expected to have a greater effect on an area's biodiversity than if located in proximity to an area dominated by hard substratum (Tews et al. 2004). Redundancy analysis was used to explore the relationships between the species assemblages and biodiversity observed on Rockall Bank and the environmental variables.

Building on the first chapter, in the second, models were built that included selected landscape metrics to predict species distributions and biodiversity over the complete extent of the acoustic maps. In order to examine specific species-environment relationships and select only ecologically meaningful environmental descriptors, a suite of bottom-up statistical techniques (redundancy analysis, random forests and maximum entropy) commonly used in predictive habitat mapping were compared and assessed using both a more traditional split-sample assessment as well as a more robust, though time consuming, independently acquired dataset (collected by colleagues from Herriot-Watt University). Recent advances in mapping

approaches were considered in order to employ all of the information provided by the statistical models and build more robust predictions for application in a current conservation issue.

For the third chapter, megafaunal spatial patterns in biological characteristics (abundance, species richness, biodiversity and presence of cold-water corals) were described in relation to bathymetry-derived environmental descriptors such as slope, aspect, rugosity, various measures of curvature and bathymetric position index calculated over multiple scales. The importance of sediment type, as extracted from the imagery, was also addressed. Both general additive models and random forests were employed to create full coverage predictive and uncertainty maps. An independently collected dataset was acquired via a collaboration with Ghent University to ensure that a thorough model assessment could be carried out. A particular focus was given to identifying verticals walls with environmental conditions showing high suitability to harbour highly diverse biological communities or cold-water corals.

As the choice of scale, appeared as a constant factor influencing the conclusions which could be drawn from a particular set of results, the fourth chapter of this thesis focused on the change in information contained in data acquired or processed using different resolutions. Making use of the three bathymetric resolutions available in Whittard Canyon, the datasets were coarsened to obtain a wide range of resolutions and examine changes in Shannon entropy (the amount of information needed to encode a signal) and Kullback–Leibler divergence (a measure of the difference between two probability density functions) (Brunsell et al. 2008). The analyses were also carried out on the Rockall Bank dataset, where estimates of percentage cover of various substratum types were obtained from benthic imagery as well as sediment interpretation maps of sidescan and multibeam backscatter data. For each scale, statistical models were built to determine the effects of resolution on variable selection and prediction outputs.

The approaches presented in this thesis are aimed at modelling species distributions and biological characteristics, but also address some general ecological hypotheses in the context of Rockall Bank and Whittard Canyon:

Hypothesis 1: Areas of higher heterogeneity are expected to provide a higher number of niches and result in higher diversity and species turnover, but which environmental variables may be responsible for the heterogeneity observed at any given scale is not always know a priori. Finer-resolution habitat characterisation over a small extent will cause different environmental variables (sediment composition) to be identified as driving the patterns observed as opposed to coarser-resolution broader-extent surveys (topographic variability).

Hypothesis 2: Environmental variables will act over a range of specific scales to drive biological patterns even when derived from the same original raster (bathymetry or backscatter). The influence on species distributions of an environmental variable measured at one scale may show a different response curve when examined at another scale.

Hypothesis 3: Species, communities and macroecological properties (e.g. abundance, species richness and biodiversity) will vary over different scales even within the same geographical area. Species are likely to require the finest-scale habitat characterisation followed by communities and macroecological properties.

The datasets employed in this thesis gave a rare opportunity for a deep-sea study to examine the importance of multiple spatial scales using environmental variables ranging in resolution from cm, as extracted from imagery, to 100m, as obtained with shipboard mapping. Although incredibly high details may now be achieved, it is clear that the highest resolutions may require too much time to acquire and sometimes provide an unwieldy quantity of information. Hence as a conclusion, I examine the trade-offs between the developed methods and more traditional faster, but less detailed approaches. Applications and best practice suggestions are also put forth for the incorporation of these methods into management and conservation measures.

Chapter 2:
Megafaunal distribution and biodiversity in a heterogeneous landscape: The iceberg scoured Rockall Bank, NE Atlantic

Katleen Robert[1], Daniel O.B. Jones[2] and Veerle A.I. Huvenne[2]

[1] School of Ocean and Earth Science, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, UK

[2] National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

2.1 Abstract

Species distributions are influenced by spatial structure in environmental factors, but the scales at which these dependencies occur and the effect of habitat patch diversity, connectivity and spatial arrangement have rarely been investigated in deep-sea settings. In this study, spatially-limited photographic transects collected from Rockall Bank, Northeast Atlantic, were combined with sidescan and multibeam sonar maps to model spatial patterns in species distribution and biodiversity. Sediment interpretation maps were created and canonical ordination techniques were used to examine relationships between fine-scale sediment characteristics extracted from the digital stills as well as landscape metrics describing the patch mosaic structure of the surrounding areas. Fine-scale sediment characteristics explained 45.1% and 63.8% of the variation in species composition and biodiversity (H') respectively. This survey effectively captured variation in species distribution resulting from iceberg ploughmarks, occurring at a scale of <50m which would normally go undetected by traditional ship-based studies. Our study suggests that fine-scale environmental information is required to capture the spatial heterogeneity of complex seafloor areas in sufficient detail to model species distributions and biodiversity.

2.2   Introduction

Long-term stability in the deep sea was initially suggested to explain the higher than expected species richness observed (Sanders 1968). However, the importance of spatial heterogeneity (the uneven distribution of environmental variables over space) was soon recognized (Jumars 1976). In an heterogeneous environment, a higher number of niches are available which allows for resource partitioning to occur, reducing competitive exclusion between species (Williams 1964). The hypothesis that increased spatial heterogeneity leads to higher species richness has been examined in terrestrial (Pickett & Cadenasso 1995, Tews et al. 2004) and shallow water environments (Hewitt et al. 2005, Mellin et al. 2012), but for many deep-sea areas spatial heterogeneity has not yet been mapped at sufficient resolution to represent fine-scale biodiversity patterns over large spatial extents (Thrush et al. 2008, Levin & Sibuet 2012).   Yet this information would have significant advantage for the implementation of management measures, where precautionary decisions may have to be made based on limited evidence. In this case, seafloor heterogeneity, which can be rapidly described via acoustics surveys, could be employed as a proxy for biological diversity or conservation priority, reducing the significant time investment associated with biological data collection, identification and manual quantification (Schoening et al. 2012).

Spatial patterns in species distribution can arise from interaction between organisms (e.g. intraspecific: reproduction or recruitment and interspecific: predation or competition) or be induced by spatial structures exhibited by environmental factors (Legendre 1993). The characterization of relationships between environmental factors and species distributions is a first step in accurately predicting species distributions and creating fine-scale habitat maps. Environmental variables vary over different spatial scales to form a mosaic of interspersed habitat patches on the seafloor (Jumars 1976). However, in the deep sea the scales at which these factors influence species distributions have not yet been thoroughly investigated. At broader spatial scales (~100km - 1,000km), deep-sea studies on the relationships between environmental variables and biodiversity showed that water mass and current related factors (e.g. oxygen and temperature), as well as flux of organic material, have the strongest influence on biodiversity (Levin et al. 2001,

Sellanes et al. 2010, Williams et al. 2010). At medium scales (~km - 10km), the presence of large geomorphological features such as submarine canyons, nodule fields or habitat forming biological structures are significant (Henry & Roberts 2007, Sellanes et al. 2010, Vetter et al. 2010). The importance of sediment properties (e.g. hardness, grain size), food resources (e.g. organic matter content) and disturbance rate appears only at finer spatial scales (~m - ~km) (Vetter & Dayton 1999, Gutt & Piepenburg 2003, McClain & Barry 2010).

The spatial relationships between habitat patches and their effects on the distribution and composition of the deep-sea benthic fauna has even more rarely been examined (Wedding et al. 2011). The field of 'landscape ecology', developed in terrestrial environments, has focused on developing metrics to describe the geometry and spatial arrangement of habitat patches and their relationships to ecological processes (Turner & Gardner 1991). Class metrics are used to describe properties of patches from a single substratum type, while landscape metrics consider all patches present within a landscape (McGarigal et al. 2012). In shallow marine environments, a landscape approach has been used to examine effects of habitat fragmentation in seagrass beds (Jackson et al. 2006), the importance of patch size and connectivity in coral reefs (Grober-Dunsmore et al. 2007), spatial patterns in rocky benthic species assemblages (Garrabou et al. 1998) as well as the multi-scale influence of landscape structure on the spatial distribution of fish species (Pittman et al. 2004, Monk et al. 2011). Metrics such has fractal dimension have also been found useful in describing irregular shapes such as spatial patterns within mussel beds (Commito & Rusignuolo 2000) and the morphology of marine branching sessile organisms (Kaandorp 1999). Although application of spatial metrics in deeper marine ecosystems had been limited owing to difficulties associated with underwater mapping, it was successful in explaining fine-scale (<m) benthic species assemblages in the Antarctic (Teixidó et al. 2002, Teixidó et al. 2007). In deeper sites, the landscape is often characterized by sediment type, and not by vegetation or biogenic structures (with the exception of cold-water corals, carbonate mounds and sponge aggregations) (Klitgaard 1995, Howell et al. 2011), resulting in more subtle changes in seafloor structure making it inherently difficult to delineate benthic habitats (Zajac 2008). However, recent studies in automated seabed classification based on sidescan or multibeam sonar backscatter have greatly facilitated creation of high-resolution sediment

interpretation maps (Wilson et al. 2007, Lucieer 2008, Brown et al. 2012). Analysis of these maps from a landscape perspective has the potential to increase the amount of ecologically meaningful information extracted. Since no additional data collection is required, this approach has the potential to reduce the time needed to gather sufficient information to address management issues.

As the anthropogenic footprint of activities such as trawling extends deeper into our oceans, detailed descriptions of seafloor habitats and the species they harbour become increasingly important in order to establish the baseline state of this ecosystem (Levin & Sibuet 2012). As the environmental variables responsible for the spatial structuring of species distributions vary over different scales, choice of resolution and spatial extent will affect the ecological processes which can be examined (Levin 1992). We need to find sampling resolutions which retain enough fine-scale variation to describe species distributions while remaining coarse enough to be economically feasible (Przeslawski et al. 2011).

This study used a hierarchical survey to examine a highly heterogeneous seafloor area characterized by iceberg ploughmarks on Rockall Bank, Northeast Atlantic (200 - 400m depth). A fine-scale analysis of photographic transects was first carried out to identify megafaunal species and map their distribution and biodiversity. A landscape approach was then used to examine whether the inclusion of metrics describing the spatial arrangement of habitat patches could improve the explanatory power of models using environmental variables to describe species composition and distributions. The amount of variation explained by survey techniques of varying resolutions was also examined and the scale of variation in the biological data was used to identify the process potentially responsible for the spatial structure captured by the survey.

2.3   Methods

2.3.1   Study site

Rockall Bank, Northeast Atlantic (Figure 2.1), is a shallower part of the larger Rockall Plateau, a subsided and submerged microcontinent which includes Hatton-Rockall Basin and Hatton Bank to the northwest, and is
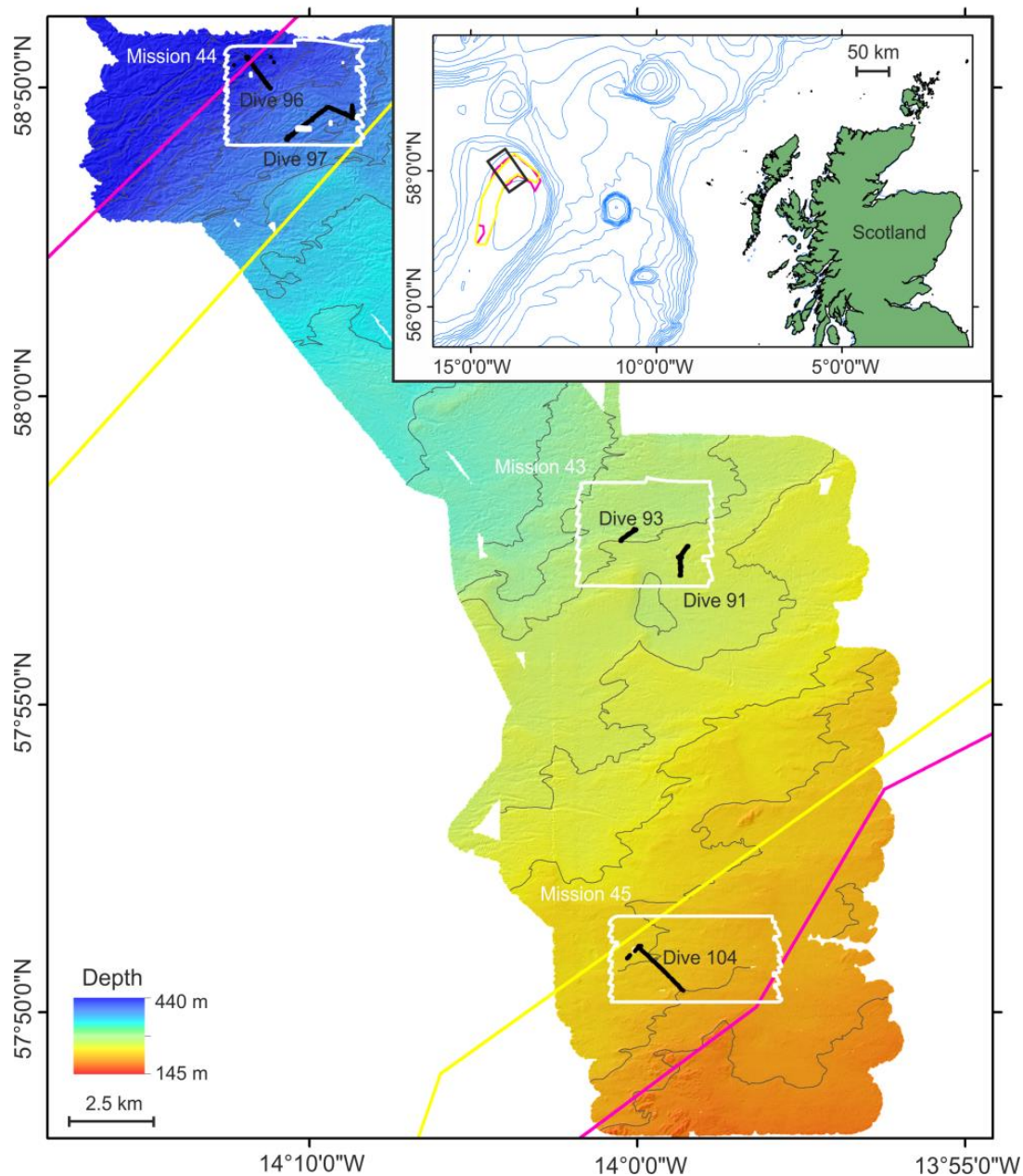
Figure 2.1: Hierarchical survey carried out on Rockall Bank, Northeast Atlantic. Ship-based bathymetry displayed with superimposed outlines of the sidescan sonar data (white) collected during three autonomous underwater vehicle (AUV) missions. The five remotely operated vehicle (ROV) imagery transects are shown in black. Insert shows the position of Rockall Bank in relation to Scotland; the black rectangle represents the location of the survey. The boundaries of the 2007 fisheries closure area and candidate for 'Special Area of Conservation' are illustrated in yellow and pink respectively. Background bathymetry of the Northeast Atlantic from GEBCO (General Bathymetric Chart of the Oceans,(IOC IHO and BODC 2003).

separated from the Scottish continental margin by Rockall Trough (Roberts 1971, 1975). The seabed of the Rockall Bank has a long history of

investigation (Thomson 1874). The shallower areas of the western bank have a heterogeneous seabed, including partly buried rock outcrops, boulder and cobble fields as well as large areas of carbonate sand cover (Roberts 1975). The deeper western and southern flanks (250 to 450m water depth) are incised with deep scours (<8m) from iceberg keels having ploughed the seabed during periods of Quaternary low sea levels (Sacchetti et al. 2012). Over time soft sediment filled the centre of the ploughmarks, while coarse debris remained at the scar edges. The presence of dispersed hard substratum has allowed colonization by the cold-water coral *Lophelia pertusa* (Wilson 1979a, b, Rogers 1999). A high diversity of organisms has been reported associated with *L. pertusa* patches (Jensen & Frederiksen 1992, Wienberg et al. 2008), while there has been little investigation of the fauna of the softer sediments (Wilson & Desmond 1986). In 2007, the North East Atlantic Fisheries Commission established a Fisheries Closure on the north western flank of Rockall Bank. In 2010, a nearly overlapping area was proposed, by the Joint Nature and Conservation Committee (JNCC), an adviser group to the UK government, as a candidate 'Special Area of Conservation' (SAC) under the EU Habitat's Directive, with the aim to protect the cold-water coral communities from extensive deep-water fishing activities (Howell et al. 2009, JNCC 2010a)

2.3.2   Survey Design and Data Collection

2.3.2.1   Map Creation

A nested seafloor survey of the north-western flank of Rockall Bank (Figure 2.1) (200 - 400m depth), was carried out during the *JC*-060 cruise in May - June 2011 as part of the 'Marine Environmental Mapping Programme' (MAREMAP: http://www.maremap.ac.uk/index.html) and the 'COmplex Deep-sea Environments: Mapping habitat heterogeneity As Proxy for biodiversity' project (CODEMAP: http://www.codemap.eu/). Three missions (M43, M44 and M45) by the autonomous underwater vehicle (AUV) Autosub6000 were carried out, mapping three distinct areas of seafloor (12.0, 12.0 and 13.0km^2 respectively) using an EdgeTech high frequency (410kHz) high-resolution (resulting pixel size of 0.5x0.5m) sidescan sonar. These were located in areas within (M43) and immediately outside (M44 and M45) the Fisheries Closure, but still within the candidate SAC (Figure 2.1). Ship-board bathymetry of the

surrounding area (380km²) was also collected using a Kongsberg EM710 multibeam echosounder (128 beams; resulting pixel size of 20x20m).

Thematic maps representing seafloor substratum composition, based on an unsupervised classification of the sidescan sonar backscatter, were produced. Unsupervised classification attempts to identify structure within the data and segment it into units without prior recourse to *in situ* reference points (Brown et al. 2012). This classification used mean backscatter, average grey level difference, and variance within a 9x9 pixel moving window to assign each pixel to one of six classes: soft, mixed or hard substratum, coral stand or rubble, and exposed bedrock. Mean backscatter represented substratum hardness, while variance at fine spatial scales was indicative of more complex substratum structures such as coral stands, where strong contrasts occur between the high backscatter of the corals in combination with the low backscatter caused by their shadows (Huvenne et al. 2002). Sediment patches of less than 12 pixels were filtered out; pixels being assigned to the sediment class represented by the majority of neighbouring pixels.

2.3.2.2   **Biological Imagery**

Within the areas surveyed by the AUV, five remotely operated vehicle (ROV) photographic transects (Stations 91, 93, 96, 97 and 104) were conducted using a SAAB SeaEye Lynx at a speed of ~0.1m/s. Digital stills (Figure 2.2) were taken every ~40sec using an oblique mounted downward-looking Kongsberg OE14-208 camera (focal length: 7.188mm and maximum aperture: *f*2) equipped with parallel lasers for scale (10cm separation), and used for analysis of megabenthic invertebrates. Using the image processing software Image J (freely available online: http://rsbweb.nih.gov/ij/), all images from a transect were imported using the 'Image Sequence' function to form a 'stack' (multiple images displayed consecutively in a single window); each organism was identified, marked and its pixel position recorded to avoid risks of double counting. Images of suboptimal quality were removed as well as those collected when ROV altitude varied beyond the 1-2.25m range. Each image was then georeferenced using the ROV's ultra-short baseline (USBL, accuracy 1% of depth) navigation system which also recorded its depth and altitude. The ROV position was estimated based on a moving average of the navigation with a 4-reading subset, corresponding to 1sec. Average depth,

transect lengths and number of images collected per transect are reported in Table 2.5. The ROV was also equipped with a SeaKing CTD which showed average bottom temperatures of $9.26\,^{\circ}$C (SD: $0.10\,^{\circ}$C) and salinities of 34.8‰ (SD: 0.025‰). Images located in proximity to the the sidescan sonar Nadir line were removed due to the associated noise.

A) B)
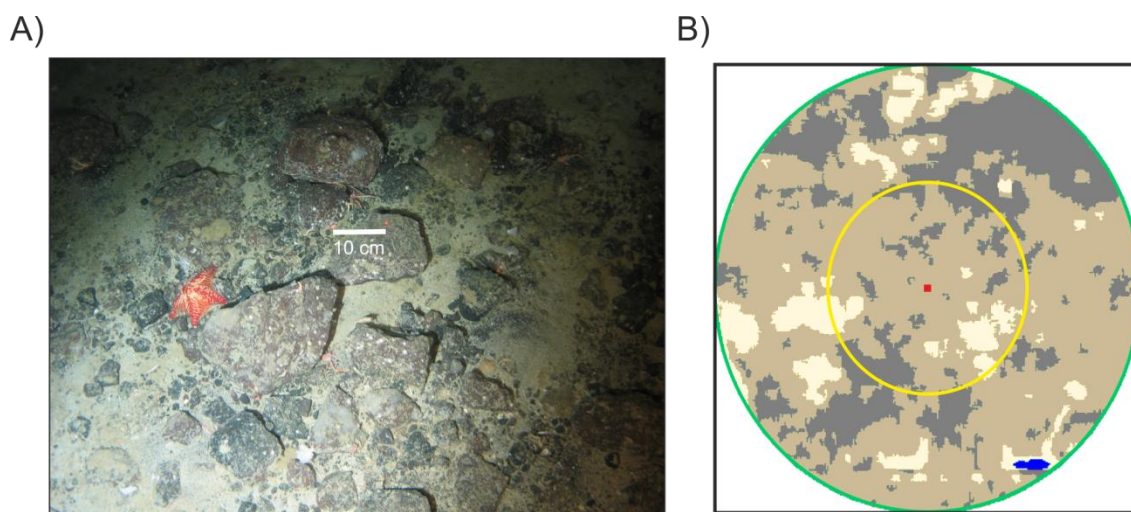
Figure 2.2: Benthic imagery and associated sediment interpretation map. A) Image acquired via remotely operated vehicle illustrating megabenthic organisms and a mixed substratum. Laser beams are separated by 10cm. B) Associated circular areas (green and yellow; 75m and 30 m radii) as represented by the sediment interpretation map. Soft, mixed and hard substratum types are represented in grey, tan and cream respectively, while coral stand are shown in blue. The red square illustrates the approximate area covered by the image.

When species level identification would have required sample collection, visually distinctive megafaunal taxa were identified to the lowest possible taxonomic unit and assigned to morphospecies. The use of morphospecies is somewhat problematic as cryptic species will be considered under a single grouping, while other groups showing greater morphological variation (e.g. Sponges) may be split into more groups, but their use is common in marine studies where imagery was used as the main sampling tool (Soltwedel et al. 2009, Schlacher et al. 2010, Compton et al. 2013). Although the use of higher taxonomic levels is another option to estimate biodiversity (Roy et al. 1996), these tend to be more useful when investigating broader-scale patterns and in the case of our dataset, even family level identification of sponges is

problematic without specimen collection. As such the diversity measures presented are likely underestimated, but represent the closest estimates that can be achieved. Consultation with the image catalogue compiled by Howell and Davies (2010) for morphospecies encountered in the surrounding area was carried out and a set of voucher images was assembled. This set of images was deposited in the publicly available SERPENT media archive (http://archive.serpentproject.com/). In the case of encrusting species, individual colonies (distinct clumps separated by surrounding substratum) were counted. For community analysis, only morphospecies for which at least 10 individuals were observed and a presence recorded in at least 10 images were retained, but every record was retained for biodiversity calculations (Shannon Index of Diversity, H'). To examine the effect that cryptic species may have on estimates of biodiversity, a percentage (5% and 20%) of the number of taxa listed in OBIS (Ocean Biogeographic Information System, http://www.iobis.org/) for the North Atlantic ocean was used as a potential number of present, but undistinguishable cryptic species. When two or more specimens of a morphospecies of Ascidians, Bryozoans, Cerianthids, Cormatulids, Sponges, Ophiuroids, Sabellid worms or Sipuncula worms were observed within an image, each individual was randomly assigned, with replacement, to a potential cryptic species and biodiversity estimates recalculated. This analysis was conducted in order to determine whether the uncertainty created by the use of morphospecies was high enough to effect the conclusions of this study. However, as the trends observed did not change, we argue that the use of morphospecies is appropriate for this study (results presented in supporting information, 2.8.1)

2.3.2.3   Imagery-Derived Environmental Variables

Composition of the substratum was visually assessed in each image based on grain size classes: soft sediment, gravel, pebbles (4-64mm), cobbles (64-256mm) and boulders (>256mm) (Wentworth 1922). Pebbles of 4mm could be measured in the imagery, but only a visual distinction in texture could be used to separate soft sediments from gravel. The overall composition was first assigned to one of 7 seabed facies: sand, sand and pebbles or gravel, sand and cobbles, cobble dominated, coral stand, coral rubble and exposed bedrock (example images are provided in supporting information, 2.8.2). To

obtain a quantitative description of substratum composition, percentage covers were obtained by importing images in the freely available statistical software R (R Development Core Team 2014) and drawing 100 randomly located points for each image. Using a custom-made R code with an interactive prompt, the substratum (grain size class, exposed bedrock, coral rubble or coral stand) at each location was recorded.

2.3.2.4   **Sonar-Derived Environmental Variables**

The spatial structure (e.g. size, shape, composition, spatial arrangement and diversity) of the seafloor habitat patches represented in the sediment interpretation of the sidescan sonar maps (complete extent figures available as supporting information, 2.8.3) was described using the class and landscape metrics listed in
Table 2.1. Metrics were grouped into five general groups 'Area and Edge', 'Shape', 'Core Area', 'Aggregation' and 'Diversity' (Peng et al. 2010, McGarigal et al. 2012). The first group is related to size of the patches and amount of edge, the second group is used to characterize the geometry of the different patches, while the third grouping examines patch sizes when the edge cells (only the first one in this study) are removed. Spatial arrangement of patches to each other is described by the fourth group, while diversity measures, only available for landscape analysis, form the last group. Refer to McGarigal et al. (2012) for a thorough description of each measure and equations. In the present study, circular areas of seafloor (1, 5, 10, 20, 30, 40, 50, 75 and 100m in radius) were delimited around each image (Figure 2.2). Within those circular areas, class and landscape metrics were computed using the 'SDMTools' (Species Distribution Modelling Tools) package in R. Compared with the seafloor visible in the fine-scale imagery, these class and landscape metrics provided a description of the broader seabed surrounding each image. The use of increasingly larger area sizes was carried out in order to determine which one would be most appropriate for the calculation of class and landscape metrics as explanatory variables for species composition and biodiversity. Using the statistical techniques described below (redundancy analysis or linear regression with forward selection) metrics were selected for each circle size and the amount of variation explained calculated. Variation partitioning (explained below) was used to select which two circle sizes

provided the most appropriate metrics to describe the biological variation
observed.

Table 2.1: List of class and landscape level metrics considered in the analysis. Class metrics were calculated for each of the six substratum classes present in the sediment interpretation maps: soft, mixed or hard substratum, coral stand or rubble, and exposed bedrock. For formulas and descriptions see McGarigal (2012).

Class Metrics		Landscape Metrics	
Area and Edge Metrics	*Core Area Metrics*	*Area and Edge Metrics*	*Aggregation Metrics*
Patch Number	Total Core Area	Patch Number	Proportion of Like Adjacencies
Total Area	Mean Core Area	Patch Density	Aggregation Index
Patch Density	Smallest Core Area	Edge Length	Landscape Division Index
Edge Length	Largest Core Area	Mean Patch Area	Splitting Index
Edge Density	*Aggregation Metrics*	Smallest Patch Area	Effective Mesh Size
Mean Patch Area	Proportion of Like Adjacencies	Largest Patch Area	Patch Cohesion Index
Smallest Patch Area	Aggregation Index	*Shape Metrics*	*Diversity Metrics*
Largest Patch Area	Landscape Division Index	Landscape Shape Index	Patch Richness
Shape Metrics	Splitting Index	Largest Patch Index	
Landscape Shape Index	Effective Mesh Size	Mean Shape Index	
Largest Patch Index	Patch Cohesion Index	Minimum Shape Index	
Perimeter Area Fractal Dimension		Maximum Shape Index	
Mean Perimeter Area Ratio		*Core Area Metrics*	
Minimum Perimeter Area Ratio		Total Core Area	
Maximum Perimeter Area Ratio		Mean Core Area	
Mean Shape Index		Smallest Core Area	
Minimum Shape Index		Largest Core Area	
Maximum Shape Index		Mean Core Area Index	

For each image, the statistical mean and variance in backscatter as obtained from the original sidescan sonar maps were calculated based on all pixels present within the two circular areas. Area weighted averages for slope, curvature, aspect, surface-area ratio and bathymetric position index (BPI) were calculated for the 4m surrounding the position of each image. The layers were derived from a 20x20m resolution base surface of the multibeam data (spatial reference: World Geodetic System '84 Universal Transverse Mercator Zone 28N). As BPI varies depending on neighbourhood size, two layers were created: coarse and fine, based on 10 pixels and 2 pixels neighbourhood radii (Wilson et al. 2007). Layers were generated in ArcGIS using the 'Spatial Analyst Extension' as well as the 'Land Facet Corridor Tools' and the 'DEM Surface Tools' developed by Jenness Enterprises (Jenness 2012a, b).

2.3.3   Statistical Analysis

Owing to the number of statistical methods employed only a general description is provided below and additional details are provided in supporting information, 2.8.4. Analyses were carried out in the statistical software R using the libraries 'vegan', 'labdsv' and 'gstat'. A methodological flowchart is also presented (Figure 2.3) with each major step represented by numbers. An extensive review of statistical techniques for spatial analysis of community data, including the ones employed in this study, is available in Dray et al. (2012).

2.3.3.1   Environmental Variable Selection

To determine which explanatory variables (e.g. sediment percentage cover, class and landscape metrics, and multibeam sonar-derived layers) could best explain morphospecies distribution, (**1**) redundancy analysis (RDA) was used. RDA is a type of constrained ordination which allows the regression concept to be applied to a multivariate response variable (such as a species matrix) (Legendre & Legendre 1998). The morphospecies abundance matrix was log(x+1) transformed before the analysis to reduce the influence of abundant species (Clarke & Warwick 2001). (**2**) Forward selection was carried out to obtain the most parsimonious model and variance inflation factors (VIF) were used to exclude additional explanatory variables that showed strong collinearity with others present within the model (Borcard et al. 2011). To compare the information obtained from sampling tools of differing resolution (ROV images: <m, versus acoustic maps: 1m-20m), this step was carried out once using all explanatory variables and once using only sonar-derived environmental variables. As only lower resolution survey methods are generally available over large extents, (**3**) variation partitioning (Peres-Neto et al. 2006) was used to examine the amount of variation that was no longer captured when imagery-derived environmental variables (e.g. sediment percentage cover) were no longer available. The amount of variation in species composition and biodiversity (H') explained by the sediment class present at the location of the image and the values of the multibeam-derived layers was also calculated. This determined whether the inclusion of landscape metrics increased explanatory power when compared to the use of sediment interpretation maps as abiotic proxies for species assemblages. These steps

were also carried out using the Shannon index of diversity (H′) (Shannon 1948) as a response variable, but linear regression was used instead of RDA as the response variable was now univariate.
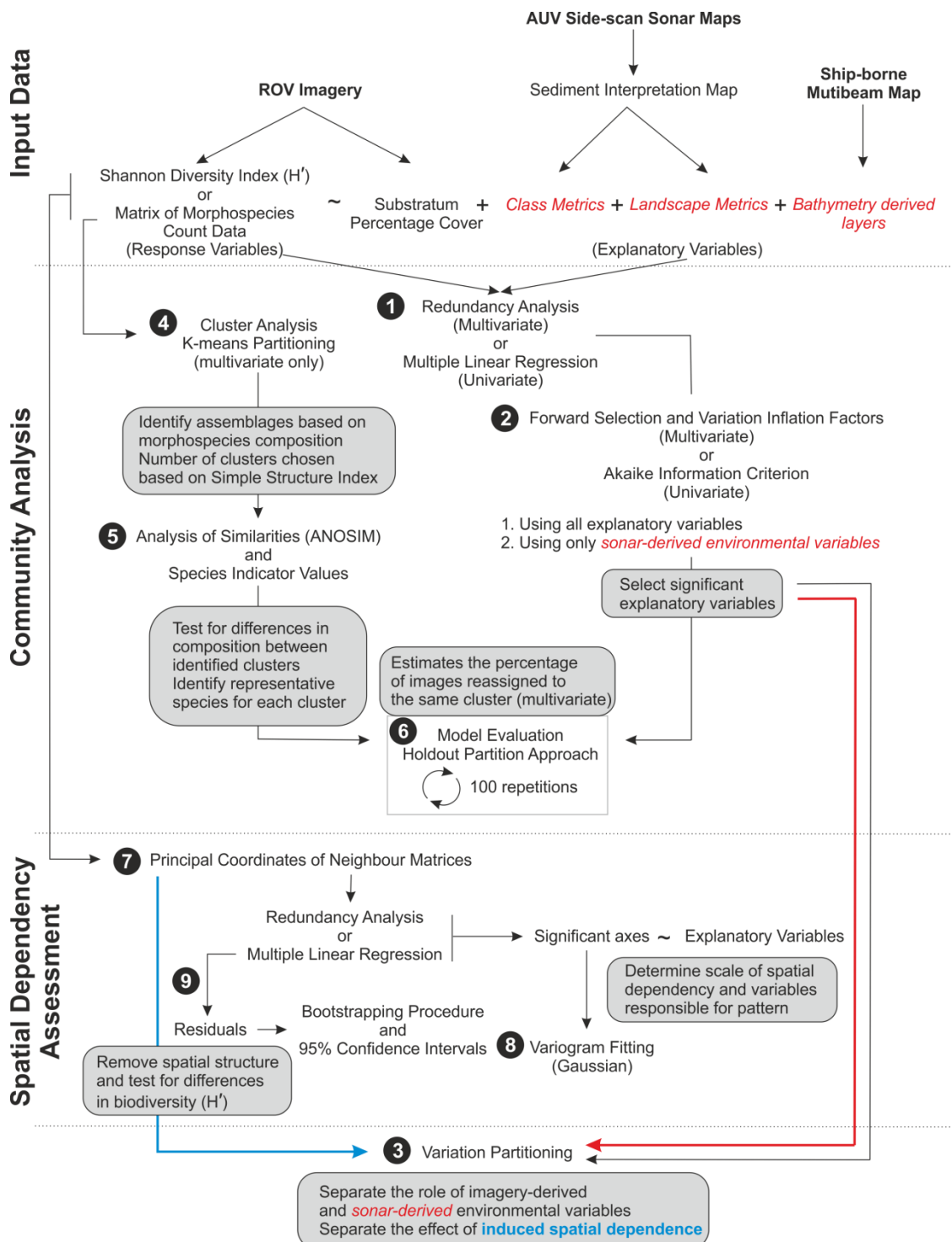
Figure 2.3: Methodological flow chart representing the statistical analysis. Rounded rectangles contain short descriptions of the aims of specific steps.

2.3.3.2 Species Assemblages

To identify clusters of images of similar morphospecies composition, (**4**) K-means partitioning was employed. This method aims at separating observations into a predefined number of clusters by minimizing the distance between individual samples and the center of their assigned cluster (Hartigan & Wong 1979). The optimal number of cluster was selected using the 'simple structure index' (Dimitriadou et al. 2002). (**5**) Analysis of Similarities (ANOSIM) was used to assess significant differences in morphospecies composition between clusters and representative species for each cluster were identified based on 'species indicator values' (Dufrêne & Legendre 1997).

2.3.3.3 Model Evaluation

As an additional independent dataset was not available, (**6**) model evaluation was carried out using a 'holdout partition' approach. A subset of the data (300 randomly selected images) was removed from the original dataset and the RDA parameters were recomputed. The removed data points were then reclassified and the percentage of data points reassigned the same cluster was calculated. This process was repeated 100 times to estimate variability in the results obtained. Marine benthic studies employing similar methodological approaches include Hewitt et al. (2004), Teixidó et al. (2007) and Verfaillie et al. (2009)

2.3.3.4 Spatial Dependency Assessment

To determine how much of the spatial structure present in the species dataset can be explained by spatial structuring of environmental variables (induced spatial dependence), (**7**) Principal Coordinates of Neighbour Matrices (PCNM) were used (Borcard & Legendre 2002, Dray et al. 2006). This approach yielded synthetic representations of potential spatial structures based on distances between sampling sites. These synthetic representations were then compared to the spatial structure present in the species dataset and those found to be accurate representations were included in the variation partitioning step (**3**) to assess how well the environmental variables described the biological spatial structures modelled. (**8**) Gaussian variogram creation of the selected spatial representations was used to determine the scale of variation captured (the distance required for two points to be considered

independent). As PCNM represent the spatial structure present within the species dataset, they can also be used to alleviate issues associated with autocorrelation. This property was used to compare differences in biodiversity at a broader scale, between transects and substratum facies.

2.3.3.5 **Broader-Scale Spatial Patterns**

In order to assess broader-scale patterns, class and landscape metrics were calculated within a 200m buffer around each transect line. Linear regressions of diversity (H′), organism abundance and substratum percentage cover observed in each image were carried out against depth. To examine differences in biodiversity (between transects and between seabed facies), we first accounted for spatial autocorrelation by filtering out the spatial structure previously modelled (**9**). The biodiversity indices (H′) were first regressed against the significant PCNMs and the residuals were used as response variables in the analysis. Because transects length differed, data resampling with replacement was carried out to obtain 75 images from each transect. The 95% confidence intervals around the mean residual biodiversity for each transect were built using a bootstrapping procedure with 999 repetitions. For the seabed facies, the same bootstrapping procedure was applied, but sample size was standardized to 30 images.

2.4   Results

In the 1,222 images analysed, a total of 7,267 individual organisms were observed from 81 morphospecies (list provided as supporting information, 2.8.5). Many of these were rare and only 35 morphospecies were seen in more than 10 images (Figure 2.4). The squat lobster *Munida sarsi* made up the largest percentage, 33.0%, of the organisms observed and was the dominant species in all transects with the exception of transect 96, which was dominated by the holothurian *Parastichopus tremulus* (40.3% of fauna). However, across all transects, *P. tremulus* only represented 2.6% of the total observations. Only the bryozoan *Reteporella* sp. (13.1%) and colonies of yellow (9.8%) and white (6.2%) encrusting sponges composed more than 5% of the total observations.
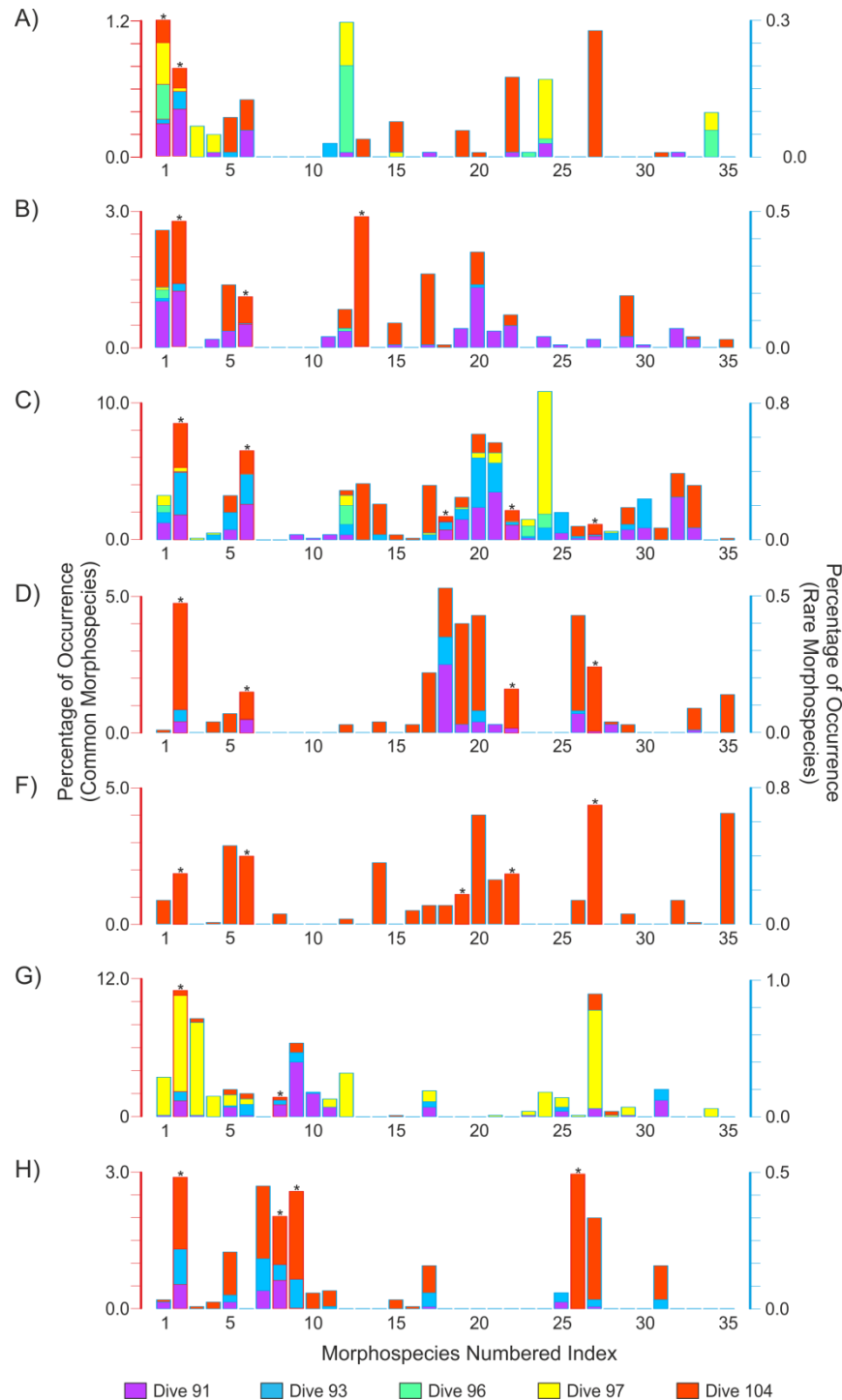
Figure 2.4: Percent occurrence for each of the 35 morphospecies retained for the community analysis (total number of individuals counted for a given species divided by the total number of individual organisms observed multiplied by 100). Values are colour coded by transect and separated by seabed facies: A) sand, B) sand and pebbles or gravel, C) sand and cobbles, D) cobble dominated, E) bedrock, F) coral rubble and G) coral stands. Two scales were used in order to improve visualization: bars with red outlines and marked with asterisks show the more common morphospecies, while bars with a blue outlines are associated with the scale on the right and represent rarer morphospecies. Morphospecies numbers refer to the list presented in Table 2.4.

The class and landscape metrics calculated at varying scales showed that the most variation was captured using a radius of 40m for both morphospecies composition and biodiversity. However, when two circle sizes were considered, the highest variation explained was achieved using metrics derived from areas with radii of 30 and 75m for morphospecies composition, and 20 and 75m for biodiversity (results included as supporting information, 2.8.6). For simplicity, sonar-derived metrics calculated at 30 and 75m were used for the analysis.

Following forward selection, a large number of environmental variables (substratum percentage cover, class and landscape metrics as well as backscatter and bathymetric variables) significantly explained morphospecies composition and biodiversity (H') (Table 2.2 and 2.3). The most parsimonious models including all environmental variables explained 45.1% (P-value: 0.001) and 63.8% (P-value: <0.001) of the variation for morphospecies composition and biodiversity respectively. When models relied solely on sonar-derived environmental characteristics lower percentages of variation were explained: 24.7% (P-value: 0.001) and 40.6% (P-value: <0.001) for morphospecies composition and biodiversity (H') respectively. However, these percentages were much higher than what was obtained when only the sediment interpretation class and multibeam-derived information were employed: with variation explained of 13.0% (P-value: 0.001) for morphospecies composition and 22.9% (P-value: <0.001) for biodiversity (H').

The first 14 axes of the redundancy analysis were significant, but only the first 2 axes are illustrated in the ordination graphs. Substratum percentage cover, as obtained from the imagery, best explained variation in the morphospecies count data (Figure 2.5A). The first canonical axis illustrates a gradient in images dominated by soft to hard substrata, while the second axis represents the presence of coral rubble and stands. The vectors representing species scores (Figure 2.5B) separate into three subgroups: upper right quadrant characterized by squat lobsters (*M. sarsi*), a species of Actiniaria, yellow encrusting sponges and Sabellid worms, lower right quadrant showing predominance of other small encrusting sponge colonies and bryozoan species as well as the asteroid *Porania pulvillus*. The lower left quadrant was represented by the holothurian *P. tremulus*. K-means

partitioning showed a similar trend by identifying an optimum of three clusters. However, the use of four clusters allowed the separation of species occurring on both corals and hard substratum from those more closely associated with live coral stands. As the four clusters differed significantly in their morphospecies composition based on ANOSIM analysis (P-value: 0.001, R-statistic: 0.85), they were retained.

Table 2.2: Selected environmental variables for the most parsimonious models for the morphospecies count matrix (with and without including fine-scale imagery-derived information).

	Sonar	Imagery and Sonar
Explanatory Variables	β Number of Patches (75m)	ε Percentage Cover Cobbles
	α Proportion of Like Adjacencies Sand (30m)	ε Percentage Cover Rubble
	α Total Area Coral (30m)	ε Percentage Cover LiveCoral
	γ Mean Backscatter (30m)	ε (Percentage Cover Cobbles)^2
	α Mean Shape Index Coral (75m)	ε Rock (Factor)
	α Effective Mesh Size Bedrock (30m)	ε Rock - Sand (Factor)
	δ Aspect	ε Sand - Cobbles (Factor)
	β Mean Shape Index (75m)	ε Cobbles - Sand (Factor)
	α Minimum Patch Core Area Coral (30m)	α Proportion of Like Adjacencies Sand (30m)
	α Patch Density Bedrock (30m)	α Mean Shape Index Coral (75m)
	α Patch Density Rubble (30m)	ε Percentage Cover Boulder
	α Landscape Shape Index Mixed (30m)	α Minimum Patch Core Area Coral (30m)
	β Largest Patch Index (30m)	ε (Percentage Cover Rubble)^2
	α Proportio of Like Adjacencies Sand (75m)	α Propprtion of Like Adjacencies Sand (75m)
	α Largest Patch Index Coral (75m)	α Effective Mesh Size Bedrock (30m)
	α Mean Shape Index Bedrock (30m)	α Total Area Coral (30m)
	α Minimum Patch Area Bedrock (75m)	α Landscape Shape Index Sand (30m)
	β Mean Core Area Index (75m)	ε Percentage Cover Gravel
	α Mean Shape Index Mixed (75m)	α Patch Density Bedrock (75m)
	α Mean Patch Core Area Hard (75m)	α Mean Shape Index Bedrock (30m)
	α Maximum Shape Index Sand (75m)	α Landscape Shape Index Hard (30m)
	γ Mean Backscatter (75m)	α Lanscape Division Index Mixed (75m)
	β Maximum Shape Index (30m)	α Patch Density Rubble (30m)
	α Landscape Shape Index Sand (30m)	ε Pebbles (Factor)
	γ Variance Backscatter (75m)	δ Aspect
	α Minimum Shape Index Sand (75m)	α Maximum Shape Index Sand (75m)
	α Mean Patch Area Sand (75m)	α Mean Patch Core Area Mixed (30m)
	α Mean Patch Core Area Mixed (30m)	α Largest Patch Index Coral (75m)
	α Mean Shape Index Hard (30m)	β Number of Patches (75m)
	α Minimum Patch Core Area Rubble (75m)	α Landscape Shape Index Mixed (30m)
		α Mean Shape Index Mixed (75m)
		β Maximum Core Area Index (75m)
		α Patch Cohesion Index Hard (75m)
		β Patch Cohesion Index (30m)
Adjusted R²	24.7%	45.1%
F Value	12.1, df = 36, 1185	26.1, df = 40, 1181
P Value	0.001	0.001

αClass and βLandscape metrics, γSidescan sonar, δMultibeam sonar and εROV imagery-derived variables

Table 2.3: Selected environmental variables for the most parsimonious models for the Shannon diversity index (H$^{'}$) (with and without including fine-scale imagery-derived information).

	Sonar		Imagery and Sonar	
Explanatory Variables	Shannon Diversity Index	**Coefficient**	Shannon Diversity Index	**Coefficient**
	α Aggregation Index Coral (75m)	0.003 ***	ε Percentage Cover Sand	-0.454 ***
	α Proportion of Like Adjacencies Sand (30m)	-0.456 ***	ε Sand - Cobbles (Factor)	0.668 ***
	α Proportion of Like Adjacencies Coral (30m)	0.188 *	α Aggregation Index Coral (75m)	0.003 ***
	γ Mean Backscatter (30)	<0.001 ***	ε Percentage Cover Gravel	-0.455 ***
	α Patch Cohesion Index Hard (75m)	-0.188	ε Rock - Sand (Factor)	0.927 ***
	δ Aspect	NA	ε Rock (Factor)	0.950 ***
	α Mean Shape Index Mixed (75m)	-0.193 **	ε (Percentage Cover Rubble)^2	-1.712 ***
	α Effective Mesh Size Bedrock (30m)	0.000 **	ε Cobbles - Sand (Factor)	0.634 ***
	α Effective Mesh Size Coral (75m)	0.039 **	α Mean Patch Area Sand (75m)	<0.001
	β Patch Cohesion Index (75m)	0.271 *	β Patch Cohesion Index (75m)	0.237 ***
	α Mean Perimeter Area Ratio Hard (75m)	-0.076 *	β Patch Richness (75m)	-0.102 ***
	β Number of Patches (30m)	0.012 ***	α Proportion of Like.adjacencies Sand (30m)	-0.263 ***
	α Minimum Patch Core Area Rubble (75m)	0.003 *	ε Percentage Cover Rubble	0.011
	α Landscape Shape Index Sand (30m)	-0.044 *	α Minimum Patch Area Coral (75m)	0.001 *
	α Mean Patch Area Sand (75m)	<0.001 **	ε Pebbles (Factor)	0.169 ***
	γ Mean Backscatter (75m)	-0.001 ***	ε Percentage Cover Coral	0.300 **
	α Minimum Patch Area Coral (75m)	0.002 *	ε Rubble (Factor)	0.121 **
	α Proportion of like Adjacencies Sand (75m)	-0.433 *	α Maximum Perimeter Area Ratio Sand (75m)	0.014 *
	α Minimum Patch Core Area Mixed (30m)	<0.001 *	α Mean Perimeter Area Ratio Hard (30m)	0.026 *
	β Mean Core Area Index (30m)	0.918 **	β Mean Shape Index (30m)	0.091
	α Number of Patches Sand (75m)	0.003	α Minimum Patch Core Area Rubble (75m)	0.002 .
	α Mean Patch Area Bedrock (75m)	<0.001	α Proportion of Like Adjacencies Coral (30m)	0.088
	β Division Index (75m)	-0.191	δ Slope	0.004
	δ Bathymetric position Index (Coarse)	0.012 .		
	α Maximum Perimeter Area Ratio Sand (30m)	0.012 *		
	α Patch Cohesion Index Mixed (30m)	-0.136		
	α Mean Perimeter Area Ratio Hard (30m)	0.036 *		
	γ Variance Backscatter (75m)	<0.001		
Adjusted R^2	40.6%		63.8%	
F Value	25.6, df = 34, 1187		94.54, df = 23, 1198	
P Value	<0.001		<0.001	

^{α}Class and ^{β}Landscape metrics, ^{γ}Sidescan sonar, ^{δ}Multibeam sonar and ^{ε}ROV imagery-derived variables

Signifiance codes: <0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.'

From 3 to 15 morphospecies were representative for each cluster, as determined by indicator species analysis (Table 2.4). However, no morphospecies represented significantly the cluster associated with soft sediment habitats, but this is likely a result of the low densities of organisms observed in this habitat. There was a clear separation in ordination space of images assigned to each K-means cluster (Figure 2.6). At the transect level, morphospecies associated with soft sediments only represented 5.3% of total organisms observed, while coral associated morphospecies represented 18.9%. The remainder of the organisms recorded represented hard bottom (40.1%) or more generalist morphospecies (e.g. associated with hard substratum, coral

stand and rubble) (38.3%). High percentages of hard bottom-associated morphospecies were recorded for transects 91 (49.0%) and 104 (46.9%).
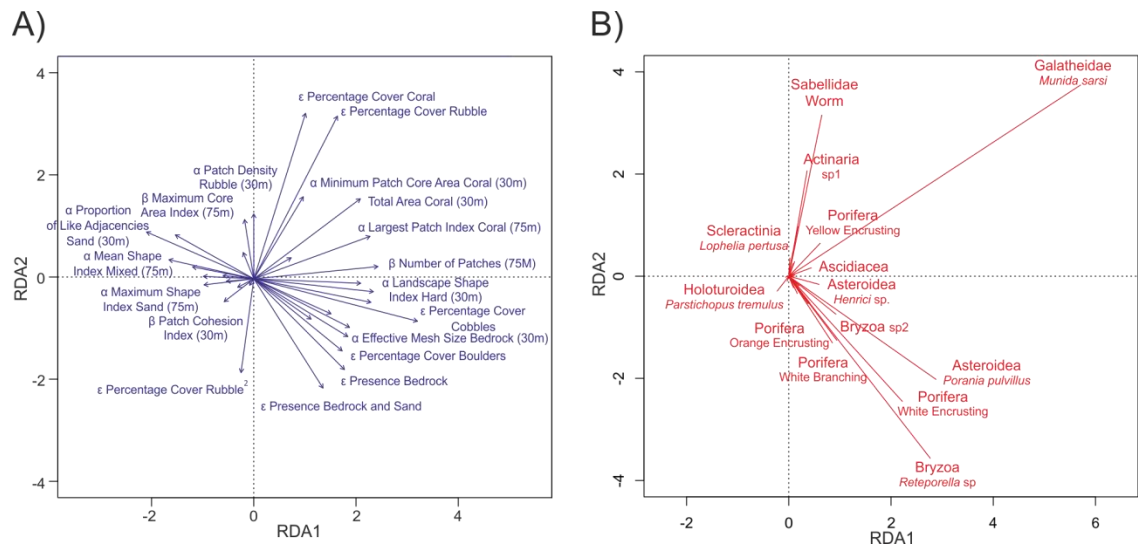
Figure 2.5: Results of the redundancy analysis using all significant environmental predictors . For clarity, the triplot is illustrated in three separate graphs (5 A, B and 6 A) and only environmental variables and morphospecies with the strongest effects are labelled. A) Environmental variables; $^{\alpha}$class metrics, $^{\beta}$landscape metrics, and derived from $^{\gamma}$side scan sonar, $^{\delta}$multibeam sonar and $^{\epsilon}$ROV imagery. B) Morphospecies.
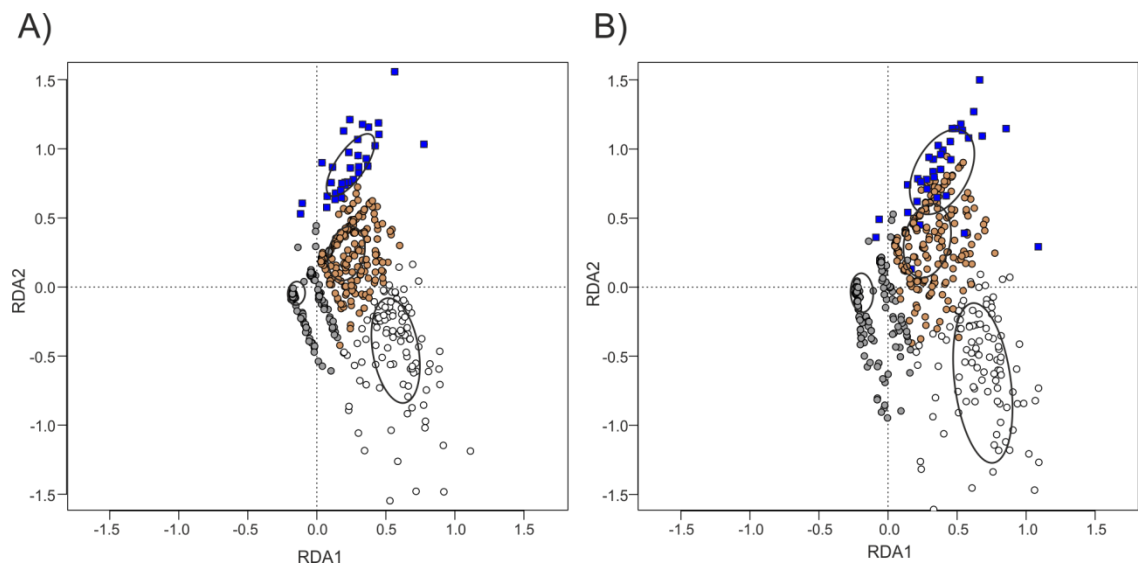
Figure 2.6: Position of each image in ordination space following redundancy analysis (plotted as 'weighted averages of species scores' for the first two canonical axes) using A) both imagery and sonar-derived environmental variables and B) only sonar-derived. Sites are colour coded based on their association to K-means partitioning clusters; soft (grey), mixed (brown) and hard (white) substratum and coral stands (blue rectangles). Black ellipses indicate the standard deviation surrounding the centroid of each cluster.

Organisms representing morphospecies associated with corals ranged between 16.7-18.9% of the observations for transects 91, 93 and 104, but fell below 5% for transects 96 and 97. Morphospecies associated with soft sediments represented 54.5% and 14.2% of total observations for transect 96 and 97 respectively, but less than 5% in the remaining transects. Morphospecies associated with the use of multiple sediment types (hard substratum, coral stand and rubble) ranged between 30.4-73.3% of total observations. Characteristic substrata and morphospecies for each of the observed assemblages are shown in Figure 2.7.

Table 2.4: Characteristic species for each species assemblages . Total counts based on 1,222 images, K-means partitioning cluster association as well as indicator values and significance for the morphospecies observed. Cluster 1: soft sediments, Cluster 2: hard substratum, Cluster 3: coral associated and Cluster 4: occurring on hard substratum, coral stand and rubble. Bold numbers indicate morphospecies presented in Figure 2.4.

Morphospecies	Count	Cluster	Indicator Value	Morphospecies	Count	Cluster	Indicator Value
1 *Parastichopus tremulus*	185	1	0.06	4 *Asterias rubens*	28	3	0.02
24 *Caryophyllia* sp. 1	92	1	0.03	5 *Henricia* spp.	117	3	0.24 ***
34 *Echinus* sp. 3 (possibly *E. elegans*)	11	1	0.01	7 *Lophelia pertusa*	39	3	0.53 ***
6 *Reteporella* sp.	942	2	0.72 ***	8 Sabellidae	272	3	0.93 ***
14 Brown Cup Sponge	942	2	0.72 ***	9 Actinaria sp. 1	227	3	0.72 ***
16 *Echinus* sp. 2	10	2	0.04 **	10 *Echinus* sp. 1 (possibly *E. acutus*)	19	3	0.24 ***
18 Cyclostomatida sp. 1	175	2	0.37 ***	11 Portunidea sp. 1	24	3	0.08 ***
19 Orange Encrusting Sponge	138	2	0.29 ***	15 *Hippasteria* sp.1	19	3	0.06 **
20 White Branching Sponge	152	2	0.43 ***	17 Translucent Tunicate	96	3	0.09 **
21 Brown Lamellated Sponge	69	2	0.16 ***	25 Yellow Columnar Sponge	28	3	0.10 ***
22 White Encrusting Sponge	445	2	0.69 ***	26 Ophiuroidea	263	3	0.16 ***
27 Yellow Encrusting Sponge	705	2	0.54 ***	31 Hydroidae	35	3	0.21 ***
28 *Porania pulvillus*	10	2	0.01	2 *Munida sarsi*	2373	4	0.32 ***
29 Blue Encrusting Sponge	38	2	0.04 *	3 *Cidaris cidaris*	59	4	0.05 .
30 Orange Branching Sponge	19	2	0.05 **	12 Shrimp sp. 1	80	4	0.03
32 Tan Columnar Sponge	48	2	0.10 ***	13 Orange Worm sp. 1	233	4	0.04
33 Yellow Spherical Sponge	36	2	0.12 ***	23 Brown Anemone	13	4	0.01
35 Red Encrusting Sponge	60	2	0.08 ***				

Signifiance codes: <0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.'

In the case of the Shannon biodiversity index, percentage cover of sand and gravel had the strongest association (P-value: <0.001) to lower biodiversity indices (Table 2.3). Sites classified as containing cobbles, exposed bedrock and pebbles had significantly higher biodiversity (P-value: <0.001).   Number of patches (P-value: <0.001) and patch cohesion, an index indicative of greater connectivity (full model: P-value: <0.001, sonar only model: P-value: <0.05), also significantly increased biodiversity, while metrics indicative of large sandy
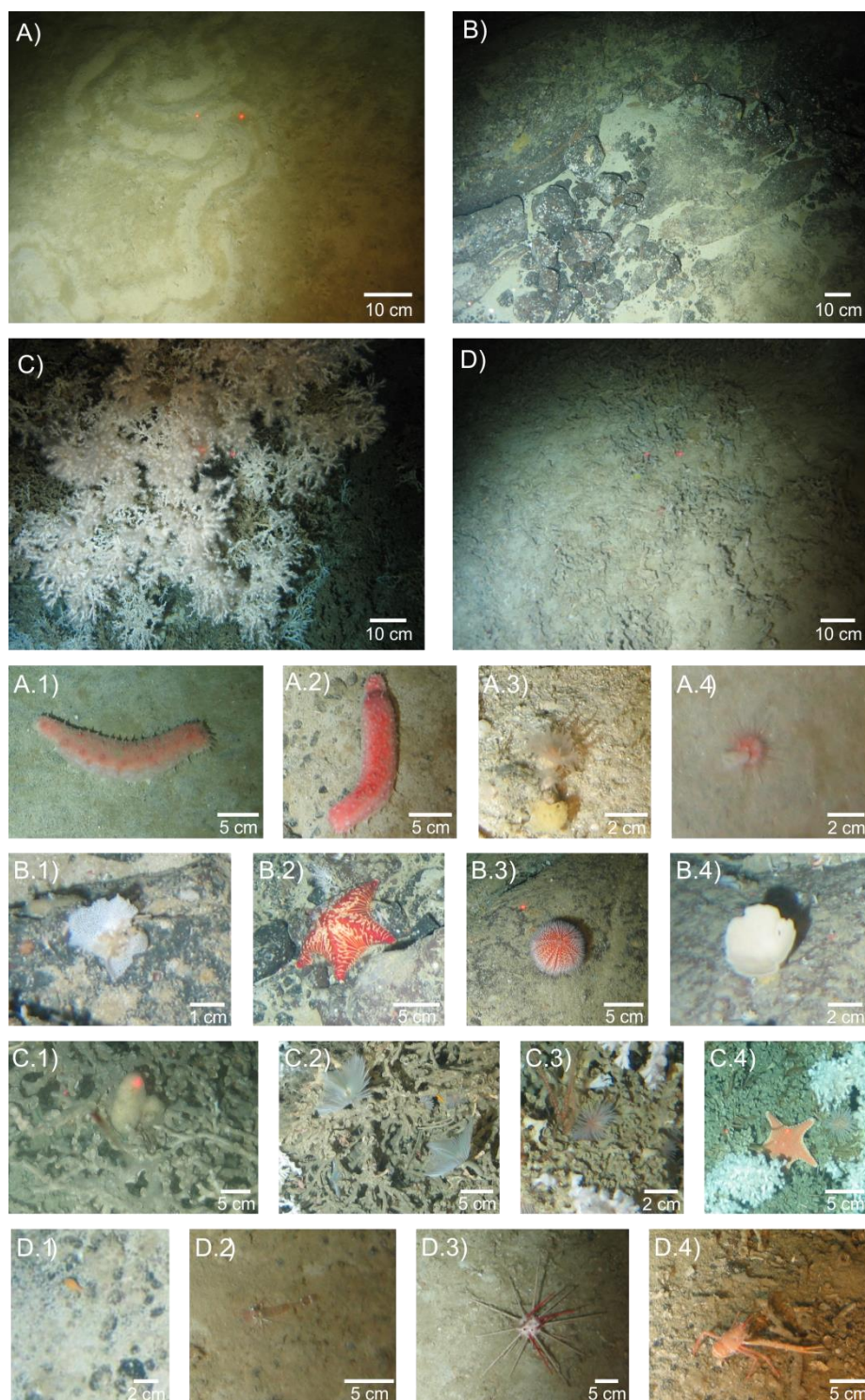
Figure 2.7: Characteristic images for substratum classes and associated morphospecies: A) Soft sediments; 1 and 2 - *Parastichopus tremulus*, 3- *Caryophyllia* sp. 1, 4- *Echinus* sp. 3 (possibly *E. elegans*), B) Hard substratum; 1- *Reteporella* sp., 2- *Porania pulvillus*, 3- *Echinus* sp. 2, 4- Porifera (cup sp. 1), C) Coral stands; 1- Porifera (yellow columnar sp. 1), 2- Sabellidea sp. 1, 3- Actiniaria sp. 1, 4- *Hippasteria* sp. 1, and D) Coral rubble; 1- Sipuncula sp. 1, 2- Caridea sp. 1, 3- *Cidaris cidaris*, 4- *Munida sarsi*, (species also frequently present on hard substratum and live coral stands).

areas (e.g. proportion of like adjacencies) were associated with a decrease in biodiversity (both models: P-value: <0.001). Percentage cover of corals, and a few associated indices indicative of larger more closely located patches (e.g. aggregation index (both models: P-value: <0.001), portion of like adjacencies (sonar only model: P-value: <0.05), mean core area index (sonar only model: P-value: <0.01) and effective mesh size (sonar only model: P-value: <0.01)), were also associated with increased biodiversity. Second-order relationships to percentage cover of rubble were also significant (P-value: <0.001) indicating that they can increase biodiversity when present in intermediate quantities.

Variation partitioning of the models allowed for the separation of the amount of variation explained by imagery versus sonar-derived environmental variables. When the model considering all environmental parameters was partitioned between imagery and sonar-derived environmental variables, fine-scale characterization of substratum percentage cover obtained from the imagery explained the most variation, contributing to a total of 38.6% for community composition and 59.0% for biodiversity (H′) (Figure 2.8). Of those percentages 16.4% and 31.0% were also explained by the selected class and landscape metrics. However, as the full models explained 45.1% and 63.8% of the variation in community composition and biodiversity respectively, it is clear that the variation explained by the class and landscape metrics was almost entirely captured by the imagery-derived substratum percentage cover. Forward selection of PCNMs resulted in two sets of 50 PCNMs which identified 31.1% and 36.0% of the variation, in the morphospecies matrix and biodiversity index (H′) respectively resulting from spatial structuring (supporting information, 2.8.7). Of these percentages associated with spatial structuring, only 5.9% (morphospecies composition) and 2.0% (biodiversity, H′) could not be explained by the environmental variables. This variation could be the result of unmeasured environmental variables or spatial autocorrelation resulting from biological interactions. Gaussian variogram fitting of site scores for significant canonical axes indicated that ranges in the spatial structure modelled varied between 20-50m.
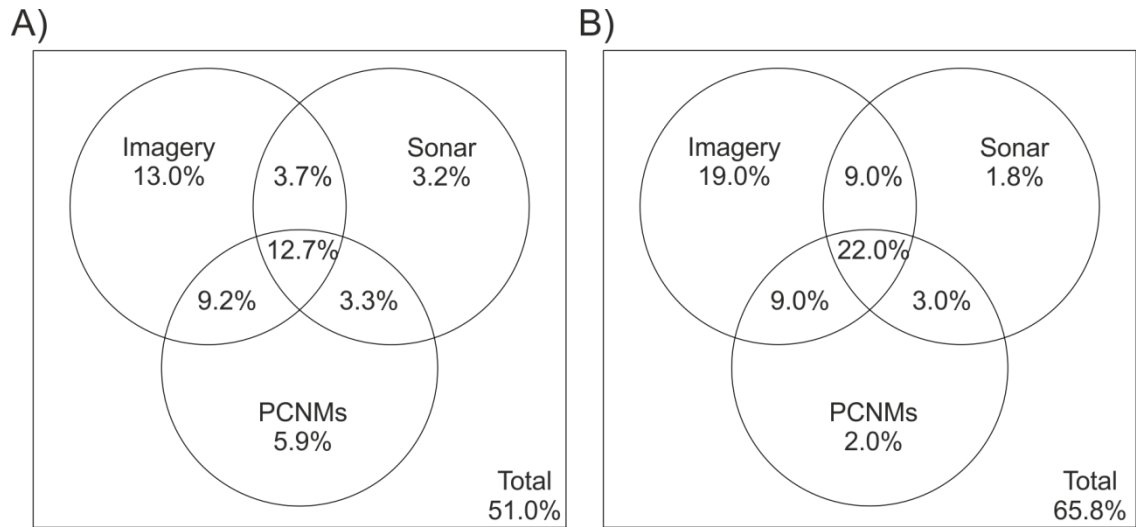
Figure 2.8 : Venn diagram showing percentages of variation in species composition and biodiversity (H'), for the full model, explained by selected environmental variables extracted from the imagery (percentage substratum cover), the sonar maps (backscatter, bathymetry, class and landscape metrics) and the Principal Coordinates of Neighbour Matrices. A) Morphospecies count matrix and B) Shannon diversity index.
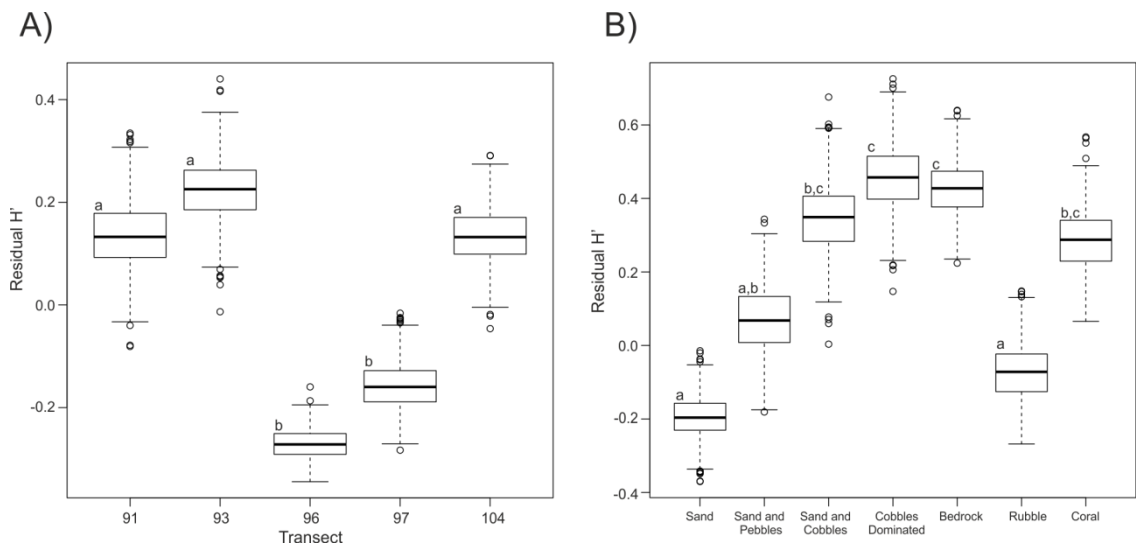
Figure 2.9: Difference in residuals of the Shannon diversity indices (H') between A) the five imagery transects and B) the seabed facies, after filtering out the spatial structure using significant Principal Coordinates of Neighbour Matrices. A bootstrapping procedure was used to standardize sample length. Letters indicate factors which are not significantly different based on 95% confidence intervals.

When model validity was assessed, following 100 repetitions, on average 81.1% (SD: 2.5%) of the 300 images removed were assigned to the same cluster as originally computed using the full model. When only sonar-derived environmental information was used, values of 71.8% (SD: 3.9%) were obtained.

At the broader scale, transects varied in morphospecies composition, diversity and substratum composition. The 95% confidence intervals for transects 96 and 97 indicated a significantly lower biodiversity (H′) than for the other three transects (Figure 2.9A). Biodiversity (H′) was lowest in soft sediment or rubble dominated images and highest when cobbles, bedrock or corals were present (Figure 2.9B). Significant negative relationships were observed between diversity (H′) (R^2: -0.167, P-value: <0.001) and organism abundance (R^2: -0.159, P-value: <0.001) versus depth. Percentage cover of hard substratum (R^2: -0.079, P-value: <0.001) and coral stands (R^2: -0.027, P-value: <0.001) also diminished with depth. Landscape metrics showed similarities across all five transects, but class metrics displayed greater variation. At the transect level (200m buffer), patch sizes averaged 256m^2 (SD: 41 m^2), but greater differences were apparent when separated by substratum types: 9,178m^2 (soft), 537m^2 (hard), 160m^2 (mixed), 21m^2 (bedrock), 14m^2 (rubble) and 20m^2 (coral). Soft sediments comprised the largest proportion of the landscape, while exposed bedrock, rubble fields and coral stands were rarely encountered. Exposed bedrock was only recorded along transects 91 and 104, while large rubble fields were only identified in the sonar maps of mission 44. Live coral stands of *L. pertusa* were recorded along transects 91, 93 and 104. Estimates of the proportion of landscape occupied by each substratum class were highly dependent on the resolution of the observation method. Greater differences were observed when comparing the high-resolution imagery to the sediment interpretation maps covering various extents (Table 2.5). Based on the imagery, soft sediments represented the dominant substratum class in all transects, but transects 91 and 104 had mixed sediments as their dominant class if sediment interpretation maps were used instead.

Table 2.5: Averaged characteristics by transects : substratum percentage cover for each sediment grain size class is given as obtained from the imagery, the circular (30m radius) area surrounding each image and the 200m buffer surrounding each transect.

Dive	No Image	Length (m)	Depth (m)	Imagery							
				Sand	Gravel	Pebbles	Cobbles	Boulders	Bedrock	Rubble	Coral
91	211	970	217.5	81.4	0.4	3.2	3.4	0.1	2.3	4.9	4.3
93	100	575	227.7	64.9	0.6	6.7	10.6	0.0	0.0	8.8	8.3
96	160	1200	317.3	98.6	0.0	0.4	0.6	0.1	0.0	0.4	0.0
97	369	2810	297.6	82.2	0.0	0.2	1.1	0.1	0.0	15.9	0.4
104	382	2490	188.6	44.7	7.1	25.6	6.0	0.8	8.3	2.1	5.4

Cont.	Thematic Maps (30m radius, Pictures)						Thematic Maps (200m radius, Transects)					
	Soft	Mixed	Hard	Bedrock	Rubble	Coral	Soft	Mixed	Hard	Bedrock	Rubble	Coral
91	41.0	46.1	12.0	0.0	0.0	1.0	40.0	45.1	14.2	0.0	0.0	0.7
93	45.4	33.0	20.1	0.0	0.0	1.5	48.3	33.5	16.6	0.0	0.0	1.6
96	55.8	22.7	20.6	0.0	0.8	0.0	51.6	22.3	24.7	0.0	1.5	0.0
97	42.5	22.7	31.9	0.0	2.9	0.0	51.2	20.2	26.1	0.0	2.5	0.0
104	22.0	42.0	20.8	14.2	0.0	1.0	18.6	49.7	21.9	9.8	0.0	0.0

2.5   Discussion

This study employed community and landscape analysis approaches traditionally employed in terrestrial settings to examine species-environment relationships based on remotely acquired data of Rockall Bank, Northeast Atlantic. This represents one of the first applications of these techniques to a deep-sea environment and the results obtained suggest potential applications for management and conservation requiring fine-scale species distribution and biodiversity information.

2.5.1   Megafaunal Distribution and Biodiversity

A strong association between morphospecies distributions and substratum types was observed, yielding three distinct assemblages: soft bottom, hard bottom and coral associated. Although not forming an independent group, the addition of a fourth cluster enabled the separation of morphospecies occurring in multiple substratum classes (hard bottom, coral stands and rubble) from those more closely associated with live coral stands.

No morphospecies were significantly associated with soft bottoms, but assemblages dominated by the sea cucumber *P. tremulus* have been reported for sandy flat areas by Howell (2010) in Rockall Trough and Buhl-Mortensen et al. (2012) in northern Norway. Howell (2010) also reported the concurrent occurrence of the pencil sea urchin, *Cidaris cidaris*, but in the current study this species was mostly observed in coral rubble fields. In this study, cup corals, *Caryophyllia* spp., were also observed on isolated cobbles or boulders in soft sediment dominated areas. The majority of morphospecies in the second assemblage were similar to descriptions of the rocky reef habitat of the central and eastern flanks of Rockall Bank (Howell et al. 2009). Being mostly dominated by encrusting sponge colonies as well as bryozoan species (Cyclostomatida and *Reteporella* sp.), taxonomic identification was not possible without sample collection and rendered comparisons between studies difficult. As previously reported from other surveys of the region, squat lobsters, *Munida* spp., were commonly observed associated with coarse sediments (gravel and cobbles) (Howell et al. 2009, Howell 2010) as well as on live coral stands and rubble (JNCC 2010a). Finally Sabellid worms and unidentified morphospecies of Actiniaria and Ophiuroidea were repeatedly observed in high densities on live coral stands. Association of suspension feeders (i.e. actiniarians, hydroids, hexactinellids and demosponges as well as crinoids and brisingiids) with *L. pertusa* has been described for the Franken Mound area on western Rockall Bank (Wienberg et al. 2008) and the nearby, though deeper, Rockall Trough (Masson et al. 2003). The distribution of similar habitats occupied by many of the same species has been reported for the nearby Hatton Bank (Roberts et al. 2008).

The assemblages observed in this study showed similarities to others described in areas characterized by iceberg reworking of the seabed (Gutt et al. 1996, Gutt & Piepenburg 2003, Jones et al. 2007a). During the Quaternary the majority of the western and southern flanks of Rockall Bank were left covered in scours averaging 2 - 2.3km in length, ranging between 50 and 200m in width and up to 8m in depth (Sacchetti et al. 2012). As found in other areas of the Northwest European continental margin (Freiwald et al. 1999, Wheeler et al. 2007), iceberg debris provided hard ground for cold-water coral colonies to establish. Association of coral colonies with hard substratum originating from iceberg activity has also been observed on the continental margin off Nova

Scotia and Newfoundland, Northwest Atlantic (Edinger et al. 2011). Fine-scale distinctions in morphospecies assemblages between the coarser and finer sediments dominating respectively the edges and centre sections of iceberg ploughmarks have been reported for other Northeast Atlantic areas such as Hatton Basin (JNCC 2012), the Wyville Thomson Ridge (JNCC 2010b) and northwest of Shetland (Bett 2001). However, studies in areas of active iceberg formation (Arctic and Antarctic) have shown that the observed patterns in benthic species assemblages and biodiversity are highly dependent on the time elapsed since iceberg disturbance, particularly with respect to slow growing organisms (e.g. sponges and corals) whose presence can affect distributions of associated species (Gutt et al. 1996, Gutt & Piepenburg 2003). Higher diversity on ploughmark edges dominated by hard sediments has also been reported for the Fimbul ice shelf region, Antarctica, and dominance of bryozoans and encrusting sponge colonies was mostly observed in less disturbed areas (Jones et al. 2007a). At broader spatial scales, when higher diversity is observed in areas of medium level iceberg activity, it often results from the coexistence of various stages of succession and recovery (Jones et al. 2007b, Teixidó et al. 2007). In the case of Rockall Bank this high diversity is more likely the result of a spatial dependence of distinct species assemblages on substratum type which varies over fine spatial scales.

Although much variation in species-depth relationships occurs at local scales, over regional scales reductions in faunal densities and biomass, attributed to a diminishing organic matter supply, and unimodal responses in diversity have frequently been reported (Rex et al. 2006). In the current study, the depth range encountered was relatively small (~100m), and showed a strong colinearity with other environmental variables. For example, higher number of coral colonies and presence of exposed bedrock were observed along the shallower transect, while the deeper transects located at the northern end of the study area were characterized by large expenses of coral rubble fields (likely resulting from previous trawling activities). Those deeper transects had the lowest Shannon diversity index. Cold-water coral stands increase habitat complexity by providing three dimensional structures to which fishing activities such as trawling can cause damage (Fosså et al. 2002). As many species are associated with cold-water corals, their destruction is expected to impact composition and diversity of benthic fauna (Roberts &

Hirshfield 2004, Henry & Roberts 2007). Our study indicated that coral rubble fields did not appear to be exploited by specific morphospecies assemblages and megabenthic diversity was generally reduced. This contrasts with results obtained by Roberts et al. (2008) who found high diversity of epifauna associated with areas of naturally occurring coral rubble on Hatton Bank. Although images where coral stands were present did not have a significantly higher biodiversity than images dominated by cobbles or bedrock, corals did harbour a specific morphospecies assemblage whose positive influence on diversity was visible at the transect level. Much of the diversity associated with cobbles or bedrock substrata was composed by hard-to-distinguish small encrusting sponge colonies, while large fish aggregations were seen surrounding coral patches, but were not included in this analysis.

2.5.2   Landscape Approach and Environmental Variables

The importance of different environmental variables depended on whether multivariate morphospecies distributions or a diversity index (H′) were considered. In both cases, environmental variables with the highest resolution (~cm), such as substratum percentage cover as obtained from imagery, were most useful in explaining the variation observed. The use of landscape metrics explained 24.7% of the variation in morphospecies distribution, values similar to those reported by Teixidó et al. (2007) using similar analyses. Using canonical correspondence analysis to examine structural patterns (characterized by landscape metrics) of successional stages of iceberg disturbance and their effect on an Antarctic benthic community, they found the two first axes to explain 11% of the variation observed. In our study, although sonar backscatter maps were available at high resolutions (0.5x0.5 m pixel size), the presence of small single coral colonies or isolated boulders which were observed within pictures, could not be identified in the backscatter. However, their presence within areas of homogeneous soft sediment did affect megafaunal composition. In this environment, these small seafloor features may represent 'keystone structures' (Tews et al. 2004) whose importance could only be captured through ROV imagery. This difference in resolution caused estimates of substratum percentage cover to vary more between acquisition methods than between the different spatial extents considered for the sediment interpretation maps. Hence in this highly heterogeneous setting, the

sole availability of high-resolution acoustic surveys, even if covering larger extent, would result in a ~35-45% decrease in the amount of biological variation that could otherwise be explained by cm-scale substratum type information. Similarly, the classification of substratum types that was achieved visually provided additional information that is not yet achievable via automated seabed classification of backscatter data. The accuracy of the sediment class assigned is also likely to be higher when based on photographs.

The high-resolution sidescan sonar was effective in mapping seafloor heterogeneity resulting from iceberg reworking and highlighting a similar spatial pattern in megafaunal distribution indicative of spatial induced dependence. Indeed, the ranges identified by variogram fitting, 20-50m, were consistent with iceberg ploughmark widths and patch sizes observed over the larger transect extents (200m buffer). Taking into account this seafloor heterogeneity by examining the surroundings of each image collected using class and landscape metrics resulted in an additional 12-17% in variation explained over the use of only sediment interpretation maps and bathymetry-derived information as proxies for morphospecies assemblages or biodiversity.

The comparatively low resolution (20x20m pixel size), small bathymetric gradient (~100m), limited extent considered and absence of topographic features may explain why other bathymetry-derived variables showed poor predictive abilities. However, in other seafloor regions where sediment characteristics do not exhibit the high spatial heterogeneity observed in this study, and where geomorphology shows greater complexity, bathymetric parameters such as slope, curvature and BPI may reveal significant broad-scale patterns in species assemblages (Jones & Brewer 2012).

The majority of the metrics selected when morphospecies distributions were considered described patch area and shape, whereas metrics describing areas of greater spatial heterogeneity, such as aggregation metrics, were selected more often when describing variation in biodiversity (H′). In northern Norway, diversity increased in areas of mixed sediment (Buhl-Mortensen et al. 2012). In our study, this increase in diversity resulted from very fine-scale partitioning of the seabed based on sediment hardness. This fine-scale

partitioning could be described by substratum percentage cover analysis of the digital stills, but was often lost in the sediment interpretation of the backscatter, yielding a transition zone of mixed sediment class with higher biodiversity, but of insufficient resolution to describe morphospecies assemblages. Effects of patch edges and transition zones on species composition and abundance have also been frequently documented in terrestrial systems showing positive, negative and neutral relationships (Ries et al. 2004). However, this important concept of landscape ecology has very rarely been examined in marine benthic ecosystems (but see Zajac et al. (2003) for macrobenthos and Anderson et al. (2009) for deep-water demersal fishes). Landscape metrics will only provide valuable information if the spatial patterns they describe can be linked to specific ecological mechanisms and the development of marine specific metrics will be required (Wedding et al. 2011). Three-dimensional metrics taking into account the structure of the overlying water masses affecting bentho-pelagic linkages would be one area of potential research. Experimental manipulation of landscape characteristics would also be required to establish links between landscape spatial patterns and specific ecological processes.

2.6 Conclusion

Our study showed that for the Rockall Bank area, a strong association between morphospecies distribution and sediment characteristics resulting from past iceberg activity exists. The identified relationships were stronger when fine-scale (<m) imagery-derived sediment percentage cover information was available. However, taking into account the surrounding (30 and 75m) spatial context in which sediment patches were located, using class and landscape metrics, nearly doubled the amount of variation in morphospecies composition and biodiversity (H') that could be explained when compared to the use of sediment interpretation maps and bathymetry alone. As the inclusion of this approach does not require that additional time be spent collecting data, it shows potential for increasing our understanding of the links between environmental structure and ecological processes, and may become a useful concomitant step for informing management decisions.

Although we currently do not have the possibility to describe large extents of seafloor to <m scale resolution levels, the scale (20-50m) of the ecological process described in this study clearly indicates that most ship-based surveys, which for the depth considered are often processed to pixels sizes of 20-50m in resolution, would be inadequate to capture the variation observed. For example, the European Nature Information System (EUNIS) classification (Connor et al. 2004) describes the western flank of the Rockall Bank area as mixed sediments separated by depth zones (deep circalittoral and upper slope). Although accurate for the spatial extent and resolution considered, finer-scale studies are still required to understand how much of the biological complexity might be underrepresented and inform conservation measures, particularly those aimed at specific species such as cold-water corals.

2.7   Acknowledgements

2.8 Supporting Information

2.8.1 Effect of Cryptic Species on Diversity
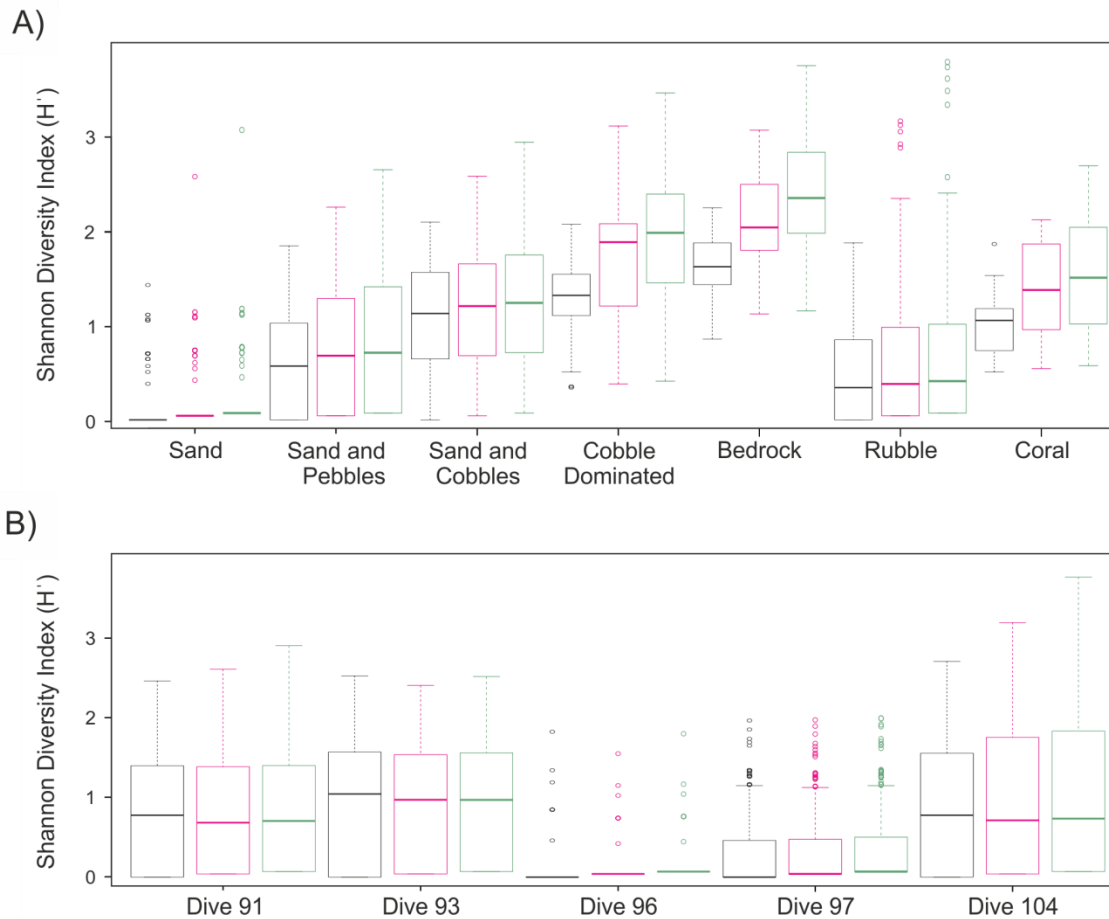
Figure S2.1: Boxplots showing variation in Shannon biodiversity index (H') observed for A) each substratum type and B) each dive (black). When two or more specimens of a morphospecies of Ascidians (111), Bryozoans (135), Cerianthids (8), Cormatulids (65), Sponges (119 or 85 for Demosponges), Ophiuroids (205), Sabellid worms (99) or Sipuncula worms (144) were observed within an image, each individual was randomly assigned, with replacement, to a potential cryptic species and biodiversity recalculated. The number of potential cryptic species (previously shown in brackets) was estimated by setting a percentage (pink, 5% and green, 20%) of the number of taxa listed in OBIS (Ocean Biogeographic Information System, http://www.iobis.org/) for the North Atlantic ocean.

2.8.2 Seabed Facies

Sand

Sand and Gravel

Sand and Pebbles

Sand and Cobbles

Cobble Dominated

Coral Rubble

Live Coral Stand

Exposed Bedrock
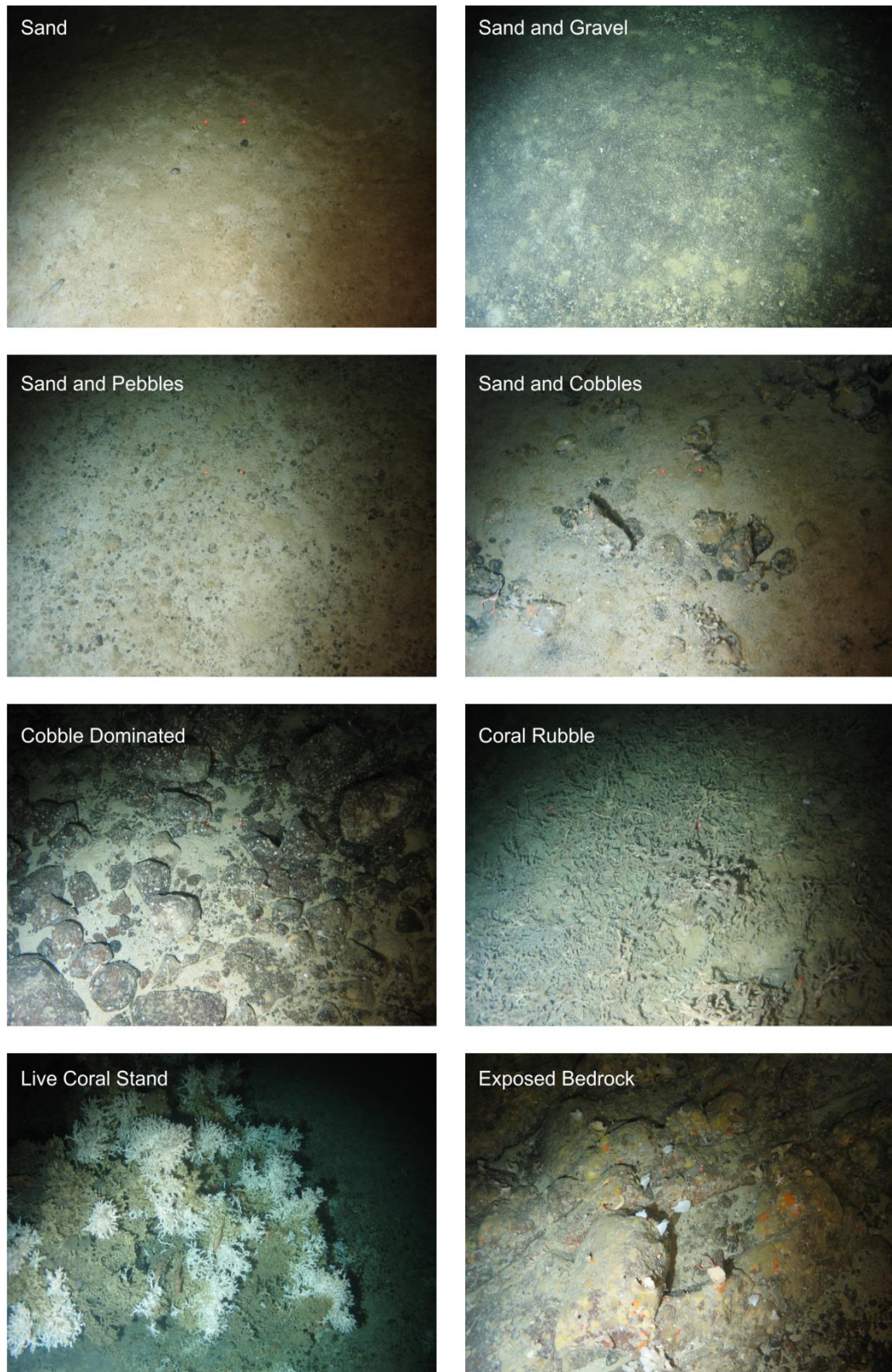
Figure S2.2: Example of images representing the different seabed facies observed along the five remotely operated imagery transects on Rockall Bank. Lasers dots are separated by 10cm.

2.8.3 Backscatter and Sediment Interpretation Maps
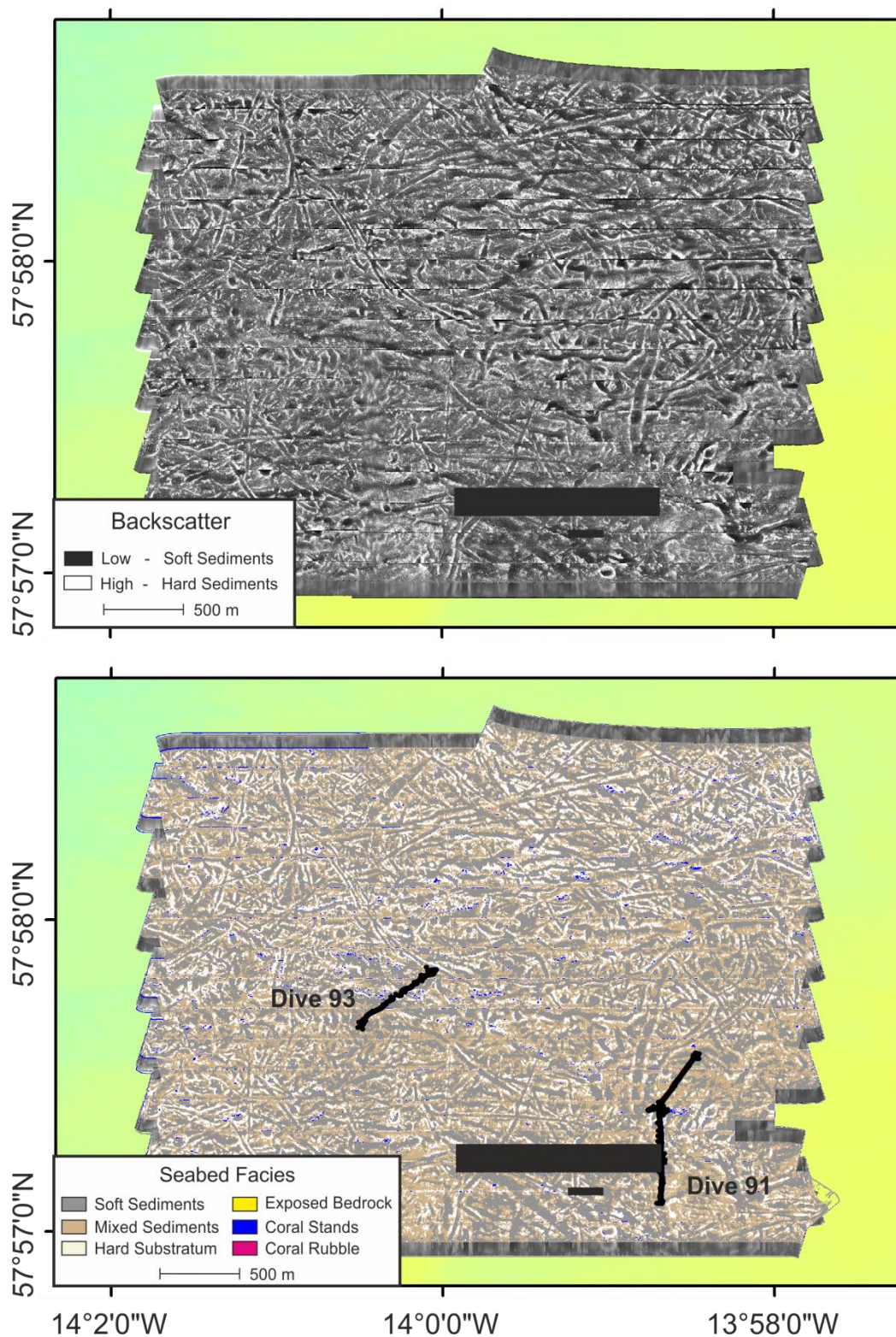
Figure S2.3: Autonomous underwater vehicle mission 43 sidescan sonar backscatter map (top) and sediment interpretation map with overlaid remotely operated vehicle imagery transects (bottom). Black boxes show areas where no data were available.
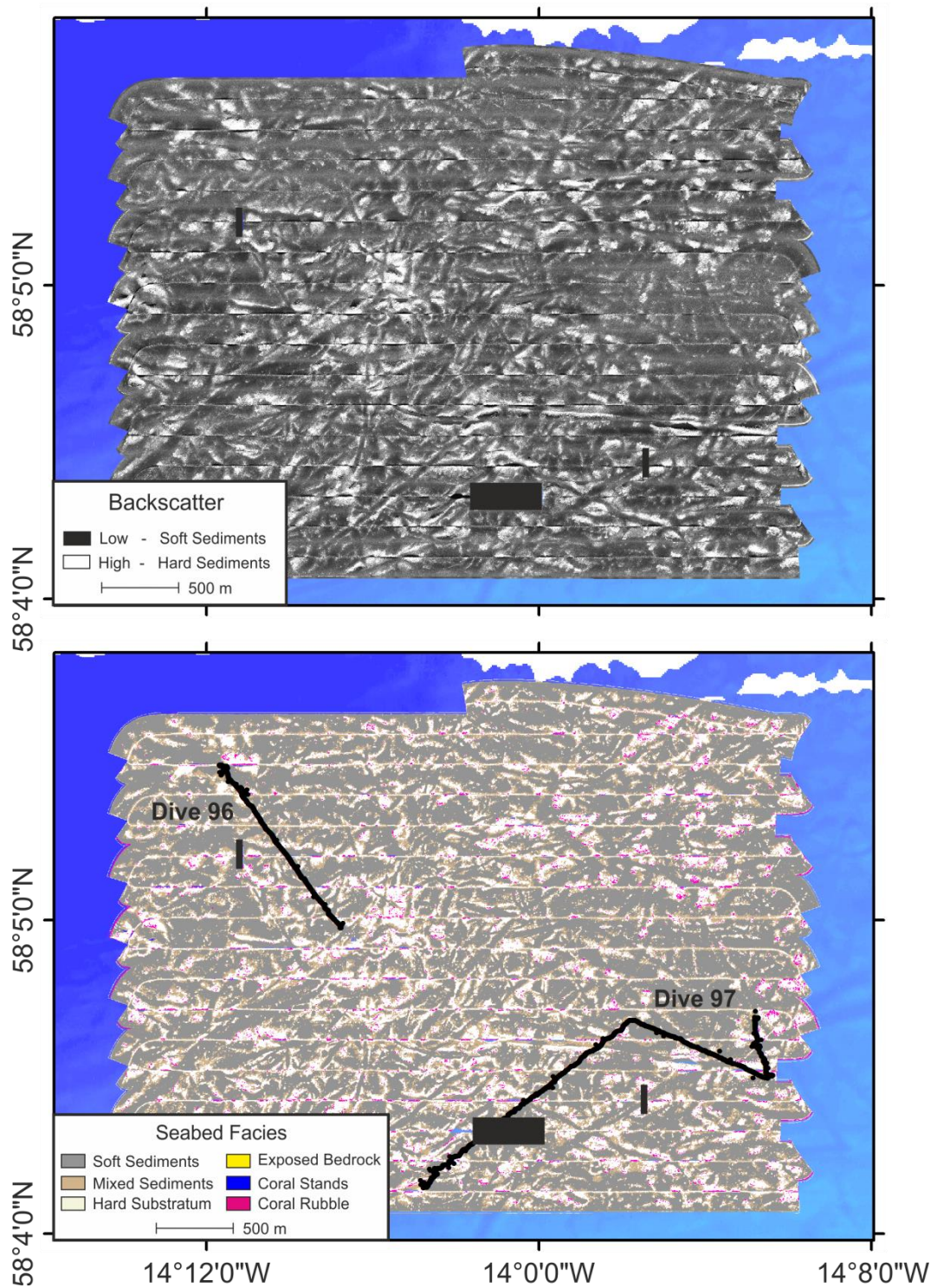
Figure S2.4: Autonomous underwater vehicle mission 44 sidescan sonar backscatter map (top) and sediment interpretation map with overlaid remotely operated vehicle imagery transects (bottom). Black boxes show areas where no data were available.
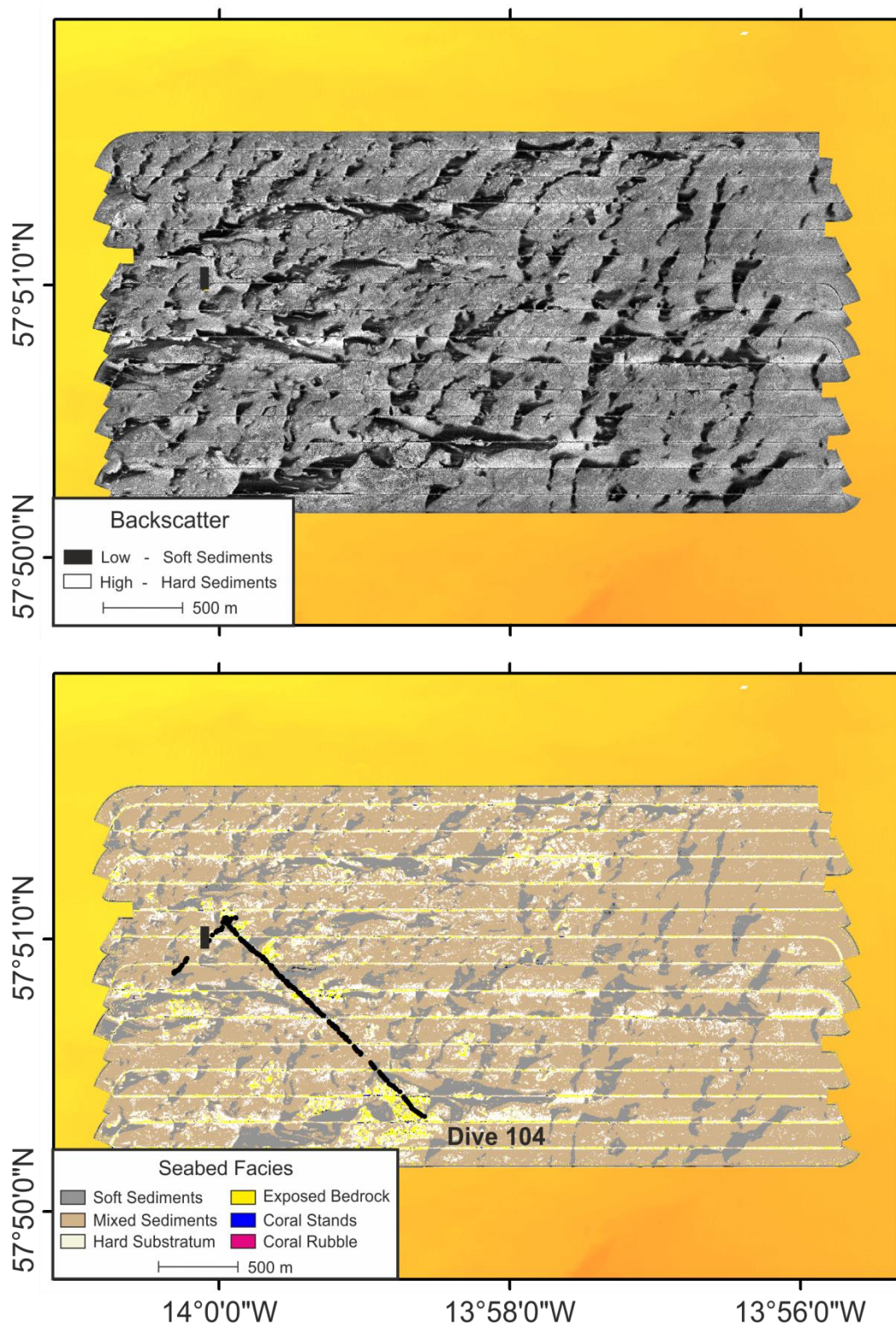
Figure S2.5: Autonomous underwater vehicle mission 45 sidescan sonar backscatter map (top) and sediment interpretation map with overlaid remotely operated vehicle imagery transects (bottom). Black boxes show areas where no data were available.

2.8.4 Detailed Descriptions of Statistical Techniques

2.8.4.1 Environmental Variable Selection

RDA (van den Wollenberg 1977) was chosen because an *adjusted R^2* could be obtained using Ezekiel's adjustment formula (Peres-Neto et al. 2006). Because any explanatory variable included in a model increases the R^2 regardless of whether it models responses or noise, this inflation must be taken into account and the resulting *adjusted R^2* used instead. This was considered particularly important for this study as a large number of environmental variables were involved (e.g. percentage covers of each substratum type, class and landscape metrics as well as bathymetric and backscatter derived variables). In RDA, species are assumed to be linearly related to environmental variables and second-degree explanatory variables for substratum composition were also introduced into the model to account for potential non-linear relationships (e.g. unimodel biodiversity relationships to percentage cover of rubble).

Model significance was assessed using permutation tests with a minimal permutation number of 999 (Makarenkov & Legendre 2002). If found significant, forward selection by permutation tests based on P-values was applied to reduce the number of explanatory variables and obtain a more parsimonious model (Borcard et al. 2011). Owing to the large number of available explanatory variables, forward selection was first applied to each environmental dataset separately (substratum percentage cover, class metrics, landscape metrics and bathymetric variables). The selected variables were then combined, a second forward selection process was applied and variance inflation factors (VIF) were used to exclude additional explanatory variables which showed strong collinearity with others present within the model. Only explanatory variables with VIF below 5 were retained (Stine 1995). When examining the influence of the previously described environmental factors on biodiversity (H'), multiple linear regression was used and forward selection was based on the Akaike Information Criterion (AIC), which is a measure based on the compromise between model simplicity and increased fit obtained when using additional parameters (Crawley 2005). For both the multivariate and univariate analyses, the forward selection process was carried out twice, once

including all environmental variables and once including only sonar-derived variables.

The morphospecies by images collected matrix was log(x+1) transformed before the analysis to reduce the influence of abundant species (Clarke & Warwick 2001). Variation partitioning was used to examine how much variation could be explained by environmental variables obtained from the imagery or the sonar maps when controlling for the effect of the other (Borcard & Legendre 2002, Peres-Neto et al. 2006). Variation partitioning is a technique which using multiple RDAs allows for the fraction of variation explained by a set of explanatory variables (e.g. imagery derived environmental descriptors) to be separated from the fraction of variation explained by another set of explanatory variables (e.g. sonar-derived environmental descriptors). As some of the variation will be explained by both sets of explanatory variables, the total amount of variation explained will be less than the sum of the two fractions examined (Borcard et al. 2011).

2.8.4.2 Species Assemblages

K-means partitioning was used on the morphospecies matrix to separate images into clusters representing different species assemblages. This method aims at separating observations into a predefined number of clusters based on sample proximity when data points are positioned in a space where each species represents an axis. Data points are assigned a cluster in such a manner as to minimize the distance between individual samples and the center of their assigned cluster (Hartigan & Wong 1979). Points are swapped between clusters and cluster centers are updated accordingly until a local optimal solution is achieved. As this is an iterative method and a predefined number of clusters is required, the analysis was successively conducted using 2 to 10 clusters and 100 random starting configurations. The number of clusters was chosen using the 'simple structure index', as defined in Dimitradou et al. (2002), and which takes into account elements describing intra and inter cluster structure. Analysis of similarities (ANOSIM) was also carried out in R to assess significant differences between the identified clusters (based on 999 permutations and Euclidian distance). The representative species for each cluster were determined using the 'species indicator values' (Dufrêne & Legendre 1997).

2.8.4.3 Model Evaluation

In order to evaluate the goodness of fit of the model, assess its ability to predict new data points and as a single dataset was available to evaluate the model, a 'holdout partition' approach was taken (Kohavi 1995, Manel et al. 1999, Verfaillie et al. 2009). In a 'holdout partition' approach, a subset of the data, in this case 300 randomly selected images (with proportional representation by transect), is removed from the original dataset and the model (RDA) parameters are recomputed. The removed data points are then reclassified based on the recomputed model and the percentage of data points reassigned the same cluster is calculated. The initial sample is then returned to the dataset and a subsequent sample is removed. The process is repeated multiple times to estimate variability in the results obtained. For this study, the environmental variables associated with each removed image were used to predict updated positions in ordination space (using RDA axis 1 and 2) based on the RDA model recomputed when the 300 randomly selected images were removed. The Euclidean distance to the centroid of each K-means cluster was then computed and the removed images were assigned the class of the closest cluster (Guisan & Zimmermann 2000, Anderson & Willis 2003, Wang et al. 2004). The percentage of removed images reassigned to the same cluster was calculated and the process was repeated 100 times.

2.8.4.4 Spatial Dependency Assessment

To determine whether the spatial structure present in species distribution was the result of spatial structuring in the environmental factors (induced spatial dependence), we carried out a Principal Coordinates of Neighbour Matrices (PCNM) approach (Borcard & Legendre 2002, Dray et al. 2006). PCNMs are a type of distance-based eigenvector maps, and are computed using a truncated Euclidean distance matrix based on sample location (distance threshold of 110m) followed by a Principal Coordinate Analysis. Only eigenvectors with associated positive eigenvalues and a significant Moran's index of spatial autocorrelation (Moran 1950, Sokal & Oden 1978, Dray et al. 2006) (an index used to test that the pattern represented is spatially structure and not random) are retained. As these eigenvectors are representations of spatial structure based on distances between sampling sites, RDA is used to select the eigenvectors which significantly represent the spatial structure also

present in the biological data. To determine which environmental variables explained the spatial structure detected, the resulting significant canonical axes were regressed against the environmental dataset. Gaussian variograms were also fitted to the significant canonical axes. Variograms describe the relationship between the variance in values observed at points with regards to their separation in space and can be used to estimate the range of spatial dependency (Bellier et al. 2007, Loots et al. 2011). Analyses were carried out using the 'vegan', 'labdsv' and 'gstat' package in R.

2.8.5 Morphospecies List

Table S2.1: List and count of the morphospecies observed along the five remotely operated imagery transects on Rockall Bank. Sample images were deposited in the SERPENT media archive (http://archive.serpentproject.com/).

Index	Name\Description	Count	Index	Name\Description	Count
1	*Parastichopus tremulus*	185	42	Porifera Cup sp1	1
2	*Munida sarsi*	2373	43	*Caryophyllia* sp2	6
3	*Cidaris cidaris*	59	44	Cephalopoda sp1	2
4	*Asterias rubens*	28	45	Hydrozoa sp2	5
5	*Henricia* spp.	117	46	Porifera White Branching sp2	10
6	*Reteporella* sp1	942	47	Holothuroidea sp1	7
7	*Lophelia pertusa*	39	48	Hydrozoa sp3	2
8	Sabellidae	272	49	Holothuroidea sp2	2
9	Actinaria sp1	227	50	Porifera Orange Lobose sp2	1
10	*Echinus* sp1 (possibly *E. acutus*)	19	51	Anthozoa sp2	1
11	Portunidea sp1	24	52	Pennatulacea sp1	7
12	Caridea sp1	80	53	Porifera Cup sp2	1
13	Orange Worm sp1	233	54	*Stylaster* sp1	2
14	Porifera Brown Cup sp1	44	55	Paguridae spp	1
15	*Hippasteria* sp1	19	56	*Ceramaster* sp1	1
16	*Echinus* sp2	10	57	Porifera Pink Encrusting spp	1
17	Translucent Ascidiacea spp	96	58	*Bolocera tuediae*	1
18	Cyclostomatida sp1	175	59	Pycnogonidae sp1	1
19	Porifera Orange Encrusting spp	138	60	Asteroidea sp2	1
20	Porifera White Branching sp1	152	61	Crinoidae spp	5
21	Porifera Brown Lamellate sp1	69	62	*Kophobelemnon stelliferum*	5
22	Porifera White Encrusting spp	445	63	Buccinidae spp	3
23	Ceriantharia sp1	13	64	Custacea sp1	2
24	*Caryophyllia* sp1	92	65	*Nephrops norvegicus*	1
25	Porifera Yellow Columnar sp1	28	66	*Caryophyllia smithii*	1
26	Ophiuroidae (rock) spp	263	67	Asteroidea sp2	1
27	Porifera Yellow Encrusting spp	705	68	Unknown 1	10
28	*Porania pulvillus*	10	69	Asteroidea sp3	3
29	Porifera Blue Encrusting spp	38	70	Porifera White Lamellate sp1	6
30	Porifera Orange Branching sp1	19	71	Ophiuroidae (Sand) spp	6
31	Hydrozoa sp1	35	72	Porifera Yellow Lobose sp1	10
32	Porifera Tan Columnar sp1	48	73	Cnidaria sp1	1
33	Porifera Yellow Spherical sp1	36	74	Brachiura sp1	1
34	*Echnius* sp 3 (possibliy *E. elegans*	11	75	Cnidaria sp2	1
35	Porifera Red Encrusting spp	60	76	Porifera Spherical sp1	1
36	*Actinauge richardi*	6	77	Porifera sp1	17
37	*Phelliactis* sp1	4	78	Asteroidea sp4	1
38	Asteroidea sp1	4	79	Unknown 2	2
39	Porifera Yellow Branching sp1	7	80	Asteroidea sp5	2
40	Porifera Orange Lobose sp1	5	81	*Paromola cuvieri*	3
41	Porifera Orange Branching sp2	2			
Total				**7267**	

2.8.6 Choice of Area for Calculation of Class and Landscape Metrics
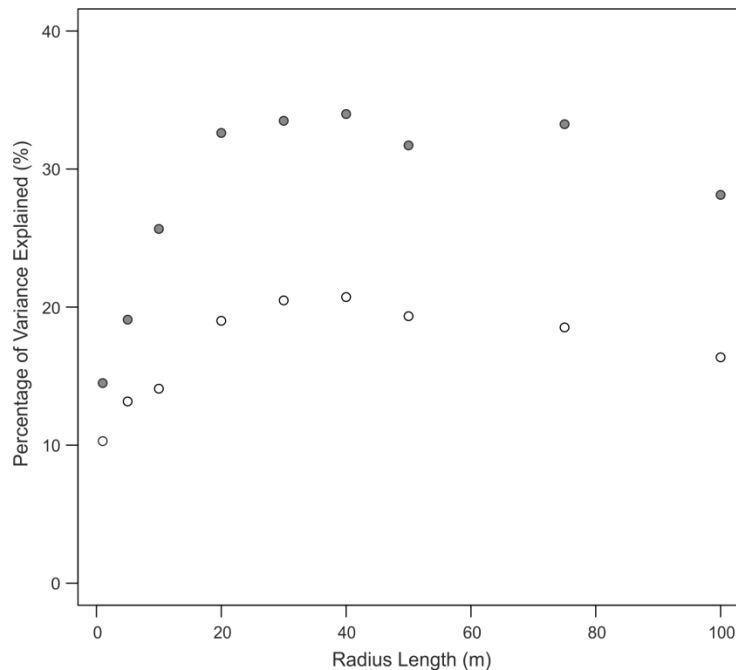
Figure S2.6: Percentage of variance explained following forward selection of class and landscape metrics calculated using circular areas of varying radius length; biodiversity (H') (grey) and morphospecies composition (white).

Table S2.2: Comparison of the percentage of variation explained when class and landscape metrics calculated from two areas of different sizes are considered in combination.

					Radius Length (m)					
Percentage of Variation Explained (%)										
		1	**5**	**10**	**20**	**30**	**40**	**50**	**75**	**100**
	1		20.0	26.8	34.2	35.2	36.0	34.1	36.9	33.5
	5	*13.2*		26.9	33.6	34.6	35.6	34.3	37.5	34.7
	10	*15.5*	*16.3*		35.4	36.3	35.3	35.5	35.5	36.5
Radius Length (m)	**20**	*20.5*	*21.4*	*21.6*		37.1	37.7	38.1	**41.2**	39.3
	30	*21.6*	*22.6*	*22.8*	*23.8*		37.0	36.8	40.7	38.3
	40	*22.4*	*23.6*	*23.8*	*25.0*	*23.9*		36.2	40.0	37.7
	50	*21.6*	*22.9*	*23.8*	*25.2*	*24.2*	*23.3*		36.8	35.0
	75	*22.0*	*23.9*	*24.2*	*26.0*	***26.3***	*25.0*	*23.7*		35.7
	100	*20.6*	*22.0*	*23.1*	*25.7*	*26.0*	*25.1*	*24.0*	*23.2*	

Species Composition (lower-left); Biodiversity (H') (upper-right)

2.8.7   Spatial Structure as Identified using PCNMs

The first 13 canonical axes of the redundancy analysis of Principal Coordinates of Neighbour Matrices (PCNMs) against the morphospecies count matrix were significant. When the first two axes were regressed against the selected set of environmental variables, those related to presence, percentage cover or proportion of landscape occupied by coral, cobbles or exposed bedrock had the strongest influence (P-value: <0.001). In the case of biodiversity (H′), the significant PCNMs showed a strong negative response to percentage cover of sand (P-value: <0.001) and a positive association to presence of bedrock (P-value: <0.001) and percentage cover of coral present (P-value: <0.01) (Figure S2.7).

A)

B)

C)
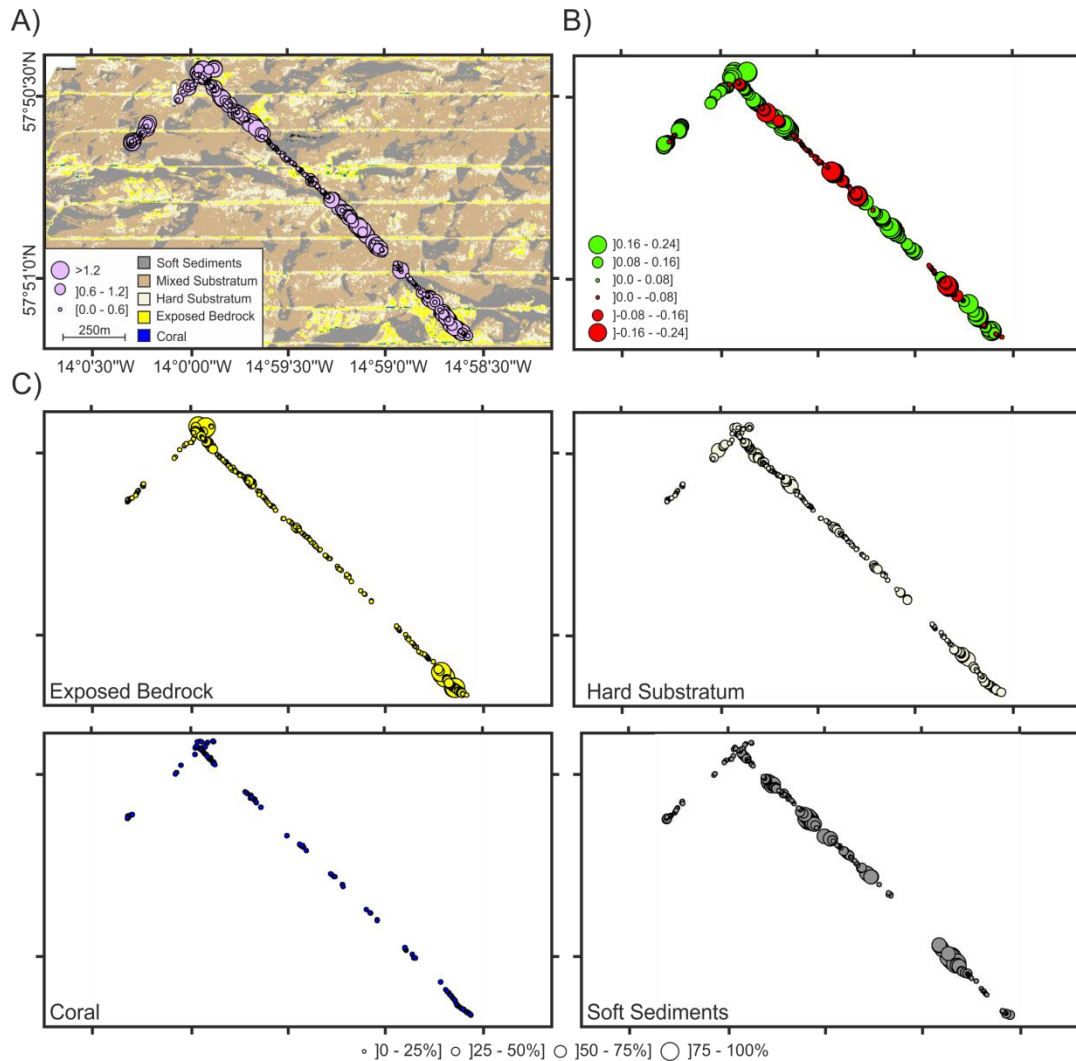
Figure S2.7: Spatial structure along transect 104 A) Distribution of Shannon diversity index for each image overlaying a section of the sediment interpretation map from mission 45. B) Fitted site scores obtained using the Principal Coordinates of Neighbour Matrices approach. These represent the strength of the spatial structure identified as present in Shannon diversity index. C) Proportions of landscape represented by different seabed facies within a 30 m radius of each image and showing similar spatial structuring as depicted by the fitted side scores.

Chapter 3:
(In) consistency of predictive mapping approaches for deep-water habitats: Considering multiple model outputs

Katleen Robert[1], Daniel O.B. Jones[2], J. Murray Roberts[3,4,5] and Veerle A.I. Huvenne[2]

[1] School of Ocean and Earth Science, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, UK

[2] National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

[3] Centre for Marine Biodiversity and Biotechnology, Heriot-Watt University, Edinburgh, EH14 4AS, UK

[4] Center for Marine Science, University of North Carolina Wilmington, 601 S. College Road, Wilmington, NC 28403-5928, USA

[5] Scottish Association for Marine Science, Scottish Marine Institute, Oban, Argyll, PA37 IQA, UK

3.1   Abstract

In the deep sea, the expenses associated with data collection often result in sparse data. As such, models capturing relationships between the observed fauna and environmental variables (acquired via acoustic mapping techniques) are often used to produce full coverage species assemblage maps. Many statistical modelling techniques are being developed, but as data are limited, there remains a need to determine the most appropriate mapping approaches and assess the ability of models to predict at new survey locations. Predictive habitat modelling approaches (redundancy analysis, maximum entropy and random forest) were applied to an heterogeneous section of seabed on Rockall Bank, NE Atlantic. The predictive maps were based on remotely operated vehicle (ROV) imagery transects, autonomous underwater vehicle (AUV) sidescan backscatter maps and ship-based multi-beam bathymetry. A year later a section was revisited and two additional ROV transects were collected to ground-truth the different predictive maps. Internal model assessment using a split-sample approach on the initially collected data showed similar fair performances for the three models tested. However, comparison with the independent transects clearly demonstrated that the heterogeneity of the area was not adequately captured. Ensemble mapping techniques, where the outputs of many models are combined, were better able to show spatial trends, and may be more useful for marine spatial management. Different statistical approaches for predictive habitat modelling possess different strengths and weaknesses and by examining the outputs of different modelling techniques, more robust predictions, with better described variation and areas of uncertainties, can be achieved and provide important information for management.

3.2 Introduction

As the anthropogenic footprint extends deeper into our oceans, reliable descriptions of the seafloor and the species present are required to devise appropriate management and conservation measures. With very limited areas of seafloor mapped at comparable resolution to terrestrial environments (Sandwell *et al.* 2006), quantitative spatial information regarding distributions of marine biotic and abiotic components is needed to build benthic habitat maps (Kostylev *et al.* 2001). Monitoring changes in species distribution or habitat coverage over time may be facilitated by the creation of reliable and thoroughly assessed high-resolution maps, improving our ability for early detection of anthropogenic impacts. Recent advances in acoustic techniques for seafloor mapping (Brown *et al.* 2011) have made it possible to create detailed geomorphological maps more rapidly. However, the biological information needed to supplement complete coverage topographic and geological maps has remained limited owing to the time-consuming process of specimen collection and taxonomic identification (Przeslawski *et al.* 2011).

As such, full coverage biological sampling is usually not an option, and hierarchical approaches involving nested survey designs are often employed. They involve a combination of broader-scale geological map creation and detailed ground-truthing biological studies covering smaller spatial extents, often taking the form of imagery transects. The two are then combined using statistical modelling approaches to form habitat maps (Brown *et al.* 2011). The manner in which the environmental and biological data are combined is crucial. Whereas in the top-down approach environmental variables are used to define patches representing areas of similar characteristics, in the bottom-up approach the relationships between fine-scale environmental variables and species distributions are modelled and then predicted across the larger extent covered by the acoustic surveys (Hewitt *et al.* 2004).

In recent years there have been an increasing number of studies employing a variety of bottom-up modelling techniques to produce predictive full coverage megabenthic invertebrate habitat maps: maximum entropy (Rengstorf et al. 2012, Ross & Howell 2012), many types of decision or classification trees (Gonzalez-Mirelis & Lindegarth 2012, Compton et al. 2013), a variety of multivariate analyses or ordination methods (Shumchenia & King

2010, Buhl-Mortensen et al. 2012), general additive models (GAMs), neural networks (Palialexis *et al.* 2011) and many more. Some of these techniques, such as maximum entropy, are based on records of presence only (with pseudo-absences selected from background points), as obtaining reliable absence data can be particularly difficult (Pearce & Boyce 2006). On the other hand, if suitable absence data is available, other methods are available, such as redundancy analysis (RDA), GAMs or random forest. Comparisons between presence only models and presence-absence models have found the latter to perform better by providing more information regarding unsuitable habitats (Brotons et al. 2004, Pearson et al. 2006). Presence only data may also be more susceptible to sampling biases, and as such background selection must be carefully considered (Phillips *et al.* 2009).

During map creation a commonly employed evaluation technique is to set aside a portion of the collected data (testing dataset), build the model using the remaining data (training dataset) and then test the model using the removed portion (Elith & Leathwick 2009). Although problematic for small datasets (Fielding & Bell 1997, Guisan & Zimmermann 2000), this is the technique usually employed in deep-sea studies. Because of the high expense associated with sample collection and considering the large areas for which samples are still lacking, collecting additional datasets is often unaffordable. As such, it is rare to have the opportunity to collect an independent dataset to assess model predictions at new locations.

In this study, we used benthic imagery data (photographs and extracted video frames), in addition to acoustic maps, to produce full coverage predictive maps for megabenthic invertebrate species assemblages and diversity, as represented by the Shannon-Wiener diversity index, H' (Shannon 1948). The study area was located on Rockall Bank, where detailed habitat maps were needed to assess which one of two proposed cold-water coral protection zones may be most effective. Three different modelling approaches were investigated for species assemblages: redundancy analysis (RDA) (ter Braak 1994), maximum entropy (MaxEnt) (Phillips & Dudík 2008) and random forest (RF) (Breiman 2001), representing three very different modelling approaches ('assemble and predict together', RDA, and 'assemble first, predict later' using a presence only model, MaxEnt, as well as a presence-absence model, RF)

(Ferrier & Guisan 2006). With these three strategies showing different strengths and weaknesses, addressed in Ferrier & Guisan (2006), we used a resampling scheme to obtain prediction variability maps as well as overall measures of model performance to build weighted ensemble maps. These maps take into account predictions and uncertainties from more than one model in order to strengthen modelling outcomes (Araújo & New 2007, Marmion et al. 2009b). Two models were also employed to examine spatial variation in biodiversity: linear regression and regressive random forest. All models were first assessed internally using a split-sample approach, but their outputs were also tested for their ability to predict outside the originally collected transects by using two additional imagery transects collected during a subsequent cruise.

3.3 Methods

3.3.1 Initial Survey Design

As part of the UK's 'Marine Environmental Mapping Programme' (MAREMAP; http://www.maremap.ac.uk/index.html) and the 'COmplex Deep-sea Environments: Mapping habitat heterogeneity As Proxy for biodiversity' project (CODEMAP; http://www.codemap.eu/), a section of the western flank of Rockall Bank, Northeast Atlantic, was mapped during the RRS *James Cook* -060 cruise carried out in May - June 2011 (Figure 3.1). Over 380km^2 of ship-based multibeam bathymetry (pixel size of 20x20m), three 12-13km^2 Autosub6000 autonomous underwater vehicle (AUV) sidescan sonar surveys (pixel size of 0.5x0.5m) and five *Lynx* remotely operated vehicle (ROV) photographic imagery (2592x1944 pixels) transects (1,222 images along ~8km using a Kongsberg OE14-208 digital stills camera) were collected. All individual organisms larger than 1cm were counted and identified, using morphospecies when species level identification could not be achieved. Parallel lasers (with 10cm separation) were mounted on the ROV to provide a scale on all recorded images. Positioning was achieved using the ROV's ultra-short baseline (USBL) navigation system. One of the main reasons for carrying out this survey was to assess which of two conservation zones would be most effective at protecting cold-water corals: a 'Fisheries Closure' established in 2007 or a nearly

overlapping, but slightly extended area put forth as candidate 'Special Area of Conservation' (cSAC) in 2011 (JNCC 2010a) (Figure 3.1).
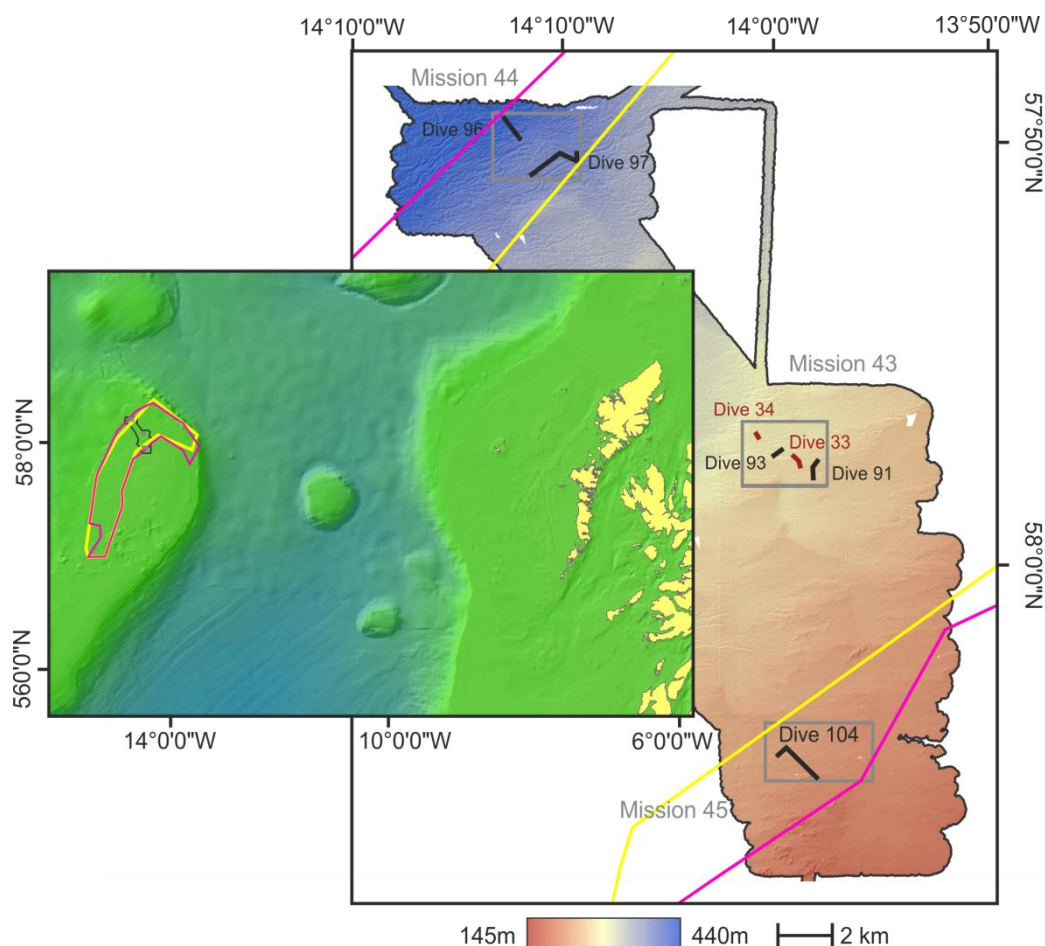
Figure 3.1: Map of the surveys carried out on Rockall Bank, Northeast Atlantic . Ship-based bathymetry displayed with superimposed outlines of the sidescan sonar data (grey boxes) collected during three autonomous underwater vehicle missions. The remotely operated vehicle imagery transects carried out during the *JC*-060 cruise are shown in black and the two from *JC*-073 in red. The boundaries of a 2007 fisheries closure area and a candidate for 'Special Area of Conservation' are illustrated in yellow and pink respectively. Background bathymetry of the Northeast Atlantic from GEBCO (General Bathymetric Chart of the Oceans,(IOC IHO and BODC 2003).

Environmental descriptors were derived from both sidescan backscatter (EdgeTech FS2200, 410kHz) and multibeam bathymetry (Kongsberg EM710) maps. The sidescan backscatter maps had been classified to sediment interpretation maps (0.5x0.5m pixel size) representing six seabed facies (soft and mixed sediments, hard substratum, exposed bedrock as well as coral stand and rubble) using an unsupervised classification (Robert et al. 2014).

From the sediment interpretation maps, class and landscape indices were derived to describe the shape, size, diversity and spatial arrangement (connectivity) of habitat patches. Class metrics are used to describe properties of patches from a single substratum type, while landscape metrics are used to characterise all patches present within a landscape (McGarigal et al. 2012). As spatial arrangement of habitat patches can have an effect on ecological processes (Turner & Gardner 1991), these indices can provide additional environmental descriptors useful in explaining species distributions. Indeed, when these datasets were included as explanatory variables in an RDA of this community, the use of class and landscape indices increased the amount of variation explained by ~40% (Robert et al. 2014).

Landscape and class metrics were calculated for each pixel of the sediment interpretation maps using moving windows (at two scales: 60x60m and 150x150m, based on the scales identified in Chapter 2). Owing to the large number of computations involved, the high performance computer cluster IRIDIS 3 (University of Southampton) was used to run an R script (R Development Core Team 2014) written for parallel computation. The R package 'SDMTools' was used to compute the metrics and the package 'Snowfall' was used to run the computations in parallel making better use of the capacities of IRIDIS 3. On smaller datasets, these computations could easily be accomplished on a regular desktop computer running R in either parallel or sequential mode (R code presented in supporting information, 3.8.1). The multibeam bathymetry was gridded at 0.5x0.5m using bilinear interpolation to allow values for each predictor variable to be extracted for each pixel of the sidescan backscatter map.

3.3.2 Predictive Modelling

Influential environmental variables to be included in the predictive models were determined using forward selection in the redundancy analysis and were used for all three modelling techniques (Table 2.2 and 2.3) (Robert et al. 2014). Predictions were made for every pixel of the sidescan backscatter maps (0.5m resolution) because of the high degree of heterogeneity in seabed composition observed in this region. The four species assemblages (soft and mixed sediment, hard substratum and coral associated assemblages) identified

and described in Robert et al. (2014) using K-mean classification, ANOSIM and 'species indicator values' were used. Although low numbers of individuals were generally found in soft sediments, the solitary coral *Caryophyllia* sp and the holothurian *Parastichopus tremulus* were most commonly observed. Hard substratum was characterized by Bryozoan species (mostly *Reteporella* sp) and various sponge morphotypes, while mixed sediments were dominated by the abundant squat lobster *Munida sarsi*. Species associated with the cold-water coral *Lophelia pertusa* included Sabellid worms, an unsampled Actinarian sp and many Asteroid spp. Analyses were carried out using the R libraries 'vegan', 'randomForest', 'dismo', 'raster' and 'caret'.

3.3.2.1   Redundancy Analysis

The models (redundancy analysis (RDA) for species assemblages and linear regression for diversity) built and described in Robert et al. (2014) were used to create the first set of full coverage fine-scale biological maps. Similarly to Oldeland et al. (2010), for species assemblage predictions, we used the estimated coefficient of the linear combination of environmental predictors to position each pixel along the canonical axes. To assign each pixel to a species assemblage, a fuzzy classification using the K-means method (Hartigan & Wong 1979) (fuzzy exponent = 2, 4 clusters), which minimises the sum of squares between individual points and their cluster centers, was carried out on the predicted pixel positions along the first four canonical axes (representing 90% of the explained variance). Fuzzy clustering was used as it returns a 'membership value' providing an estimate of the probability (and hence confidence) of each data point belonging to each of the clusters considered (Lucieer & Lucieer 2009).

3.3.2.2   Maximum Entropy

Maximum entropy (MaxEnt) predicts an index of relative habitat suitability using presence data compared to randomly selected background (pseudo-absences) points (Phillips & Dudík 2008). This technique allows for predictor variables to be fitted, using many different kinds of relationships (e.g. linear, product, quadratic, hinge, threshold and categorical), to the occurrence data. A probability density of species occurrence which minimizes the distance to the probability density of the covariates as they occur in space (relative

entropy) is created (Elith *et al.* 2011). This method aims at building distributions without imposing additional unfounded constraints, hence maximizing entropy.

The software MAXENT (version 3.3.3, freely available online http://www.cs.princeton.edu/~schapire/maxent/) was employed with sampling bias grids to help account for the transect design. Bias grids were built based on a Gaussian kernel estimation (with SD of 500m) of sampling density which resulted in a weighting surface with more weight given to areas close to sampled locations (Clements *et al.* 2012). Based on this weighing scheme, pixels in the vicinity of the ROV transects were more likely to be selected as pseudo-absence points, with progressively fewer selected as distance from transects increased. Predictions were made for each of the four species assemblages (soft and mixed sediment, hard substratum and coral associated).

3.3.2.3   **Random Forest**

Random forest (RF) is a technique that allows for the building of multiple regression or classification trees for a dataset, hence the term forest (Breiman 2001). Each tree is built based on a sub-sample of the data and at each node the data are split based on the best predictor variable, selected out of a smaller number of randomly selected variables. Once grown, the trees can be used to make predictions based on the rules developed at each node, but now each tree in the forest provides an answer and the class assigned is based on the majority. A probability estimate can be obtained based on the number of votes given to each class for a given pixel. To reduce the forest error rate a large number of trees of low correlations must be built (Oshiro et al. 2012). However, increasing the number of trees past a certain threshold will only lead to higher computational costs with limited improvements. Additionally, enough environmental variables must be available at each node to produce a tree capable of classifying the data.

Two forests were built using the *JC*-060 data: one classification case, using the same species clusters as for MaxEnt, and one regression case for biodiversity (H'). Varying numbers of trees and environmental variables were assessed and a forest containing 1,000 trees and considering 15

environmental predictors per node was selected. Predictions for each pixel of the acoustic map were then carried out.

3.3.3   Model Evaluation

The original assessment using the *JC*-060 dataset for all three models was carried out using 100 holdout partitions, where 30% of the samples were randomly removed and the models reconstructed. The habitat suitabilities predicted for the removed data points, for which the dominant species assemblage was known, were used to calculate the area under the curve (AUC) of the receiver operating characteristics (ROC) (Fielding & Bell 1997, Manel et al. 2001). Ten of the models built using the partitions were randomly selected and used to produce full coverage maps from which a mean and standard deviation were calculated for each pixel, giving a representation of prediction variability across space.

During the *JC*-073 cruise carried out in June 2012 as part of the UK Ocean Acidification programme's 'Changing Oceans Expedition', two *Holland I* ROV transects were carried out within the area surveyed during AUV mission 43 (Figure 3.1). They consisted of 0.80km (Transect 33) and 0.30km (Transect 34) high-resolution (1920x1080 pixels) video transects (Insite Mini Zeus camera with direct HDSDI fibre output). Frames were extracted every ~4m to recreate the distance separating the *JC*-60 digital stills. The biological imagery was treated as for the *JC*-060 data and organisms larger than 1cm were identified into morphospecies by the same observer. To examine comparability of the communities observed during the 2011 and 2012 surveys, a MDS plot was created.

The *JC*-073 data were used to evaluate independently the predictions from the holdout partitions. In addition, new models were created combining all data points (*JC*-060 and *JC*-073). They were again evaluated with the holdout partitions. Finally to assess the effects of the transect design and the influence of survey and video system, predictions were also made based on models where all combinations of two transects from the *JC*-060 data were removed. This was carried out to examine whether differences in the *JC*-073 data could account for the differences observed in prediction abilities by determining whether accurate predictions at *JC*-060 data locations could be

achieved. In the case of the diversity predictions, partial Mantel tests (using Euclidean distances and based on 999 permutations) (Leduc *et al.* 1992) were carried out to evaluate the correlation between the predictions and the observed diversity (H') in the images collected during the *JC*-073 cruise while controlling for spatial proximity.

3.3.4   Ensemble Predictions

Considering that different models are likely to produce different predictive outputs, but with each containing separate information and areas of uncertainties, the idea of ensemble predictions is to summarise a range of potential outcomes to produce more robust predictions (Araújo & New 2007). For each statistical approach, using the predicted habitat suitability values for each of the holdout partitions, an average was obtained weighted by the AUC result for each of the faunal assemblages. These maps were further combined by averaging the predictions of all three statistical approaches for each faunal assemblage using an inverse-variance weighted average, based on the variability captured by the 10 holdout partitions. The AUC values of the resulting ensemble maps were calculated, threshold values based on prevalence were derived (Liu *et al.* 2005) and used to calculate measures of accuracy (the proportion of correctly assigned presences and absences over total sample size).

To show the complexity and gradual changes observed between species assemblages an RGB image was created based on the ensemble habitat suitability maps using three of the four species assemblages (coral, soft sediment and hard substratum associated) (Compton *et al.* 2013).   To obtain simplified categorical maps, the species assemblage with the highest prediction for each pixel was determined. To show the importance of the high-resolution acoustic data, predictive maps were also created based on sedimentation interpretation maps coarsened to 10m resolution for which the same environmental descriptors were derived. To examine similarities between the different predictive maps of both resolutions, the Hellinger-based niche similarity metric described in Warren et al. (2008) was computed. This measure can vary from 0 (no overlap) to 1 (identical niches).

3.4 Results

A total of 46 morphospecies were observed in the two *JC*-073 transects, of which only 1 morphospecies with more than 5 individuals (a grey encrusting sponge) had not been observed in the previous cruise. Diversity (H') for transect 33 was 1.63 (based on 194 images and 0.80km), and for transect 34 was 1.78 (based on 63 images and 0.30km) as compared to the *JC*-060 data: transect 91 H' = 2.42 (based on 211 images and 0.97km) and transect 93 H' = 2.17 (based on 100 images and 0.58km). However from the MDS plot showing no clustering by transect (supporting information, 3.8.2), it is clear that the *JC*-073 data sampled the same communities as observed during *JC*-060.
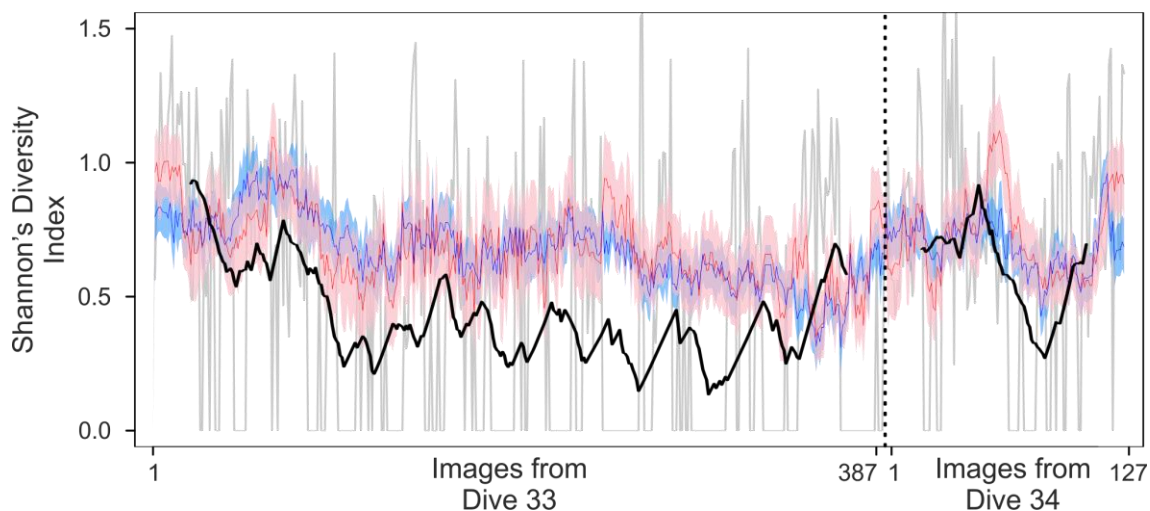
Figure 3.2: Shannon-Wiener diversity index (H') comparison between observed (light grey) and smoothed (black) using a 50 image moving average window for transects 33 and 34 with predictions based on the linear regression (blue) and regressive random forest (red). Shaded areas show the variation in predictions based on 10 holdout partitions.

Comparison of the diversity predictive maps produced using linear regression and RF models on the *JC*-060 data showed similarities (r: 0.25, simulated P-value: 0.001, Figure 3.3): both predict higher diversity in areas dominated by hard substratum and coral stands. When the outputs from the linear regression and RF models were compared to the diversity estimated from the *JC*-073 images, the two predictive models showed only limited agreement to the actual data. Neither RF nor linear regression were able to
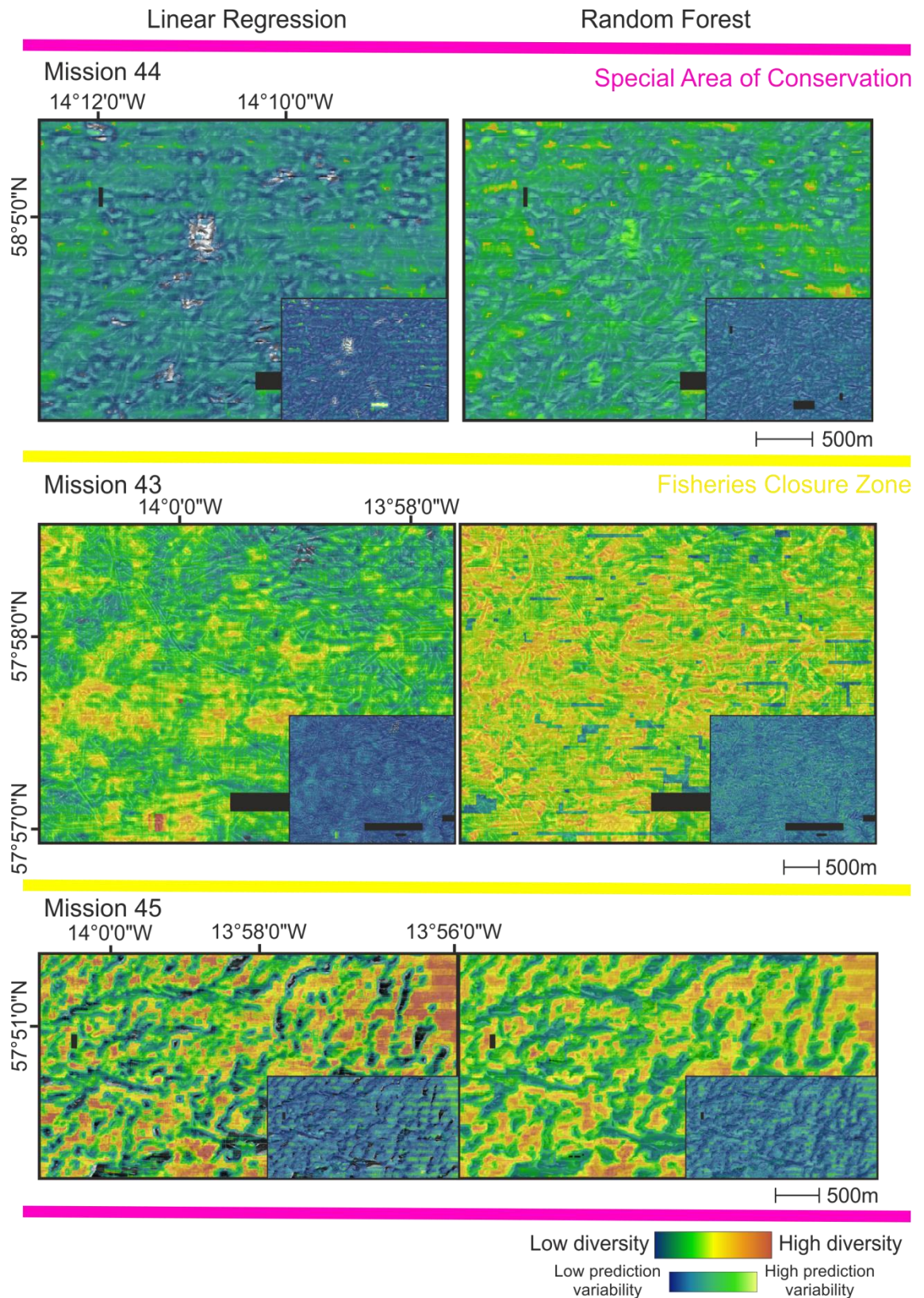
Figure 3.3: Shannon-Wiener diversity index (H') prediction maps using linear regression and regressive random forest, averaged based on 10 holdout partitions. Smaller insets show standard deviation of predictions.
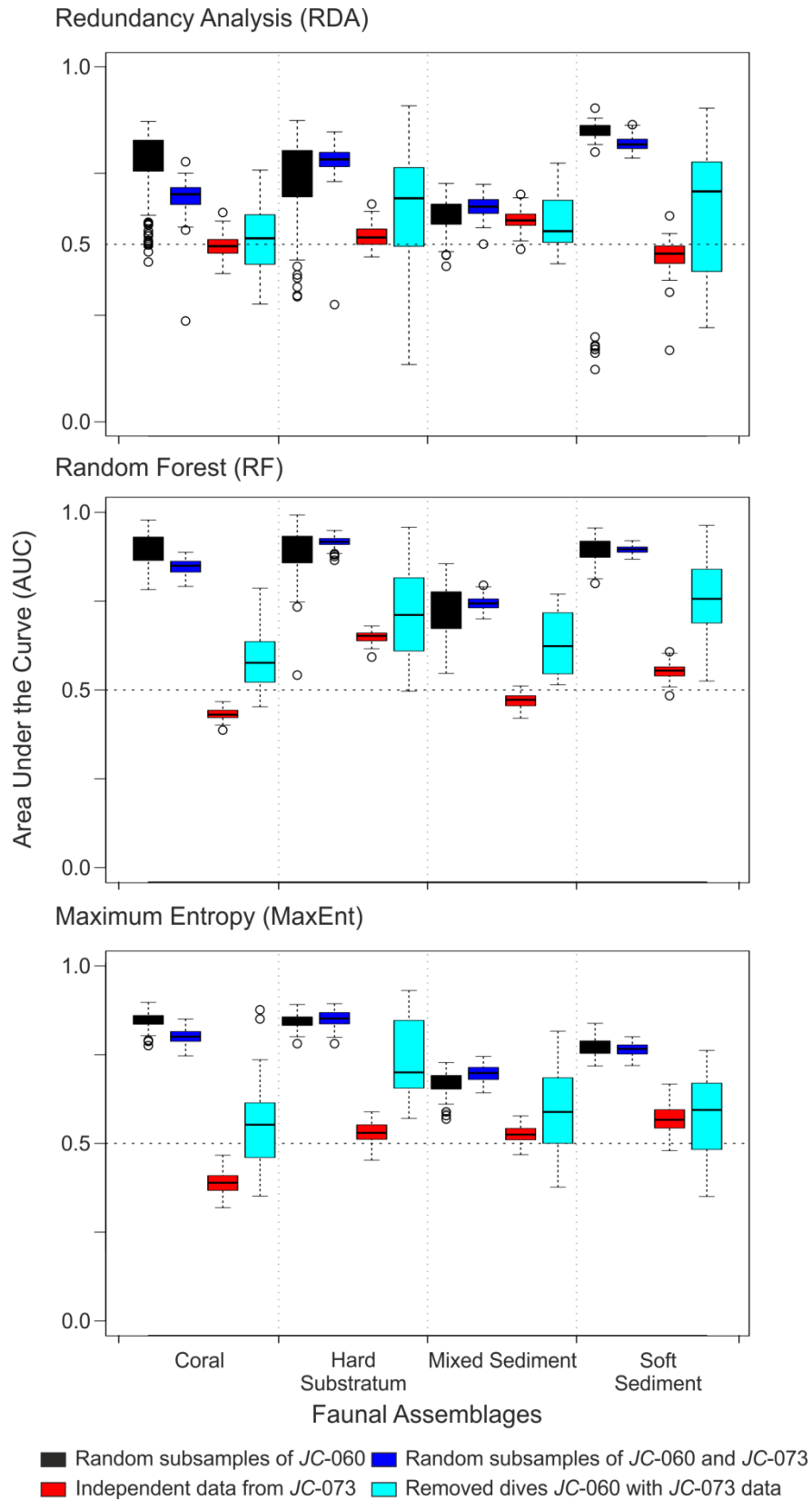
Figure 3.4: Area Under the Curve (AUC) observed based on 100 holdout partitions for four different species assemblages.

capture the finer-scale variations in biodiversity. However, when biodiversity was smoothed using a moving average window of 50 images, partial Mantel tests showed significant relationships to the predictions obtained using RF (r: 0.23, simulated P-value: 0.001) and linear regression (r: 0.20, simulated P-value: 0.001) (Figure 3.2).

For species assemblages, the three models produced fair results when internal validation was carried out on the *JC*-060 data (Figure 3.4, black boxes). The holdout partition process for the RF classifier showed average AUC values ranging from 0.72 SD= 0.06 (mixed sediments associated) to 0.89 SD= 0.05 (hard substratum associated), values of 0.59 SD= 0.05 (mixed sediment associated) to 0.79 SD= 0.14 (soft sediments associated) for RDA and 0.65 SD= 0.03 (mixed sediment associated) to 0.84 SD= 0.02 (coral associated) for MaxEnt (Figure 3.4). Addition of the *JC*-073 data into the models showed a general slight improvement on the average AUC values for most models (Figure 3.4, dark blue boxes). However, when attempting to use the *JC*-060 data to predict species assemblages at the *JC*-073 sample locations, much poorer results were obtained (Figure 3.4, red boxes). None of the models were able to predict reliably coral associated faunal assemblages, with AUC values of 0.44 SD= 0.02 for RF, 0.52 SD= 0.04 for RDA and 0.41 SD= 0.025 for MaxEnt, but different models fared better for different species assemblages. Although still poor, mixed sediments associated fauna were better captured by RDA, while RF was more useful for hard and soft sediment associated fauna. When transects from *JC*-060 were instead removed (Figure 3.4, light blue boxes), prediction performances were more variable, but in average remained higher than predictions for the *JC*-073 transects, particularly for MaxEnt and RF.

The three models produced different maps of habitat suitability for the various species assemblages (Figure 3.5), but measure of environmental niche indicated similarities between model predictions with mean values of 0.77 (for coral and soft sediment associated fauna) and 0.79 (for hard substratum and mixed sediment associated fauna). The areas of variability did also differ between models and predictions made using weighted averages combining all three models exhibited a more consistent accuracy across species assemblages (Figure 3.6). Soft sediment associated fauna showed poor predictive accuracy when *JC*-060 data was considered, but predictions performed better on the *JC*-
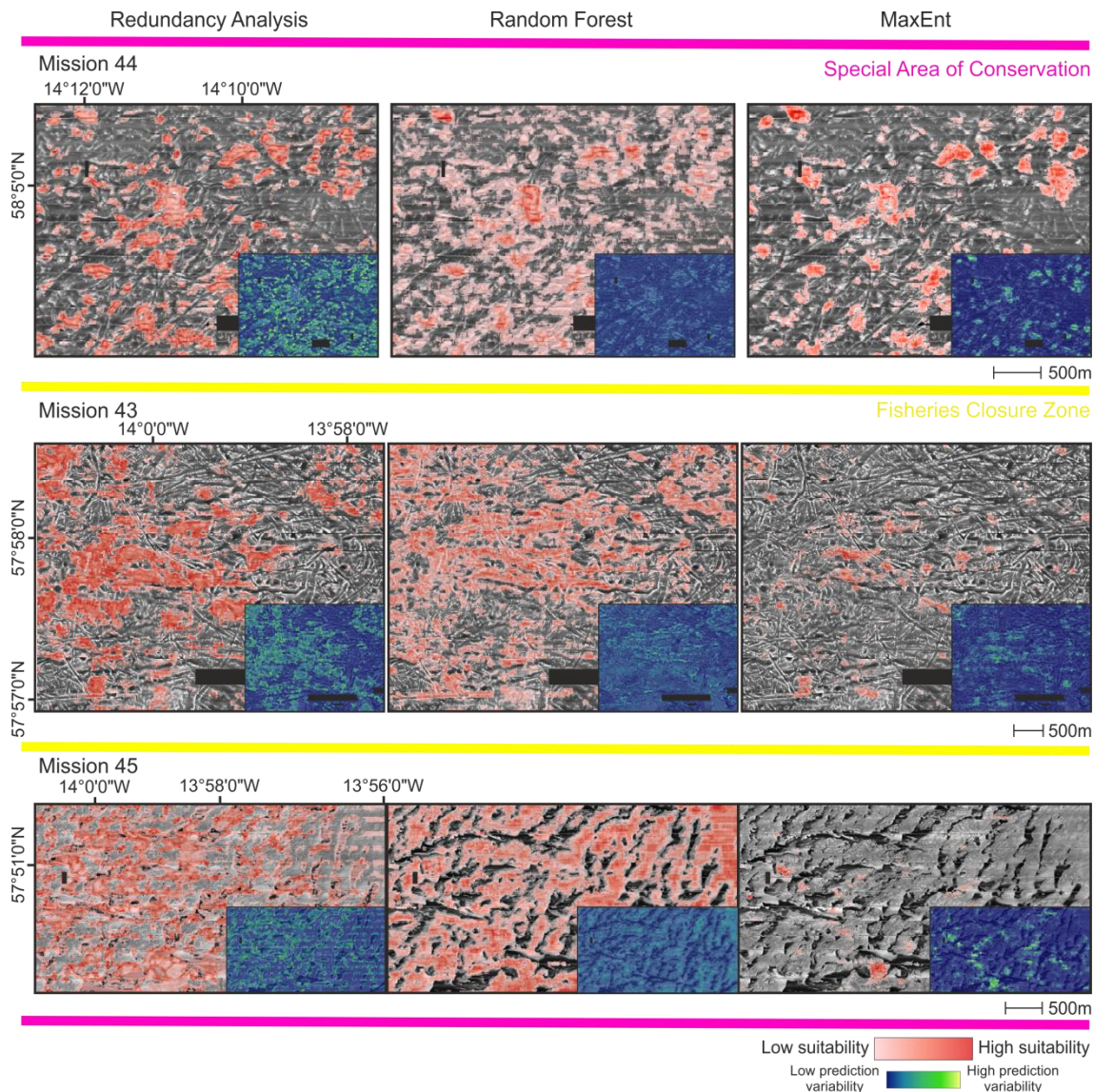
Figure 3.5: Maps showing the relative habitat suitability of the coral associated fauna for three different areas around two conservation zone boundaries on Rockall Bank based on three statistical techniques. Smaller insets show standard deviation of predictions based on 10 holdout partitions. Maps for the other three groups are presented in supporting information, 3.8.3.

073 data. On the other hand, corals in the *JC*-073 data were best captured by MaxEnt, with the other two models overestimating their occurrences. As such, the combined maps were built using the habitat suitability layer of MaxEnt for corals, but the weighted averages for the other three species assemblages (Figure 3.7). The predictive maps based on coarsened backscatter data (10m resolution) are presented in supporting information 3.8.4, and show an even lower ability to capture the variation observed in the biological imagery. Model predictions for rarer species assemblages, such as those associated with cold-

water corals and hard substratum, also showed decreases in averaged niche similarities of 0.07 and 0.03 respectively.
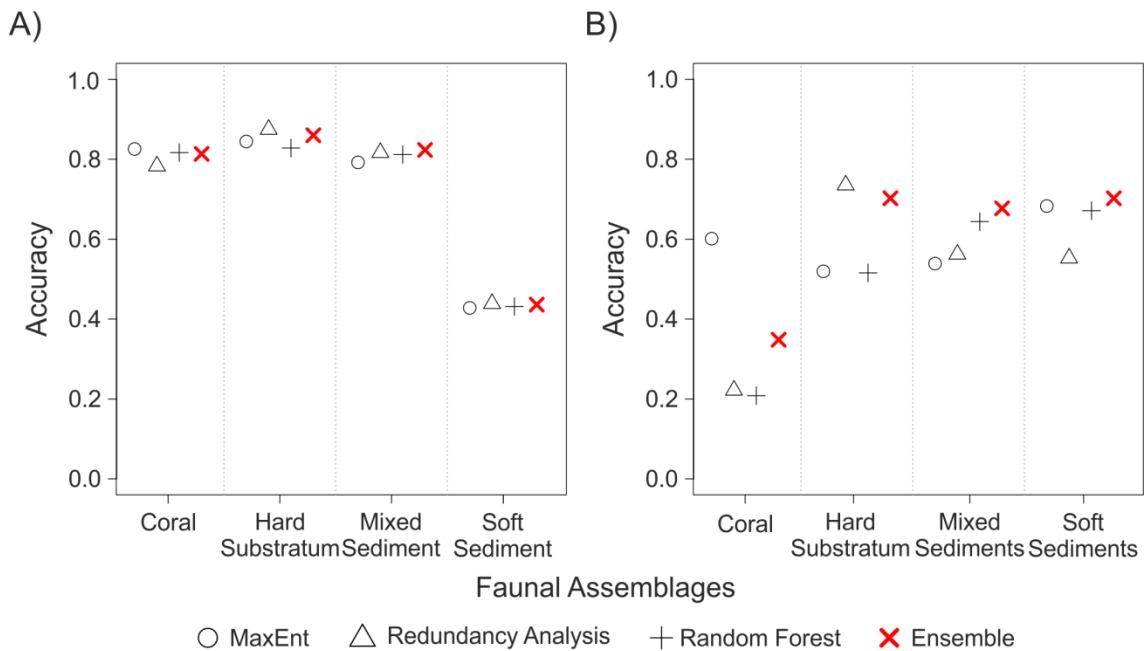
Figure 3.6: Comparison of accuracy values obtained between single statistical models and an ensemble model for four different species assemblages. A) Predictions at points removed from the *JC*-060 data and B) using the *JC*-060 data to make predictions at locations visited during the *JC*-073 cruise. Accuracy defined as the proportion of correctly assigned presences and absences over total sample size.

3.5 Discussion

Compared to terrestrial ecosystems, biological data from the deep sea tend to be particularly difficult and expensive to collect, resulting in smaller datasets with reduced options for sampling designs. As such, very thorough evaluation of predictive models ought to be carried out, as with lower quality datasets one would expect lower predictive abilities and higher levels of uncertainty (Guisan *et al.* 2007). Despite the difficulties associated with the collection of datasets in deeper marine environments, this study was able to procure a second dataset from the area originally surveyed in order to examine predictions made for unsampled locations.
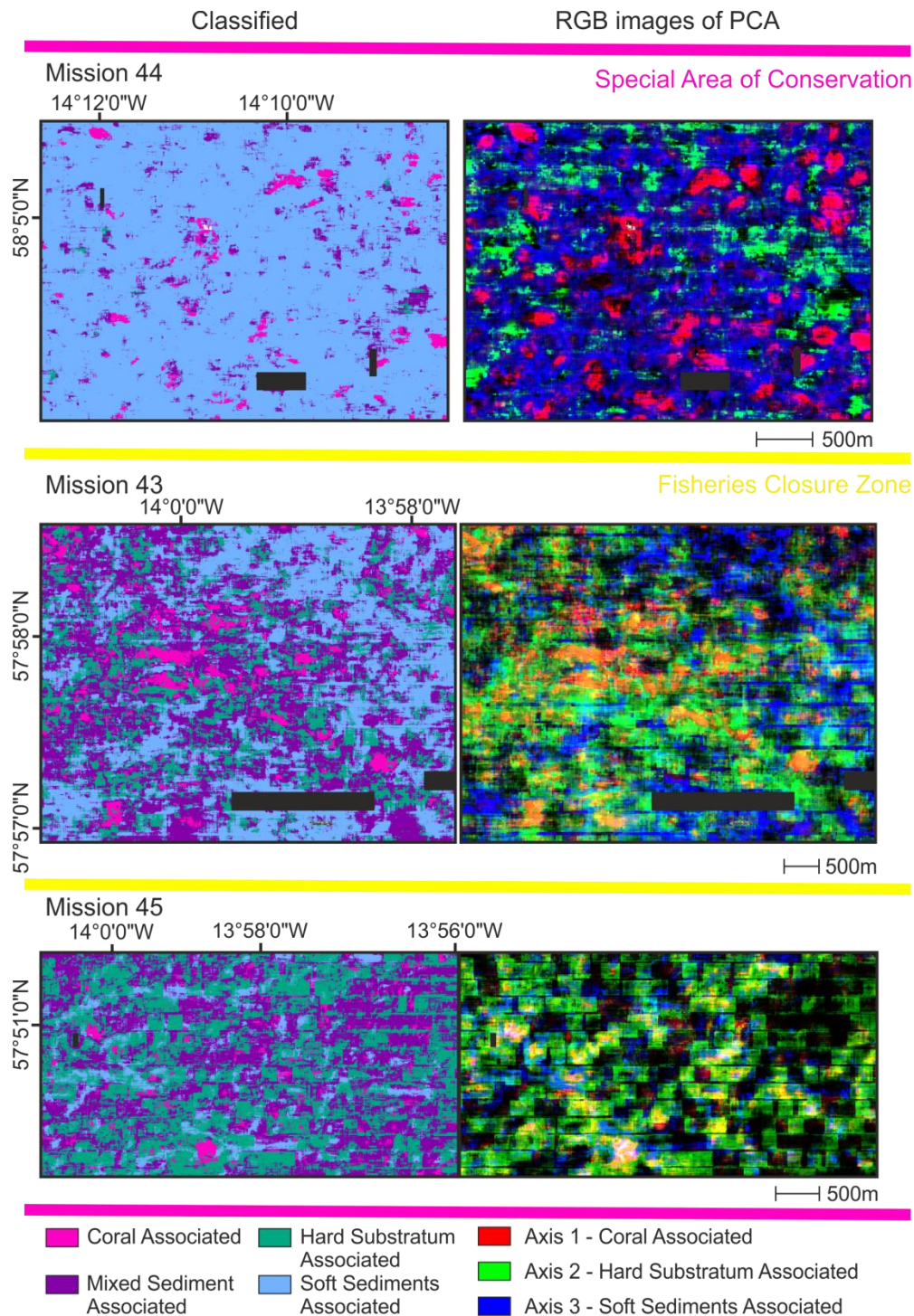
Figure 3.7: Maps showing the classification into species assemblages for three different areas around two conservation zone boundaries on Rockall Bank based on three statistical techniques. Hard classification based on the highest suitability from weighted averages taking into account 10 partitions for three statistical models (left) and an RGB image based on the weighted average predictions. Colours on the RGB image show a gradient of habitat suitability for three of the species assemblages, black indicates areas for which habitat suitability was low for the three assemblages represented.

3.5.1   Model Predictions

Of the three models (RDA, RF and MaxEnt) for species assemblage predictions compared in this study, similar split-sample AUC values were obtained, but performance for each model varied for different species assemblages. As species turnover generally occurs over a gradient, the predictions showed a similar pattern and overlap between habitat suitability predictions occurred, particularly between coral and hard substratum associated fauna. This is to be expected as cold-water corals need hard substratum for attachment (Freiwald et al. 1999) and in turn provide hard substratum to a number of species. The mixed sediment associated faunal assemblage appeared as a transition between the more defined hard substratum and soft sediment associated fauna, and as such prediction performance for this assemblage generally tended to be lower. Larger areas were predicted as suitable by either RDA or RF than obtained using MaxEnt for most species assemblages.

As additional datasets to test model performances at new survey locations are rarely available, one is often obliged to stop after this initial assessment phase. Under these conditions, it would be very difficult to determine which of the model considered may be most adequate. However, if a second independent dataset is not available, our results suggest that taking into account the output of many different models may provide more robust predictions and that a resampling approach allows for the creation of prediction variability maps.

3.5.2   Confounding Factors

Differences in imagery systems can lead to disparities in the size of the field of view, while lower resolutions could affect the minimum organism size discernible and the amount of detail visible for species identification. The extracted frames from the *Holland I* ROV HD video of 2012 were of lower resolution (5 pixels cm^{-1} versus 18 pixels cm^{-1} in line with the scaling lasers) but provided a wider field of view than the photographs acquired with the *Lynx* ROV in 2011. However, in both cases lasers were present to rescale transect width, and a minimal organism size of 1cm had been set. Although the *JC*-073 transects had a lower biodiversity, the majority of species had been observed

during the previous cruise and species assemblages were consistent between the two cruises, suggesting that temporal and instruments variability was limited. Measures ensuring similar acquisition rates, area sampled and minimal size of organism recorded should reduce the risks of sampling biases between datasets, but the use of differing systems should not be a reason to reject systematically comparisons, as limited access to survey equipment and technological improvements are unlikely to make other options easily available.

Precise spatial positioning can also be problematic for underwater vehicles, particularly AUVs operating in deeper settings, owing to difficulties associated with determining the initial position following the descent and correcting for drift (McPhail 2009). Owing to the high seafloor heterogeneity of Rockall Bank, differences in position of less than 10m could lead to very different assemblages being observed. However, as our study site was located at only ~220m water depth, limited drift of the AUV during descent would be expected. In addition, accuracy of the ROV's Sonardyne USBL navigation is expected to be better than 1% of the depth, i.e. better than 2.2m. With the landscape and class metrics calculated using moving windows of 60m and 150m, a small shift in ROV position would have had limited effects on the values of the explanatory variables.

The spatial extent for which predictions can be valid is also of importance, as predictions made for areas outside the range of environmental conditions captured by the survey design are problematic (Elith & Leathwick 2009, Menke et al. 2009). The use of transects limited the area surveyed to single narrow lines leaving most of the regions covered acoustically without any biological sampling. However, the additional 2012 transects were located only 0.2km and 1.3km away from previously collected imagery, while the *JC*-060 transects spanned a distance of up to 15km between transects.

The seabed of this area is of such high heterogeneity that the inclusion of substratum percentage cover estimates derived from imagery nearly doubled the percentages of variation explained by the RDA and linear regression (Robert et al. 2014) . As this level of detail was not available for the spatial extent of the predictive maps, the complexity of the species-environment relationships was likely still underestimated. Seabed variability

was particularly high in the M43 area and may be part of the reason why predictions for the *JC*-073 data were generally poorer. Lower prediction success was also observed when transects 91 or 93 were removed from the *JC*-060 dataset.

Transects are designed to maximise seafloor survey areal coverage for a given bottom time, but as a consequence can cause issues of spatial autocorrelation which may not be adequately captured using a split-data procedure and can lead to predictive ability being overestimated (Hirzel & Guisan 2002, Legendre et al. 2002). Analysis of the spatial autocorrelation of the present dataset showed that spatial structuring in environmental variables accounted for most of the spatial structure detected in the fauna, and that it was occurring over a relatively short range of 20-50m (Robert et al. 2014). When fine-scale percent substratum cover was available, only 5.9% and 2.0% of the spatial structure detected, for species assemblages and biodiversity (H') respectively, was not accounted for by the environmental variables. However, when only lower resolution sonar information was available these percentages raised to 15.1% and 11.0%. This combination of an increase in unaccounted spatial autocorrelation and unavailability of finer-scale substratum information is likely the main reason why none of the models considered could adequately represent the very fine-scale heterogeneity of Rockall Bank.

3.5.3   Ensemble Mapping for Coral Conservation

As each of the modelling approaches was based on different statistical approaches with different assumptions, each model possessed different strengths and weaknesses, leading to different areas of uncertainties. Using an ensemble modelling approach, where the models were combined based on weights resulting from model performance and prediction variability, the different modelling strengths were optimized to capture trends in the spatial distribution of species assemblages more reliably. As expected from what is known of the biology of cold-water corals in the region (Wilson 1979b), coral assemblages were predicted in areas characterized by hard substratum, while biodiversity appeared highest in areas dominated by hard substratum or in proximity to cold-water corals. In all cases, the use of a categorical classification, although useful for model assessment and management

purposes, clearly under-represented the complexity that could be observed in the RGB depictions (Figure 3.7).

However, the results obtained from the ensemble maps showed trends at a scale applicable for conservation and management. Although all models considered gave different maps at the very fine scale (<1m), all models produced the same answer to the question of which conservation zone boundary would be most appropriate for cold-water coral protection. Consistent with what had been observed in the imagery, coral stands were predicted throughout the M45 area, an area which would be protected only by the cSAC, in the meantime established as a 'Site of Community Importance' (http://jncc.defra.gov.uk/). Corals were also predicted within the M44 area, but the ROV imagery only found coral rubble to remain in the area, likely the result of past trawling activities. Similarly, lower biodiversity was also generally observed in that area. Even though, it is the broader-scale patterns in species distribution which may be of interest for management purposes, it is the fine-scale habitat characterisation of the environment, through high-resolution mapping, that allowed the heterogeneity of the region to be accurately captured. However, these results are based on only 37km^2 of high-resolution surveys out of a 4,365km^2 conservation zone and the uncertainty maps clearly highlight a high degree of uncertainty.

3.6 Conclusion

Predictive habitat maps can be of great use for marine management as they visually represent the best available information, but, as they are typically based on a very limited amount of available data, they should only serve as general guides until much more data become available. The presentation of uncertainty maps should help emphasize this point. If possible, model predictions and selection of proposed areas for spatial protection should then be validated using additional survey work to evaluate directly the conservation value of such areas, and to provide larger datasets for use in improving model selection and model predictions. Finally, as our current validation process is unlikely to change, owing to the difficulties associated with additional data collection in deeper areas, it is important to remember that model assessment

might be over-optimistic and that undue confidence should not be accorded to predictive modelling results without a thorough exploration of model uncertainties, and preferably some effort to validate prediction results with survey data. However, ensemble mapping approaches may help mitigate this effect by allowing the strength of different modelling approaches to be combined.

3.7 Acknowledgments

3.8 Supporting Information

3.8.1 Example R Code

```
library (raster); library (SDMTools); library (snow); library (snowfall); library (rlecuyer)

Matrix<- raster("NameOfRaster")
window.size<- 60 # Side length (in pixel number) for a square neighbourhood
res<-res (Matrix)[1]
n.rows<- nrow (Matrix)
n.cols<- ncol (Matrix)
n.cell<- ncell (Matrix)
S <- matrix (seq (1, n.cell, 1), ncol = n.cols, nrow = n.rows) # Numbers all pixels
V <- as.vector (S [(window.size+1) : (n.rows-window.size ),
    (window.size+1) : (n.cols-window.size)])
Matrix<- as.matrix(Matrix)

sfInit(parallel=TRUE, cpus=8) # Sets up R to run in parallel using 8 CPUs
sfExport ("Matrix", "V", "n.rows", "n.cell", "res" )  # Export all the necessary data to all
                                        #the slaves
sfExport ("window.size")
sfLibrary (SDMTools)
sfLibrary (rlecuyer)
sfClusterSetupRNG ()

wrapper<- function (cell) {  # Wrapper function to calculate class metrics for each pixel
n.rows<-n.rows
j <- floor (cell / n.rows) + 1
i <- cell - (floor (cell / n.rows) * n.rows)
window.size<- window.size
    b <- Matrix[(i - window.size) :  (i + window.size),
        (j - window.size) : (j + window.size)]
metrics<- ClassStat (b, cellsize = res)
return (metrics)
gc ()
}

result<- sfLapply (V, wrapper) # Parallel version of lapply
sfStop () #Stops the clusters
save (result, file = "NameOfResultFile.RData")
```

3.8.2 MDS Plot by Transects
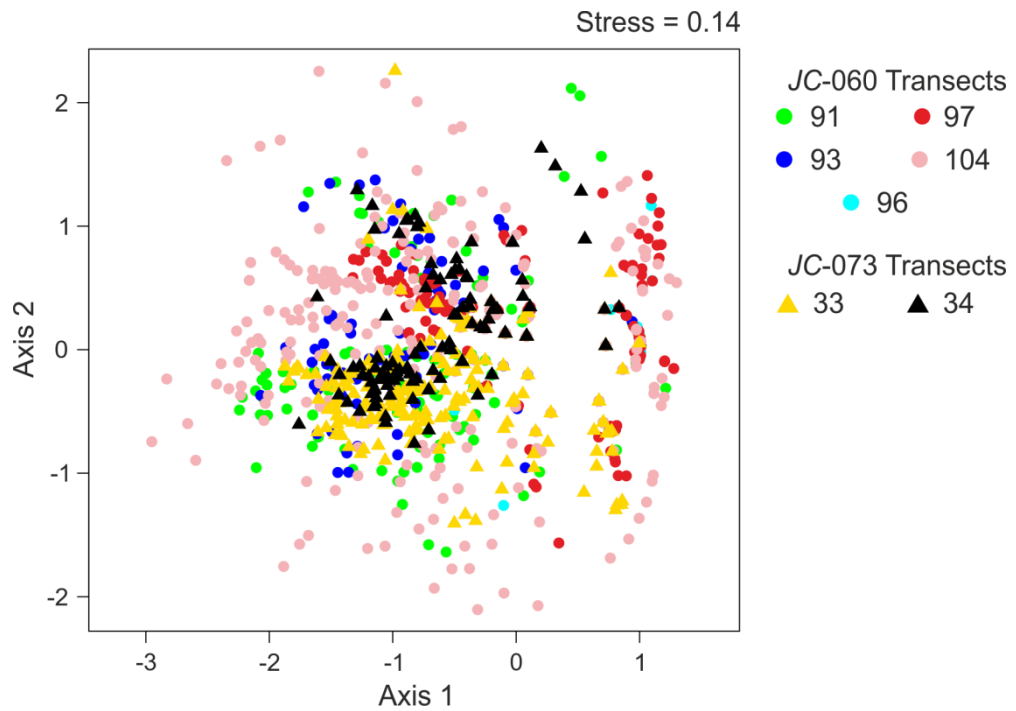
Figure S3.1: MDS plot showing no differences in species composition by transect or between cruises *JC*-060 and *JC*-073.

3.8.3 Prediction Maps by Species Assemblage

Hard Substratum Associated
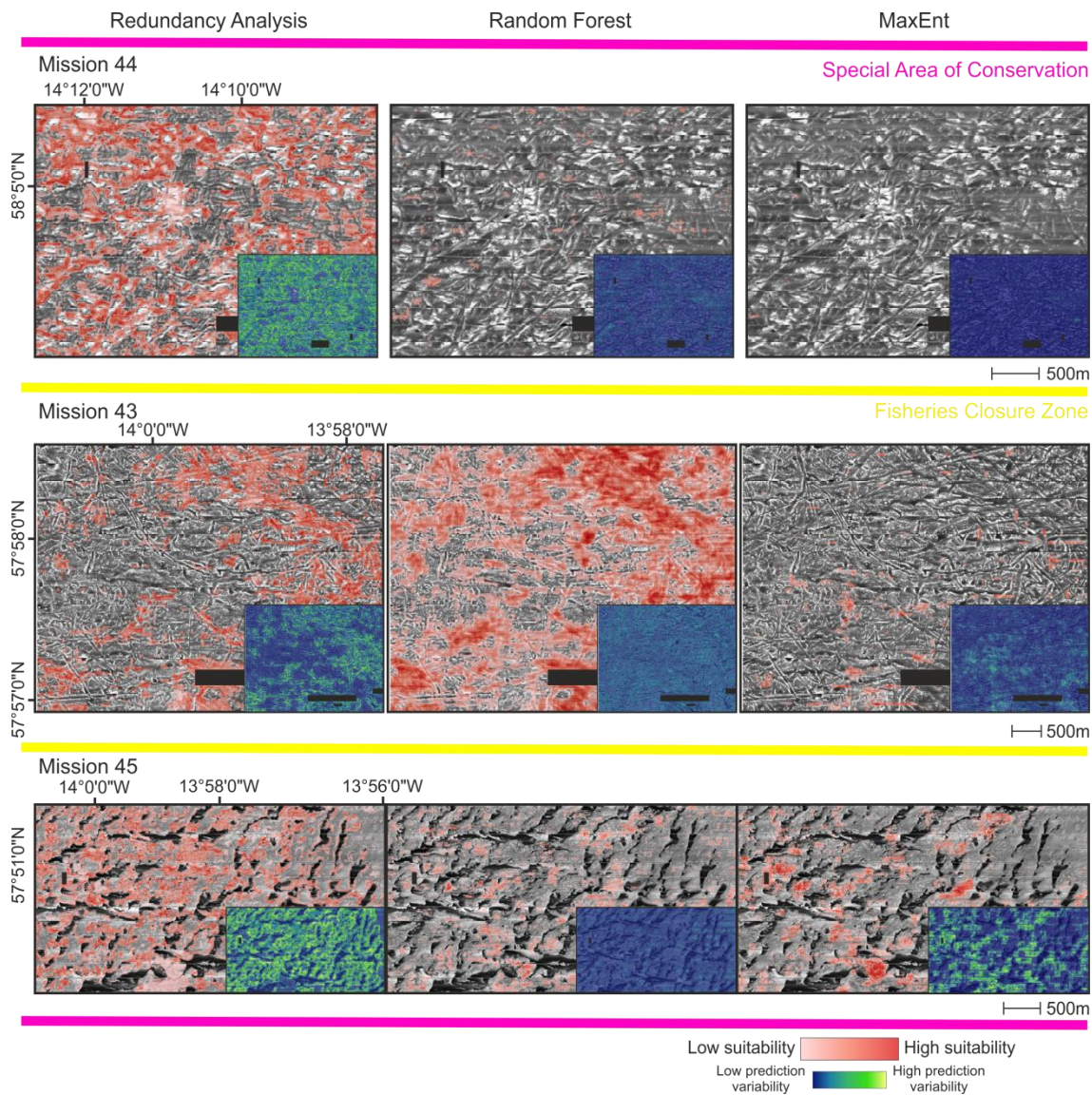
Figure S3.2: Maps showing the relative habitat suitability of the hard substratum associated fauna for three different areas around two conservation zone boundaries on Rockall Bank based on three statistical techniques. Smaller insets show standard deviation of predictions based on 10 holdout partitions.

Mixed Sediment Associated
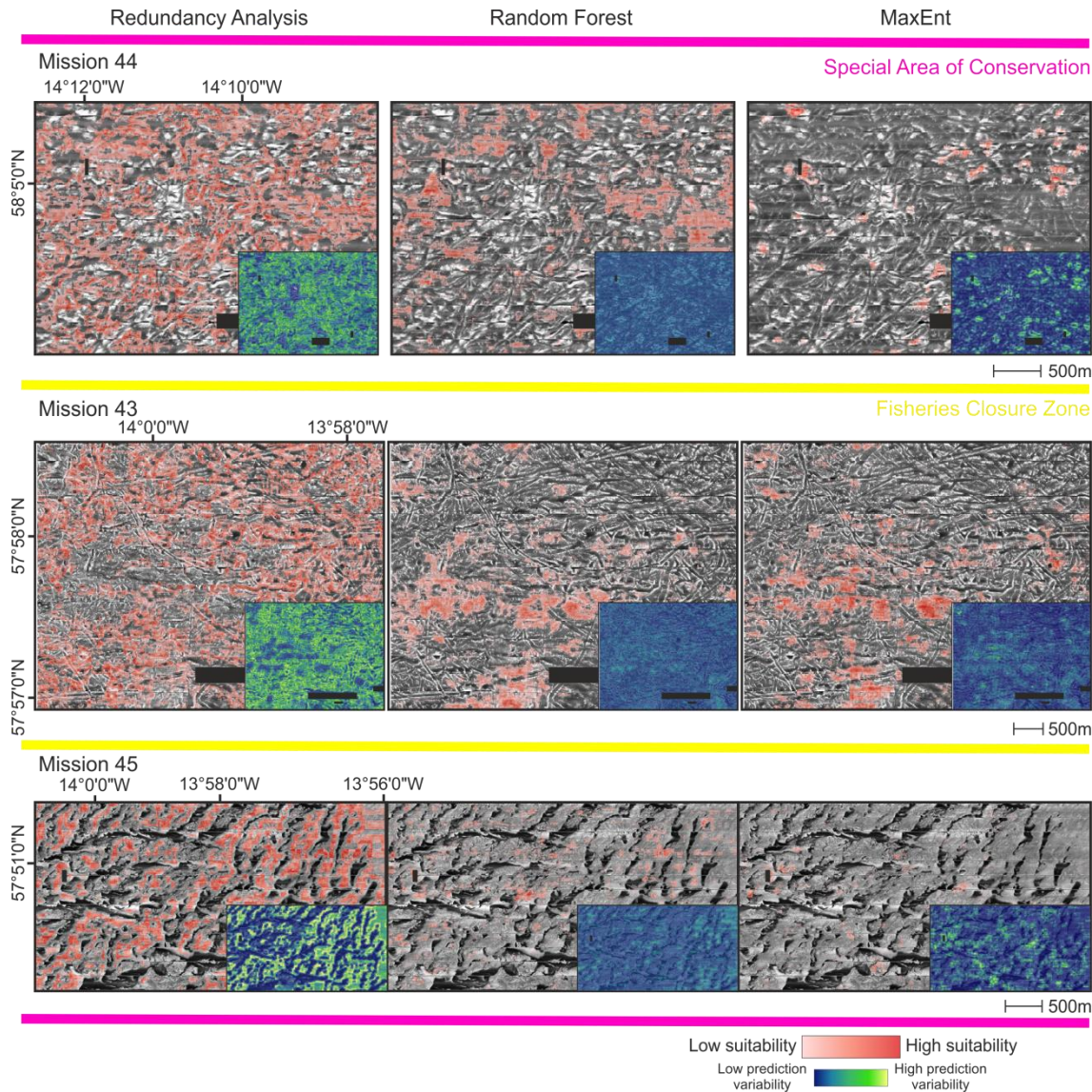
Figure S3.3: Maps showing the relative habitat suitability of the mixed sediment associated fauna for three different areas around two conservation zone boundaries on Rockall Bank based on three statistical techniques. Smaller insets show standard deviation of predictions based on 10 holdout partitions.

Soft Sediment Associated
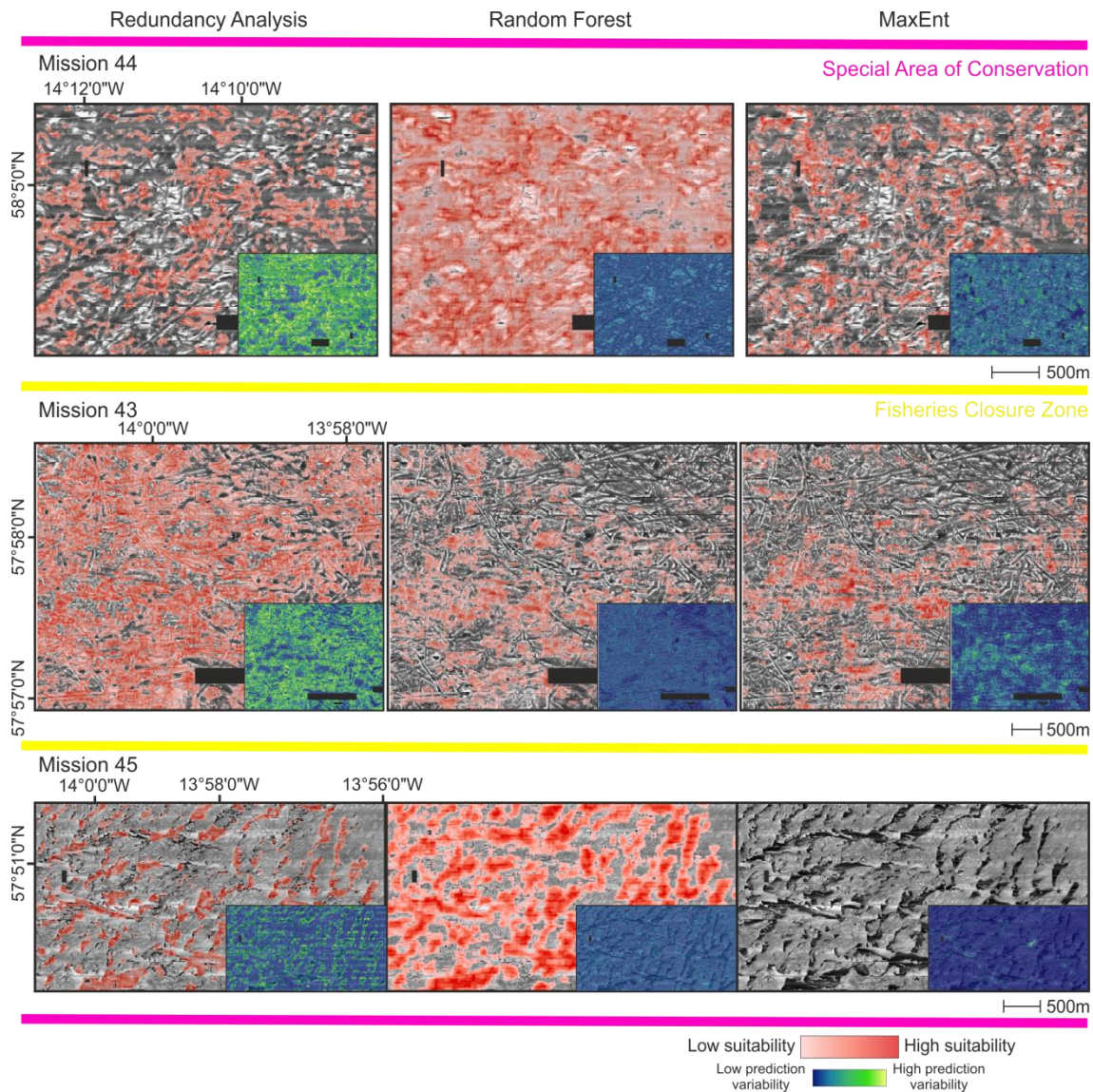
Figure S3.4: Maps showing the relative habitat suitability of the soft sediment associated fauna for three different areas around two conservation zone boundaries on Rockall Bank based on three statistical techniques. Smaller insets show standard deviation of predictions based on 10 holdout partitions.

3.8.4 Prediction Maps at Coarser Resolutions
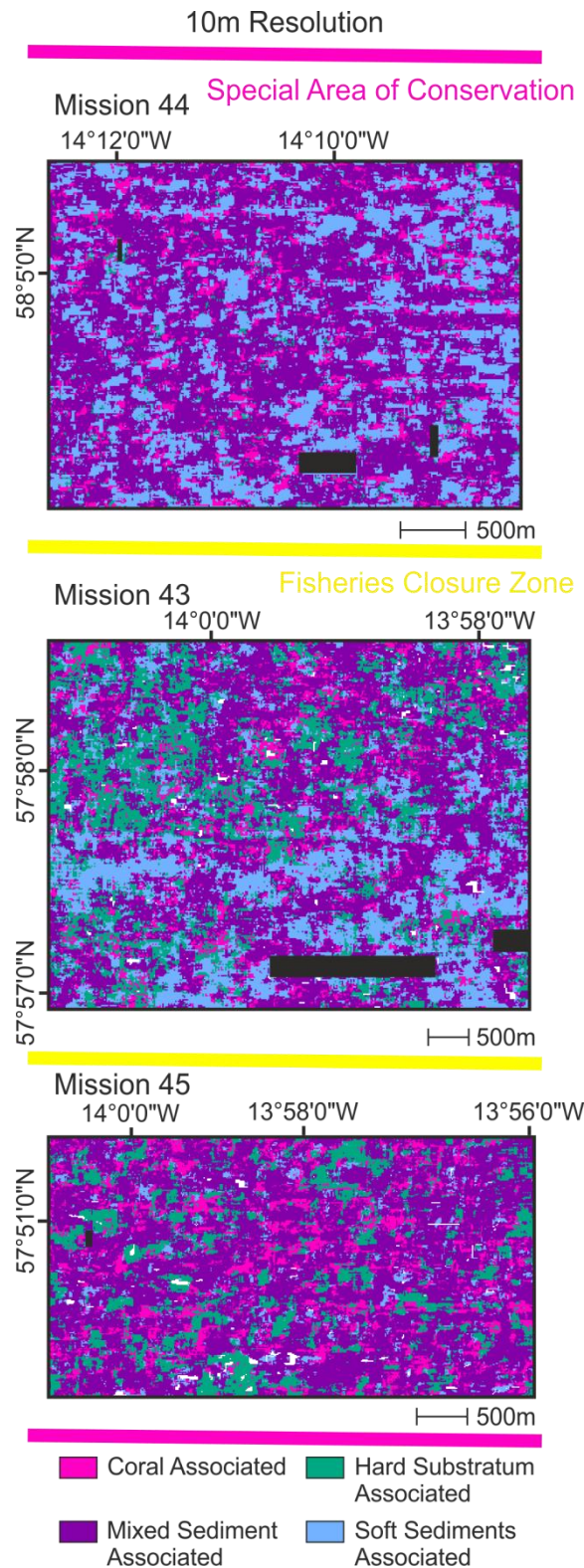
Figure S3.5: Maps showing classification into species assemblages for coarsened data (10m) for three different areas around two conservation zone boundaries on Rockall Bank based on weighted averages taking into account three statistical techniques.

Chapter 4:
Finding the hot-spots within a biodiversity hotspot:
Fine-scale biological predictions within
a submarine canyon using high-resolution
acoustic mapping technique

Katleen Robert[1], Daniel O.B. Jones[2], Paul A. Tyler[1], David Van Rooij[3], and Veerle A.I. Huvenne[2]

[1] School of Ocean and Earth Science, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, UK

[2] National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

[3] Ghent University, Department of Geology and Soil Science, Gent, Belgium

4.1   Abstract

Submarine canyons are complex geomorphological features that have been suggested as potential hotspots for biodiversity. However, few canyons have been mapped and studied at high resolution (10's m). In this study, the four main branches of Whittard Canyon, Northeast Atlantic, were mapped using multibeam and sidescan sonars to examine which environmental variables were most useful in predicting regions of higher biodiversity. The acoustic maps obtained were ground truthed by 13 remotely operated vehicle (ROV) video transects at depths ranging from 650 to 4,000m. Over 100 hours of video were collected, and used to identify and georeference megabenthic invertebrate species present within specific areas of the canyon. Both general additive models (GAMs) and random forests (RF) were used to build predictive maps for megafaunal abundance, species richness and biodiversity. Vertical walls had the highest diversity of organisms, particularly when colonized by cold-water corals such as *Lophelia pertusa* and *Solenosmilia variabilis*. GAMs and RF gave different predictive maps and external assessment of predictions indicated that the most adequate technique varied based on the response variable considered. By using ensemble mapping approaches results from more than one model were combined to identify vertical walls most likely to harbour a high biodiversity of organisms or cold-water corals. Such vertical structures were estimated to represent less than 0.1% of the canyon's surface. The approach developed provides a cost-effective strategy to facilitate the location of rare biological communities of conservation importance and guide further sampling efforts to help ensure that appropriate monitoring can be implemented.

4.2   Introduction

By comparison to other regions of the continental slope, submarine canyons have been proposed as hotspots for biodiversity. Incising the continental shelf, they are typically characterised by high spatial heterogeneity, complex hydrographic patterns and can act as conduit for larvae and organic matter from the shelf to the deep sea (Vetter & Dayton 1999, Tyler et al. 2009, Vetter et al. 2010). In addition to increasing habitat heterogeneity at the regional scale (De Leo *et al.* 2010), submarine canyons also exhibit high habitat heterogeneity at the local scale (Huvenne *et al.* 2011) which can result in further finer-scale variations in biodiversity. Near vertical walls within submarine canyons of the Bay of Biscay with particularly high percentage cover of biological growth (particularly cold-water corals, *Lophelia pertusa*, limid bivalves, *Acesta excavata*, and deep-water oysters, *Neopycnodonte zibrowii*) have been observed (Van Rooij et al. 2010a, Huvenne et al. 2011, Johnson et al. 2013). The authors suggest that, although rare and spatially limited, these structures may provide critical functions as refugia for certain species against anthropogenic impacts such as trawling. As fisheries are often associated with the head of submarine canyons (Morell 2007, Puig et al. 2012), it is important to investigate which environmental factors within submarine canyons are responsible for local increases and decreases in biodiversity so that appropriate management and monitoring can be implemented.

Establishing biological spatial patterns is a crucial step in informing conservation measures. However, this is particularly problematic in the deep sea where biological samples are sparse and spatially limited. When investigating broad-scale submarine features characterised by a high spatial heterogeneity, only small sections of seafloor may actually harbour high biodiversity or specific species of interest. Without prior knowledge of their likely distributions at high resolutions, important areas may be missed. As full coverage biological sampling is usually not feasible, techniques for producing full coverage predictive maps are being developed. These techniques are based on finding relationships between biological patterns and environmental descriptors derived from high-resolution acoustic maps (Brown *et al.* 2011). Environmental descriptors such as slope, rugosity, aspect (orientation of steepest slope), bathymetric position index (BPI, measure of the relative height

of a pixel in comparison to surrounding pixels) and curvature can be derived from bathymetric maps (Wilson *et al.* 2007), while sidescan sonar or multibeam backscatter can be used as a proxy for sediment hardness (Lo Iacono et al. 2008, Micallef et al. 2012). Additional descriptors from sidescan sonar backscatter, such as image texture indices (e.g. homogeneity and entropy), skewness or kurtosis, can also be derived to help identify seabed composition (Huvenne et al. 2007, Blondel & Gómez Sichi 2009, Isachenko et al. in press). Indices specific to canyon morphology, used to compare cross-section profiles, have also been successfully applied (De Leo et al. 2014). As acoustic maps can cover larger extents of seabed much faster than traditional biological sampling techniques (e.g. photographs or grabs), predictive maps covering entire regions can be achieved even though only limited biological information may be available (McArthur *et al.* 2010). These predictions can act as proxies for biological information until further sampling is possible.

Cold-water corals are one of the vulnerable marine ecosystems of interest as biogenic reefs can maintain particularly high biodiversity both across the reef itself as well as in comparison to the surrounding seafloor (Freiwald et al. 2004, Costello et al. 2005, Buhl-Mortensen et al. 2010, Henry et al. 2010). Submarine canyons have been proposed as potentially suitable habitats for cold-water corals as their heterogeneous seabed provides exposed hard substratum for attachment while canyon morphology creates more complex hydrographic patterns, which may be beneficial for filter-feeders (Mortensen & Buhl-Mortensen 2005, White et al. 2005, Orejas et al. 2009, Edinger et al. 2011). Cold-water corals are also particularly vulnerable to trawling (Fosså et al. 2002, Hall-Spencer et al. 2002, Roberts et al. 2008) and as such, much interest exists for mapping their distributions. Global (Tittensor et al. 2009, Davies & Guinotte 2011) and large-scale studies of the Northeast Atlantic (Ross & Howell 2012) have been successful in creating habitat suitability maps, but higher resolution predictive maps are still needed to target specific areas where cold-water corals are likely to occur.

Many studies have found acoustically-derived environmental descriptors to be useful in explaining biological distribution patterns and predicting habitat suitability (Dolan et al. 2008, Mortensen et al. 2009, Monk et al. 2010). However, few deep-sea studies have employed fine-scale species-environment

relationships to build full coverage predictive maps over the full extent of a submarine canyon. In this study, we used environmental descriptors derived from multibeam bathymetry and sidescan sonar backscatter to build full coverage predictive maps for biological characteristics of epibenthic megafauna (including abundance, species richness, biodiversity and cold-water coral presence) across four branches of the Whittard Canyon, Northeast Atlantic. Both 'general additive models' (GAMs) and 'regressive random forest' (RF) are used, and in addition to a split-sample assessment, model predictions are further tested using an externally acquired dataset. We also created ensemble maps of the results obtained to identify likely areas for which future sampling might be valuable, particularly with respect to vertical structures.

4.3   Methods

4.3.1   Acoustic Surveys

During the 2009 RRS *James Cook* -035 cruise, both multibeam ($6130km^2$, 50x50m resolution) and sidescan sonar ($3800km^2$, 3x3m resolution) surveys were conducted to map the Whittard Canyon, Northeast Atlantic. This submarine canyon is located east of the Goban Spur and links the continental margin of the Celtic Sea (200m) to the Celtic Fan and Porcupine Abyssal Plain (4,000m) (Figure 4.1). Whittard Canyon is a dendritic system composed of four main branches joining at around 3,500m water depth to form a single wider channel. During the last glacial period, it was part of the drainage system of a paleovalley (Bourillet et al. 2003, Toucanne et al. 2008), but its activity is now much reduced owing to its distance from the present-day shoreline (Reid & Hamilton 1990).

The acoustic surveys were conducted using a shipboard EM120 multibeam system and the 'Towed Ocean Bottom Instrument' (TOBI) mounted with a 30kHz sidescan sonar. The bathymetry, processed using the CARIS HIPS & SIPS software suite to a 50m resolution grid, (WGS1984, UTM Zone 29N) was used to derive additional environmental layers (50m resolution) such as slope, standard deviation of slope, aspect (orientation of steepest slope split into two
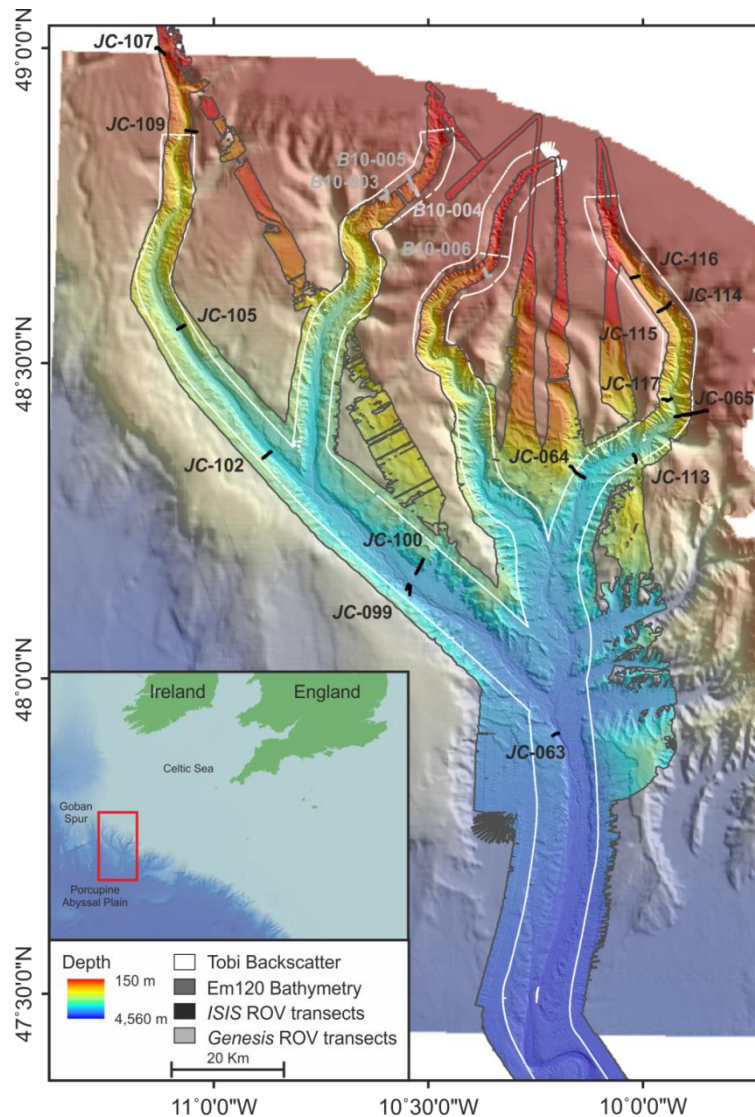
Figure 4.1: Map of Whittard Canyon and surveys carried out during the *JC*-010, -035 -036 and *Belgica* 10/17b cruises . Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).

continuous measures: eastness and northness) and curvature (general, plan and profile), based on 3x3 pixel size windows. Surface area ratio and bathymetric position index (BPI, using neighbourhood sizes of 150m, 500m, 1km and 2km) were also computed (Figure 4.2). Flow direction and length layers (downstream to an outlet or sink and upstream from the highest upslope basin point) were also derived as calculating the direction and distance along the flow path of a watershed may provide a proxy for transport within the canyon. Layers were generated in ArcMap 10 using the 'Spatial Analyst Extension' as well as the 'Land Facet Corridor Tools' and the 'DEM Surface Tools' developed by Jenness Enterprises (Jenness 2012a, b). The TOBI sidescan
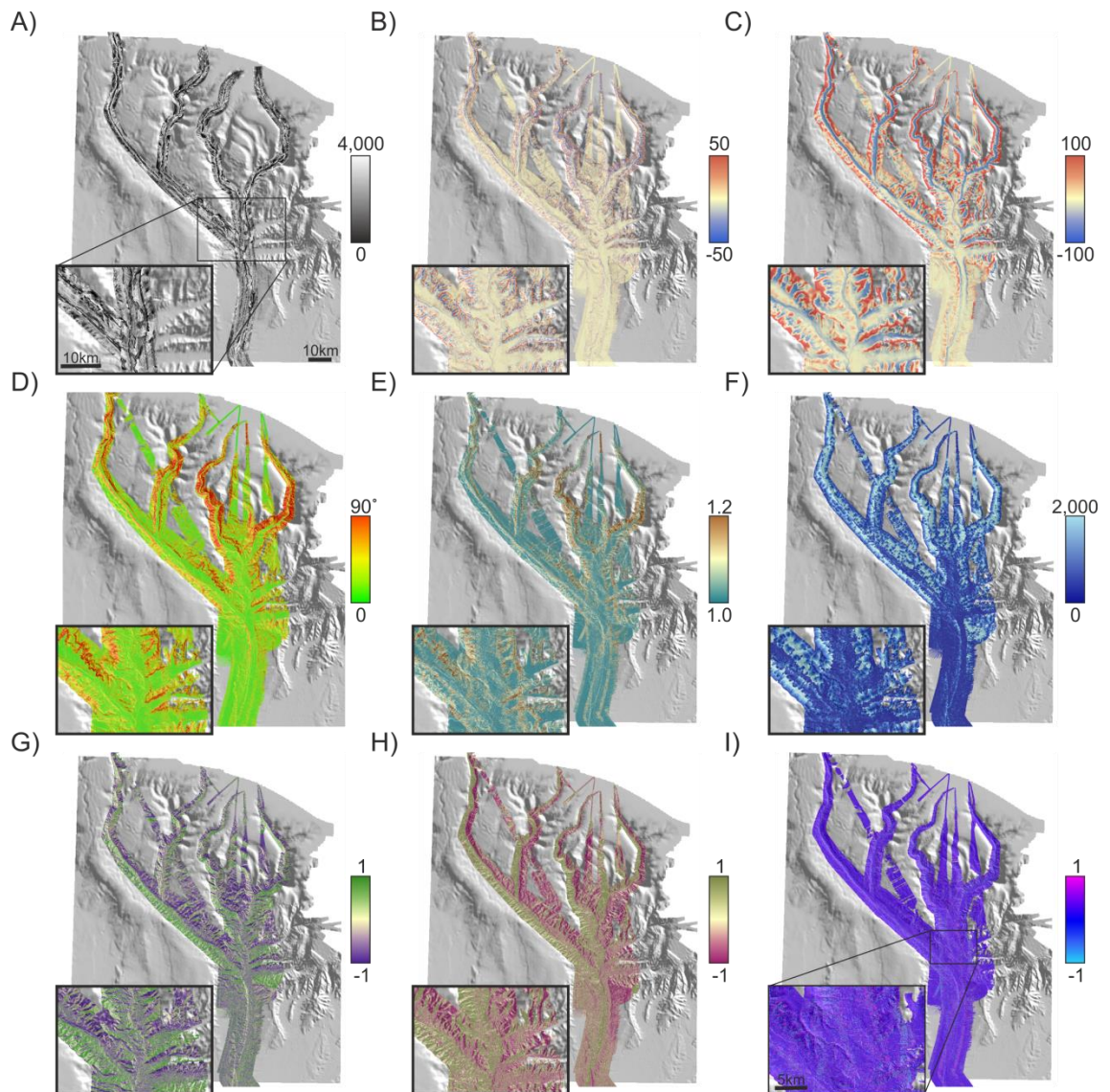
Figure 4.2: Environmental variables (50m resolution) derived from the acoustic survey : A) sidescan backscatter (low backscatter in dark), B) bathymetric position index (BPI) with neighbourhood of 500m, C) BPI with neighbourhood of 2km, D) slope (˚), E) surface area ratio, F) downstream flow length, G) northness, H) eastness and I) profile curvature. Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).

sonar data were originally processed to a 3x3m resolution grid using the in-house PRISM software suite (Le Bas & Hühnerbach 1998). A 'true slant range' correction (whereby the bathymetry is used instead of an assumed flat seabed) (Le Bas & Huvenne 2009) and an across-track equalisation of illumination on an equal range basis were applied during processing and values were normalized to a pixel value of 1,000. The higher resolution layer was aggregated to 50m resolutions by using the mean value in order to conform to the multibeam

derived layers.   Sidescan sonar backscatter was added to provide information on seabed type, high backscatter values are indicative of harder substratum types, while lower values result from soft sediment absorbing the acoustic signal.

4.3.2   Benthic Imagery

To complement the acoustic maps, 10 video transects were conducted during the *JC*-036 cruise using the work-class *ISIS* remotely operated vehicle (ROV). An additional 3 transects (dives *JC*-063, -064 and -065) previously collected in 2007 during the *JC*-010 cruise were also added to the benthic imagery dataset. These transects were divided across the western and eastern branches of the canyon at depths ranging from 650m to 4,000m (Table 4.1). With the ROV moving at an average speed of ~0.08m s^{-1} and an average height of 3m off the seafloor, the down-looking colour video camera (Pegasus, Insite Tritech Inc. with SeaArc2 400W, Deepsea Power&Light illumination) was recorded on digital tapes and later converted to .mov using a Sony digital HD videocassette recorder. In cases where a vertical section of seabed was encountered, video footage from the forward-looking wide-angle camera (Atlas, Insite Tritech Inc.) was examined instead. Over 100 hours of video were watched to identify and georeference megabenthic invertebrate organisms larger than 1cm. As identification to species level could not always be achieved, morphospecies (visually distinct taxa) were used for the analysis. High-resolution stills (Scorpio, Insite Tritech Inc., 2048x1536 pixels) and specimens were also collected to help with identifications. The position of each individual animal on the seafloor was determined using the ROV's ultra-short baseline navigation system (USBL). Substratum type (soft, hard or mixed sediments as well as live coral or rubble) and terrain inclination gradient (flat, sloping, vertical or complex) were also recorded. Distances travelled within each of these patches were measured. Two lasers representing 10cm on the seabed were present for scaling. A frame was extracted every minute (representing a 5m displacement) using QuickTime 7 Pro (Apple Inc.) and the distance separating the scaling lasers was measured in the image processing software Image J (http://rsbweb.nih.gov/ij/). These measurements were used to standardize transect widths to 2.5m.

Table 4.1: Description and biological information for the benthic video transects collected using the remotely operated vehicle *ISIS* and the ROV *Genesis* for external model assessment. For biological characteristics, averages based on 50 m sections are shown in parenthesis.

Cruise	Dive	Branch	2D Length (km)	3D Length (km)	Avg Depth (m)	SD Depth (m)	Average Slope (°)	Video Duration HH:MM:SS	No. Individuals	Species Richness	Reciprocal Simpson Diversity
JC-010	063	Lower	1.60	1.68	-3,868.7	120.0	28.7	05:19:22	352 (10.3)	17 (2.32)	2.55 (1.74)
JC-010	064	Eastern	4.13	4.39	-2,959.7	341.4	32.5	08:25:35	1,386 (15.7)	21 (1.35)	3.23 (0.96)
JC-010	065	Eastern	7.17	7.17	-1,432.3	690.5	45.3	14:30:36	13,134 (85.8)	73 (5.08)	6.32 (2.23)
JC-036	099	Western	3.15	3.24	-3,638.3	254.0	17.8	09:55:43	125 (1.9)	13 (1.24)	1.77 (1.18)
JC-036	100	Western	2.93	3.01	-3,396.2	389.0	19.4	08:58:18	325 (5.3)	35 (3.26)	9.15 (2.68)
JC-036	102	Western	1.94	2.04	-3,095.9	232.5	26.7	06:46:06	335 (8.2)	24 (2.85)	4.34 (1.81)
JC-036	105	Western	1.60	1.73	-2,550.1	521.6	36.6	07:00:30	2,095 (56.6)	41 (3.89)	6.75 (2.01)
JC-036	107	Western	2.45	2.45	-778.2	170.4	27.3	09:49:48	1,350 (26.5)	27 (2.67)	3.03 (1.80)
JC-036	109	Western	2.09	2.09	-1,123.6	243.6	46.4	08:30:52	1,287 (28.5)	60 (6.89)	8.36 (3.41)
JC-036	113	Eastern	2.09	2.35	-2,334.2	1,189.4	47.5	09:08:10	2,270 (48.3)	33 (2.79)	2.64 (1.47)
JC-036	114	Eastern	1.50	1.64	-1,462.1	164.4	39.0	11:21:54	9,571 (290.0)	67 (9.97)	1.54 (2.68)
JC-036	115	Eastern	1.65	1.69	-1,409.5	414.7	19.7	06:07:30	235 (6.9)	27 (2.32)	8.83 (1.76)
JC-036	116	Eastern	2.08	2.38	-1,235.4	186.6	52.3	08:59:23	7,399 (160.8)	52 (6.28)	4.24 (2.5)
JC-036	117	Eastern	2.34	2.67	-2,199.2	188.3	51.9	11:00:46	3,414 (63.2)	49 (4.85)	3.87 (2.21)
Be-10/17b	003	Inner	3.89	4.03	-821.8	105.6	16.0	03:57:28	11,856 (146.4)	24 (5)	2.04 (1.81)
Be-10/17b	004	Inner	2.45	2.47	-922.8	83.2	9.5	02:24:32	9,563 (191.3)	37 (5.48)	3.03 (2.11)
Be-10/17b	005	Inner	1.39	1.42	-943.6	79.9	13.2	00:51:08	1,352 (48.3)	14 (4.93)	2.9 (2.30)
Be-10/17b	006	Inner	1.88	1.99	-746.3	186.5	19.0	01:57:34	2,251 (56.3)	25 (2.93)	1.9 (1.42)

Transects were subdivided into segments of 50m in length and the species records consolidated. However, as the topography of Whittard Canyon varied greatly, the distance between the geographic coordinates of two points would not accurately depict their three dimensional separation. The vertical dimension was taken into account in ArcMap 10 by creating a 3D surface from the bathymetry using the 'Add Surface Information' tool in the '3D Analyst Toolbox'. This tool calculates the distance between two points based on the topography of a surface. Using the bathymetry raster as the input surface, the transects were separated into 50m sections which took into account seabed topography. Differences between 2D and 3D transects length are reported in Table 4.1. Abundance, species richness, and the reciprocal of Simpson's index (1/D, Simpson 1949), were calculated for each transect and 50m section. This index was chosen because it is more sensitive to changes in dominant species, while the importance of rare species is captured by species richness (Hill 1973). Components of Beta diversity, *a'* (percentage of species occurring in both a focal and neighbour sample), *b'* (percentage of species occurring in focal sample but not in neighbour sample) and *c'* (percentage of species occurring in neighbour sample but not in focal one) were also calculated between ROV dives (Koleff *et al.* 2003).

4.3.3　Predictive Modelling

4.3.3.1　General Additive Models (GAMs)

'General additive models' (GAMs) are similar to 'general linear models' (GLMs) in that they allow the building of models with different error structures and link functions, but they are not limited to modelling relationships for which the form is known *a priori* (Crawley 2007). Instead, they make use of non-parametric smoothers, such as regression splines or tensor products, which allows for the building of more complex (e.g. non-linear, non-monotonic) relationships (Guisan *et al.* 2002). GAMs have the potential to explain additional variation, as compared to GLMs. However, consideration must be given to the appropriate level of smoothing to avoid over-fitting (Wood & Augustin 2002).

GAMs were used to build predictive maps for each of the biological characteristics calculated (abundance, species richness, reciprocal Simpson index and cold-water coral presence). Abundance was log(x+1) transformed prior to modelling to improve normality and was modelled using normally distributed errors. The reciprocal Simpson index was also modelled using normally distributed errors. In the case of species richness, since the data was only composed of positive integers, a Poisson distribution was used, while presence-absence of cold-water coral colonies was modelled using a binomial distribution. Absences were assumed when no corals were observed in a 50m section of video; however, it is important to note that corals could have been present in the remaining unsampled area of the given pixel. For coral presence, only colonial, framework-building scleractinians were considered (*Madrepora oculata*, *Lophelia pertusa* and *Solenosmilia variabilis*), as they can be considered habitat-forming species. Environmental variables were assessed by forward selection, the variables resulting in the highest deviance explained were added one step at a time until no more statistically significant (P-Value <0.05) variables could be added. The significance of the addition was assessed by comparing the reduction in deviance caused by the additional variable in comparison to the previous model using the χ^2 statistics (Guisan *et al.* 2002). Using the most parsimonious models, full coverage maps were created by predicting values for each pixel of the bathymetry. To avoid rescaling non-linear relationships, such as diversity-area, the pixels show the

expected value for a section of seabed of the same area as sampled, i.e. 2.5x50m or 125m^2. A measure of prediction variability was generated by mapping the standard error associated with the expected value for each pixel.

Spatial correlograms, using Moran's I index of spatial autocorrelation (Legendre & Fortin 1989), of the 50m sections dataset suggested that most of the spatial autocorrelation present occurred at scales <100m, except for cold-water coral presence which occurs at <50m (results shown in supporting information, 4.8.1). As such, the 50m sections were systematically subsampled to create distances of at least 100m, resulting in half the dataset being considered for model building (similar result were obtained using either set of subsamples), except in the case of coral presence where subsampling was not carried out. The 50m sections not used for model building were subsequently employed for model assessment. Statistical analyses were done using the statistical package R (R Development Core Team 2014) and the library 'mgcv'.

4.3.3.2   Regressive Random Forest

'Random forest' (RF) is a technique whereby multiple decision trees are built based on random subsamples of the data and the environmental predictors, leading to the construction of a 'forest' (Breiman 2001). Trees are built by binary splits where the data are recursively separated into smaller and smaller groupings based on the best predictor variable available. As each tree will be built based on a different set of samples and environmental predictors, each tree will be different and once grown can be used to make predictions based on the rules developed at each node (Cutler *et al.* 2007). Each tree provides an answer and the average (in the regressive case) is the expected value.

For each of the biological characteristics measured, a total of 1,000 trees were built with 12 variables being randomly selected at each node. Full coverage maps were obtained by making predictions for each pixel of the bathymetry and uncertainty maps were calculated by taking the standard deviation of the 1,000 predicted values (one from each tree) for each pixel. Variable importance was measured using the out-of-bag (OOB) data, the part of the randomly subsampled data not used for building the tree. As values for

the OOB data are known, a measure of accuracy (the mean squared error) can be obtained by comparing the known values to the estimates obtained when these samples are regressed along the trees (Hastie *et al.* 2009). Each predictor variable is then permuted, the accuracy measure recomputed and the mean difference averaged over all trees. The same dataset as for the GAMs was used and functions from the R package 'randomForest' were employed.

4.3.4   Model Comparison

The performance of each model was assessed by calculating the amount of variation explained when the predicted values were regressed against the known values from the 50m sections which had not been used to build the models. The root-mean-square error (RMSE) was also calculated to give a more easily comparable measure between the observed and predicted values (Knudby *et al.* 2010). However, as previous work suggested that split-sample assessment methods may yield overly optimistic results (Araújo et al. 2005, Randin et al. 2006), comparisons to an independently collected dataset were also carried out. This dataset was comprised of four ROV video transects collected using the 1,400m-rated inspection-class sub-Atlantic Cherokee ROV *Genesis* on board the RV *Belgica* during the 2010 cruise 10/17b. These transects (*B*10-03, *B*10-04, *B*10-05 and *B*10-06) were acquired at an average speed of ~0.3m s^{-1} using 4 video cameras (including forward looking) and a digital Canon Powershot colour stills camera (250W Q-LED illumination) and were located in the two inner branches of the canyon (Figure 4.1). Two parallel laser beams with a distance of 10cm were used as a scale during seabed observations, and the seabed positioning was recorded using the IXSEA USBL GAPS system. These transects were analysed in the same manner as the ROV *ISIS* video transects. Using subsections of 50m, the results obtained were regressed against the values predicted by the different models and the RMSE was calculated.

4.3.5   Vertical Structures

As an association between vertical walls and high biodiversity was known for Whittard Canyon (Huvenne et al. 2011, Johnson et al. 2013), areas with slopes >35˚ were selected and polygons representing individual steep vertical walls were created. Owing to the depths surveyed, the coarser ship-board

bathymetry resolution causes slope values to be underestimated. Based on the imagery available, slopes >35° could potentially represent near-vertical walls (Huvenne *et al.* 2011). Pixels predicted to be in the top 90% for at least 4 of the maps created for abundance, species richness or reciprocal Simpson index for either GAMs or RF were identified. These selected pixels represented areas likely to harbour diverse communities with high abundances. The number of such pixels found within each polygon was calculated and standardized based on the surface area of the polygon. The same was carried out for pixels which showed high suitability for presence of corals (0.6 for GAMs and 0.8 for RF, values chosen by inspecting the histogram).

4.4   Results

A wide variety of seabed environments were observed within Whittard Canyon (Figure 4.3) with a median substratum patch size <100m and beta diversity measures indicating that only a relatively small percentage of species (<40%) tended to be shared between transects (results are shown in supporting information 4.8.2). The high turnover in species assemblages is one of the main reasons this study focused on biological characteristics as opposed to specific species assemblages. A total of 42,934 individuals and 202 morphospecies were observed in the initial *JC* survey, with the most commonly observed taxa (representing ~60% of individuals) being xenophyophores (likely *Syringammina fragilissima*), *Pentametrocrinus* sp., *Acanella* sp., *Lophelia pertusa*, cerianthids and *Anthomastus* sp. The eastern and western branch harboured similar numbers of species, but abundance was much greater in the eastern branch (on average 82 individuals per 50m section compared to 17). A greater percentage of the western branch was composed of flat soft sediment (82% versus 53%) and deposit feeders were more common in this branch. In shallower (<1,000m) and deeper transects (>3,800m) cerianthids dominated the observations. Holothurians and ophiuroids increased in relative abundance at 2,700-3,700m depth, with the former composing a greater percentage of observations in the western branch. Soft corals, particularly *Acanella* sp. (or possibly *Chrysogorgia* sp.) and *Anthomastus* sp., were predominant at 1,500-3,500m in depth, particularly in more morphologically complex areas. Stalked crinoids appeared more frequent on hard bottoms, while *Pentametrocrinus* sp.

was frequently observed in flat areas. Xenophyophores composed a large
portion of observations in flat soft sediments at 1,200m.
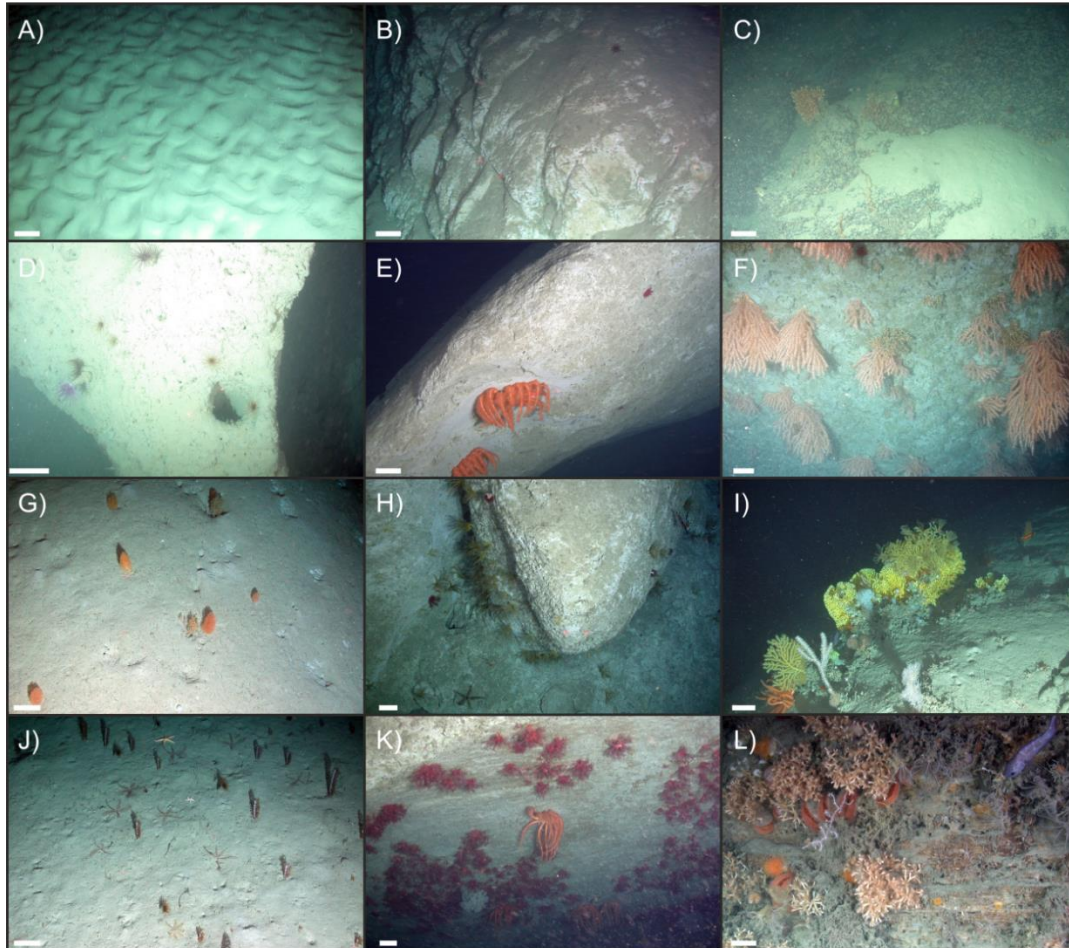
Figure 4.3: Example images of environments encountered and organisms observed : A) hummocky
sediment, B) rocky outcrop, C) exposed bedrock with single coral colony, D) burrowed wall with
Cerianthids, E) steep gully with Brisingids, F) vertical wall with *Primnoa* sp.and *Solenosmilia variabilis*, G)
soft sediments with *Acanella* sp. H) rocky outcrop with Cormatulids, I) Stylasterid coral and associated
community, J) soft sediments with Pennatulacea and *Pentametrocrinus* sp, K) vertical wall with
Anthomastus sp and Brisingids, and L) vertical wall with *Lophelia pertusa*, *Acesta* sp., *Actinauge* sp. and
Comatulids. Scale bars show 10cm.

In total 58 morphospecies and 25,022 individuals were observed in the
four transects of the *Belgica* cruise 10/17b, of which 8 morphospecies had not
been recorded in the previous survey. In these transects, the most commonly
observed morphospecies were cerianthids, *Kophobelemnon* sp., and *Madrepora
oculata* (representing ~80% of individuals). Although the majority of the
morphospecies observed in the *Belgica* transects had been observed during the

previous cruises, their relative composition differed. Soft corals, possibly owing to the shallower transect depths, were not as frequently observed throughout the inner canyon branches. Instead echinoids (*Cidaris cidaris* and *Phormosoma placenta*), sea pens (*Kophobelemnon* sp.) and cerianthids were more prevalent. Large aggregations of sea pens (particularly *Kophobelemnon* sp., but also *Pennatula aculeata*) were observed, at depths of 800-900m and 900-1,000m respectively, in areas dominated by soft sediments. Sea pen meadows composed of *Pennatula* sp. have been reported for submarine canyons of the western Atlantic (Baker *et al.* 2012), and *Kophobelemnon* sp. in other canyons of the Bay of Biscay (Davies et al. 2014). Few instances of areas dominated by mixed or hard sediments were recorded for the *Belgica* transects, but areas dominated by coral rubble or live corals in flat areas were more prevalent. Detailed species composition by dive, substratum type and depth band are presented in supporting information 4.8.3, 4.8.4 and 4.8.5. When available, a representative image for each observed morphospecies was deposited in the online media archive of the SERPENT project (http://archive.serpentproject.com/view/sites/sea_14.html).

Transect lengths differed between dives (Table 4.1), with dive *JC*-065 being the longest and the one with the overall highest abundance and species richness. However, the highest diversity at the transect level was observed in dives *JC*-100, *JC*-115 and *JC*-109. When averages based on the 50m sections were compared instead, dives *JC*-114 and *JC*-116 had the highest abundance, species richness (with dive *JC*-109) and diversity as measures by the reciprocal Simpson index (with dives *JC*-100 and *JC*-109). Lowest abundance and species richness was found in dive *JC*-099 with low diversities (1/D), based on averaged 50m sections, in dive *JC*-064. With dive *B*10-05, these transects represented those having the highest percentages of soft sediments in flat areas.

In both GAMs and RF, depth was the most important environmental variable for predicting abundance (Figure 4.4 and 4.5). Although a general decrease in abundance with depth was observed, peaks occurred at ~1,200, 2,200m, 3,000m and 3,700m. Coarse scale BPI (2km) indicated that lower abundances could be expected in flat regions and a similar trend was observed for curvature. Steeper slopes were expected to have higher abundances and

although not selected by GAMs, RF also found surface area ratio to be a useful predictor. Downstream flow length indicated a general decrease in abundance towards the canyon thalweg. Differences in prediction extent between certain models were caused by the smaller extent of the TOBI backscatter layer.

Depth and slope were also significant predictors of species richness in both GAMs and RF. An overall negative relationship with depth was observed, with the highest species richness recorded at 1,200m, while steep slopes (>20°) had a positive effect (Figure 4.4). Higher species richness was characteristic of regions with high standard deviations of slope, but a lower
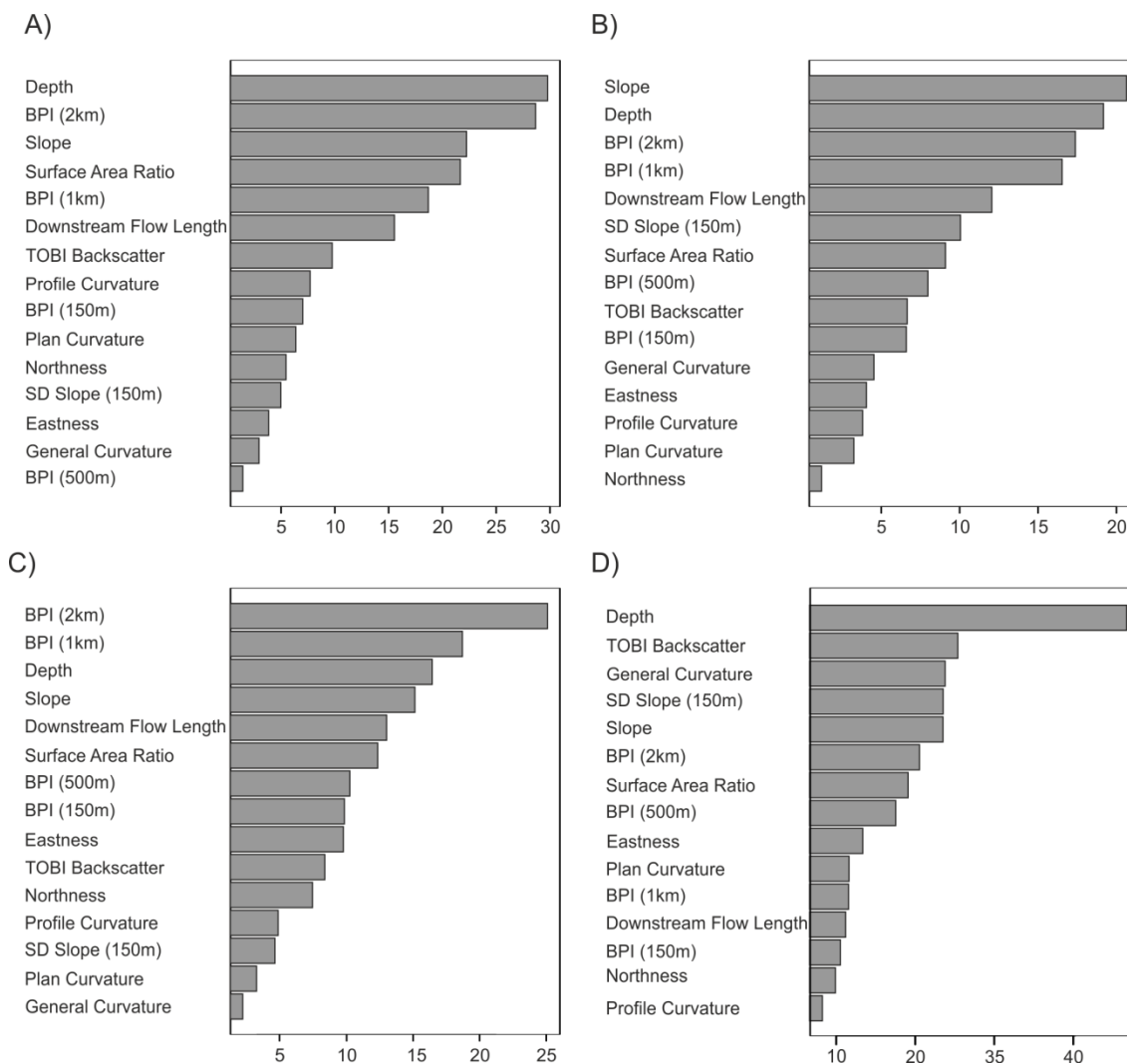
Figure 4.4: Variable importance for the random forest models for A) abundance, B) species richness, C) reciprocal Simpson index and D) coral presence. Variable importance is reported as the percentage increase in mean squared errors.
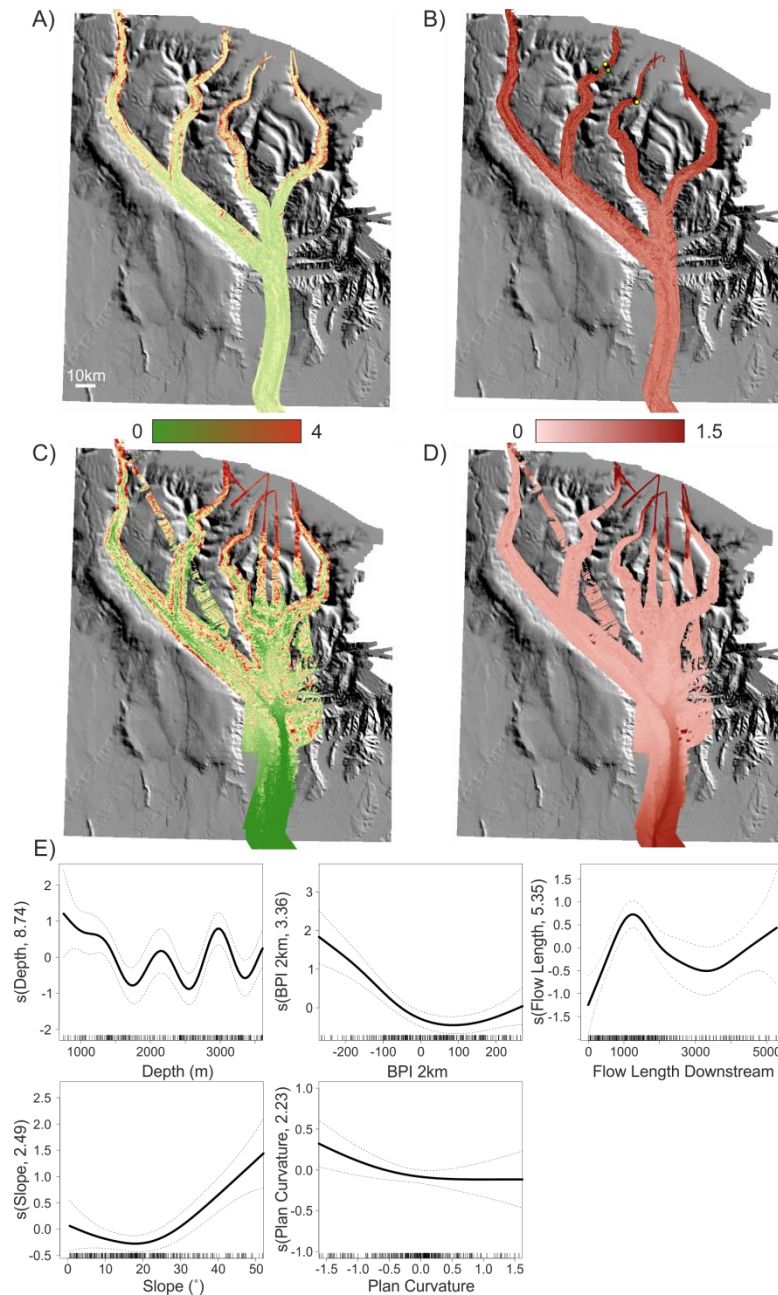
Figure 4.5: Model output for predicting log abundances within a 125m^2 section of seabed across the extent of the multibeam survey: A) random forest prediction map and B) uncertainty map (SD), C) general additive model prediction map and D) uncertainty map (SE). Abundance prediction map represent Log(x+1) transformed data. E) Relationship (centred smooth component, black line, and estimated degrees of freedoms in parenthesis) for the selected environmental variables using GAMs with the dashed lines representing the standard error estimated for the predictions. Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).

number of species was expected on strongly east or west facing slopes. Coarse scale BPIs (1km and 2km) were good predictors in both GAMs and RF, and indicated lower species richness in flat areas, a trend also supported by
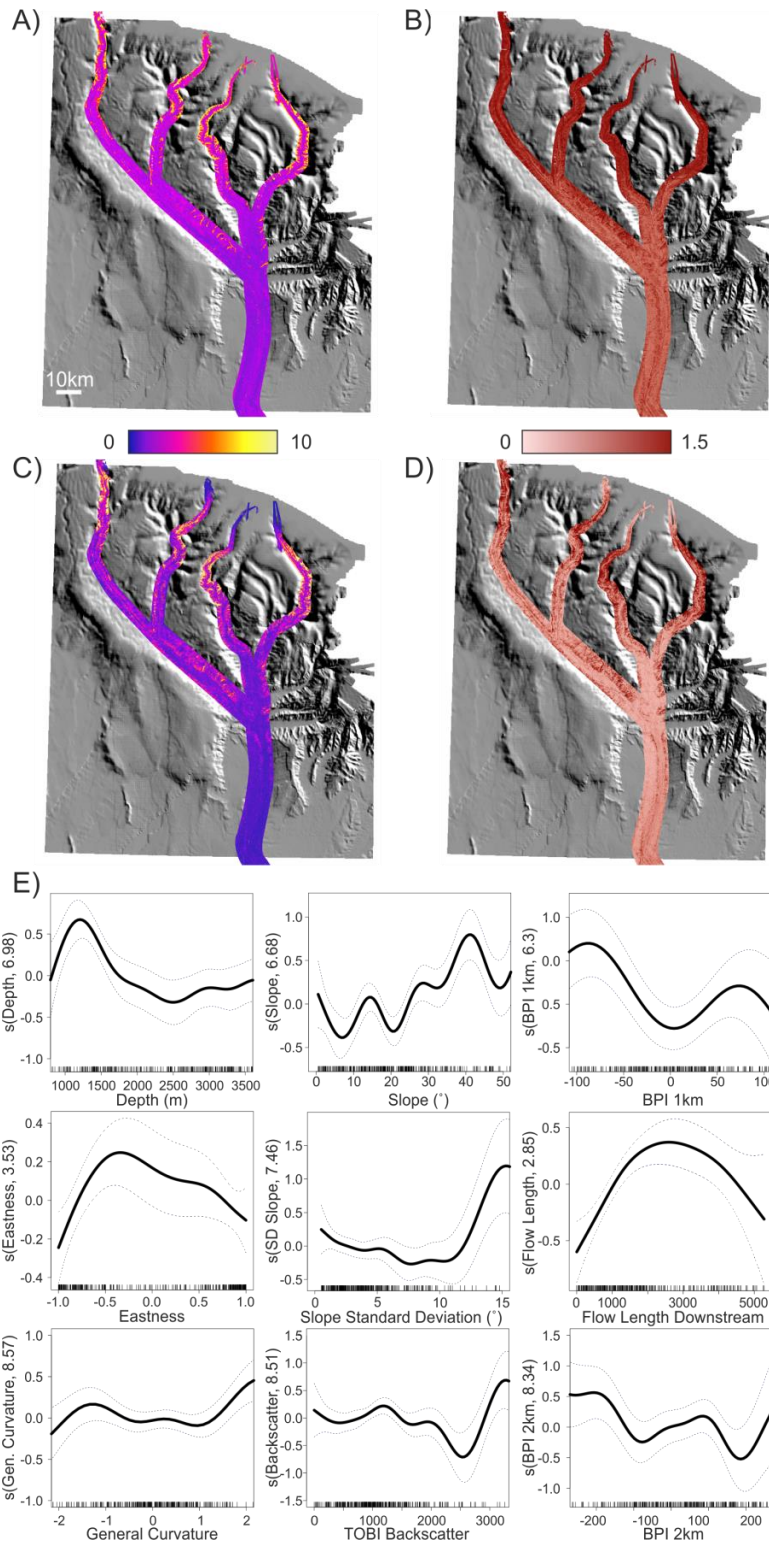
Figure 4.6: Model output for predicting species richness within a 125m² section of seabed across the extent of the multibeam survey: A) random forest prediction map and B) uncertainty map (SD), C) general additive model prediction map and D) uncertainty map (SE). E) Relationship (centred smooth component, black line, and estimated degrees of freedoms in parenthesis) for the selected environmental variables using GAMs with the dashed lines representing the standard error estimated for the predictions. Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).

the relationship observed with general curvature. The canyon thalweg also harboured lower species richness as indicated by the relationship with the downstream flow length and the decrease in species richness in low backscatter areas.

For the reciprocal Simpson index, coarse scale BPIs (2km or 1km) were more significant than either depth or slope in both GAMs and RF (Figure 4.4 and 4.7). As for abundance and species richness, lower biodiversity was found in flat areas and a positive relationship to slope was observed. Although not selected by GAMs, RF also found depth to be important in predicting biodiversity. The relationship to plan curvature suggests that very fine-scale ridges may harbour higher diversity (1/D) than valleys.

Three species of colony forming cold-water corals were observed: *Madrepora oculata* (39 colonies), *Lophelia pertusa* (3,337 colonies) and *Solenosmilia variabilis* (383 colonies). Dive *JC*-116 in the eastern branch, imaged the 120m high coral wall mapped by Huvenne et al. (2011), while a smaller, sparser and less diverse coral wall was also found in the western branch during dive *JC*-109. Presence of cold-water coral was predicted in association with finer-scale variations in seabed morphology, with fine-scale BPI (150m) being the first predictor selected by GAMs (Figure 4.8). RF also found fine-scale descriptors to be of greater importance than for the previous biological characteristics, with the standard deviation of slope and general curvature as most useful (Figure 4.4). Cold-water corals were more likely to occur in areas of complex topography with higher rugosity as opposed to flatter areas. This was also apparent at the coarser scale as shown by the BPI (2km) and the sharp decrease in coral presence towards the canyon thalweg. RF also highlighted the importance of depth and slope. Both RF and GAMs found TOBI backscatter to be a useful predictor of cold-water presence, the relationship identified by GAMs indicates a higher likelihood of coral presence in areas characterised by higher TOBI backscatter which is indicative of hard substratum.
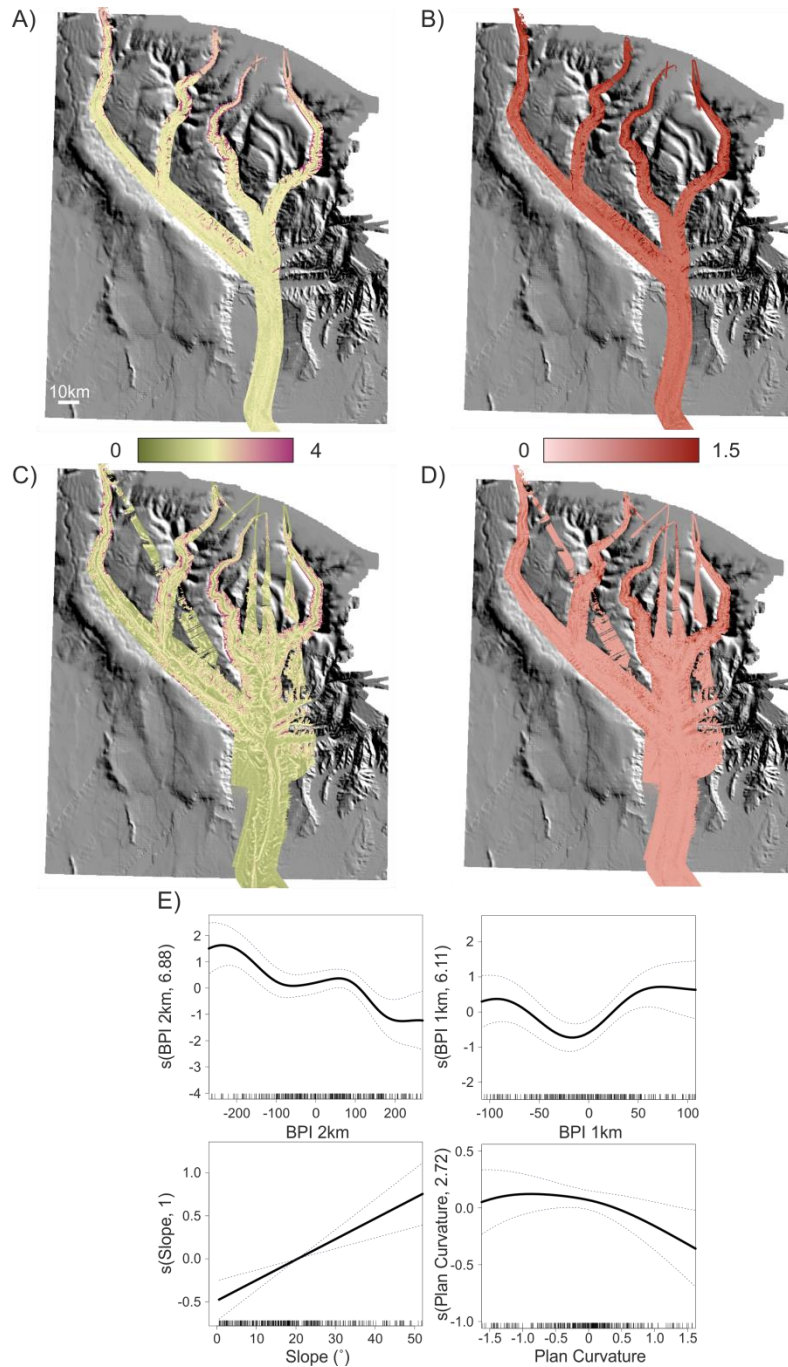
Figure 4.7: Model output for predicting the reciprocal Simpson index (1/D) within a 125m^2 section of seabed across the extent of the multibeam survey: A) random forest prediction map and B) uncertainty map (SD), C) general additive model prediction map and D) uncertainty map (SE). E) Relationship (centred smooth component, black line, and estimated degrees of freedoms in parenthesis) for the selected environmental variables using GAMs with the dashed lines representing the standard error estimated for the predictions. Background bathymetry (201m resolution) Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).
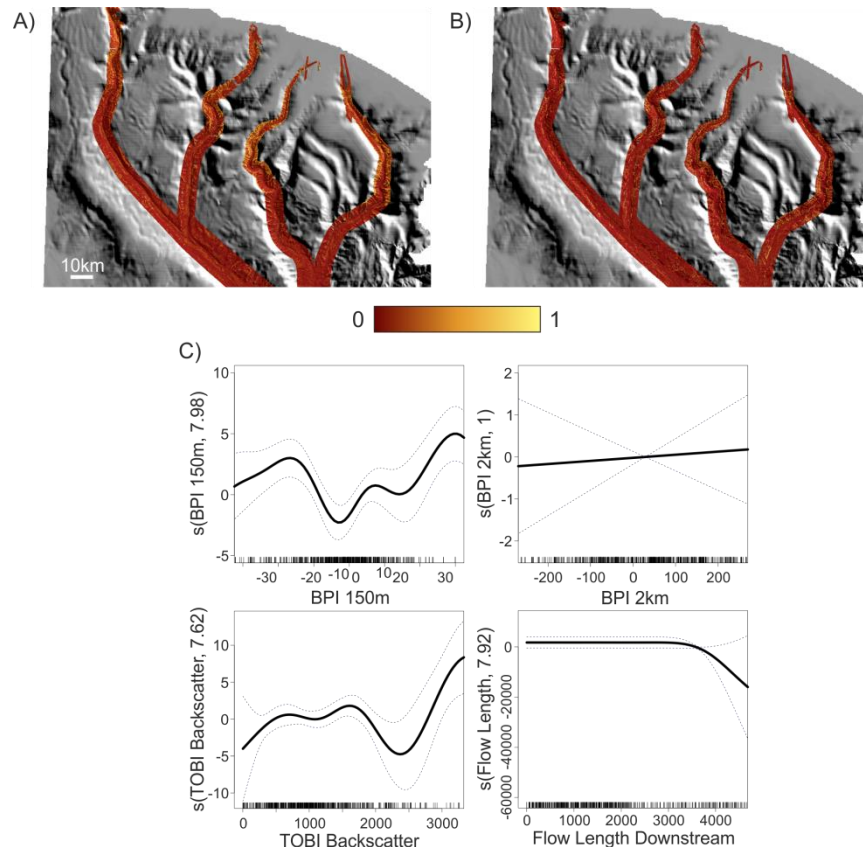
Figure 4.8: Model output for predicting the presence of scleractinian corals within a 125m^2 section of seabed across the extent of the multibeam survey: A) random forest prediction map and B) general additive model prediction map. C) Relationship (centred smooth component, black line, and estimated degrees of freedoms in parenthesis) for the selected environmental variables using GAMs with the dashed lines representing the standard error estimated for the predictions. Background bathymetry (201m resolution) provided by the Geological Survey of Ireland (GSI Dublin).

In the case of abundance, uncertainty (Figure 4.5 B, D) was highest in shallower and deeper sections of the canyons past the limit of the available data. Higher variability in the predictions of species richness (Figure 4.6 B, D) and the reciprocal Simpson index (Figure 4.7 B, D) was observed in association with the canyon walls in elevated areas characterised by steep slopes. Higher overall variability was typically observed for RF predictions.

Overall, GAMs appeared better able to model the biological characteristics of interest using the available environmental variables (Table 4.2). GAMs were able to explain 59.3% of the variation in species richness, 43.2% in abundance and 29.7% of the reciprocal Simpson index, compared to 30.6%, 33.1% and 23.4% when RF were considered. However, RF consistently scored higher than

Table 4.2: Model performances in percentages, values reported represent deviation explained for GAMs and variation explained for random forest, except for the evaluation of cold-water coral presence against independent data which represent measures of accuracy. For the modelled data, percentages in parenthesis represent the additional variance explained when fine-scale environmental variables extracted from the imagery were also considered. For model agreement the first value is the percentage agreement when only values at sampled locations were compared, while the percentages in parenthesis represent a comparison using the entire extent of the prediction maps.

| | | Model Performance (%) | | | |
		Modelled Data	Split-Sample Data	Independent Data	Model Agreement
Abundance	GAM	43.2 (16.9)	33.2	4.3	56.1 (29.3)
	RF	33.1 (13.5)	40.9	11.1	
Species Richness	GAM	59.3 (3.4)	31.5	NS	64.3 (31.0)
	RF	30.6 (8.7)	39.9	NS	
Reciprocal Simpson Index	GAM	29.7 (0.7)	15.4	14.3	54.8 (40.0)
	RF	23.4 (1.8)	28.8	9.4	
Coral Presence	GAM	59.8	NA	65.3	17.7 (4.5)
	RF	41.0	NA	71.4	

GAMs when the removed portion of the dataset was compared. When a completely independent dataset was used, models for abundance and reciprocal Simpson index were only able to explain a small percentage of the variation observed, while the model for species richness was not significant. Differences in relative species composition were apparent (supporting information, 4.8.3 and 4.8.4) and may in part result from the shallower depths (depths for which highest abundances and richness were recorded during the *JC* cruises were not sampled during the *Belgica* cruise) and relative differences in substratum type encountered in the two inner branches. Accuracy measures for the prediction of cold-water coral presence indicated that RF outperformed GAMs, obtaining values of 71.4% as opposed to only 65.3%. RMSE values indicated a clear difference between the methods of model assessment, with the use of an external dataset consistently showing greater deviations from predicted values (Figure 4.9). Model agreement between GAMs and RF varied around ~55%, with the most similar model produced for species richness, with a 64.3% agreement, when values at the position of the data points used to build the models were compared. When the predictions for the entirety of the canyon were compared, model agreement fell to ~30% with the highest agreement for diversity at 40%. Although coral presence was found to have the highest percentages of deviance (GAMs) and variance (RF) explained, 59.8%

and 41.0% respectively, agreement between the two models was lowest at just 17.7% for the sampled locations.
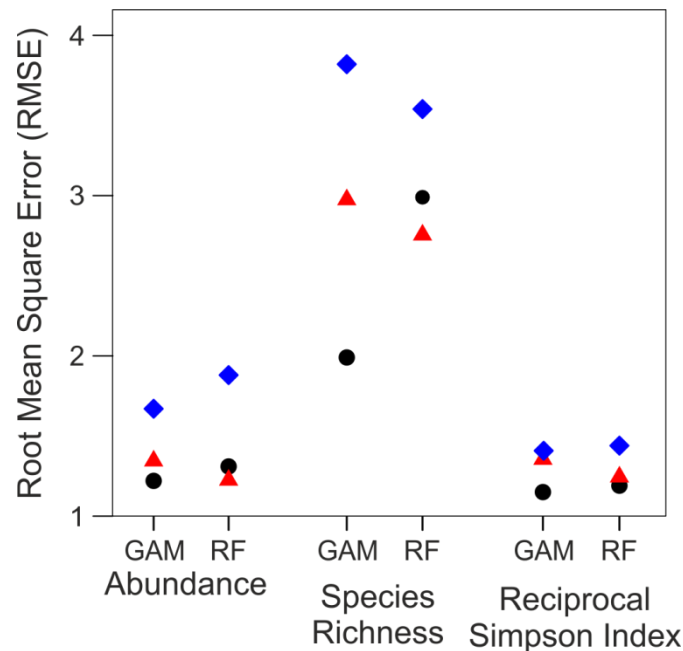
Figure 4.9: Root-mean-square-error (RMSE) for the different types of model assessment, modelled data (black circles), split-sample data (red triangles) and independent data (blue diamonds)

If finer-scale substrata and terrain gradient information as derived from the imagery was incorporated into the model, performance was increased by an additional 16.9% and 13.5% for GAMs and RF respectively in the case of abundance, and 3.4% and 8.7% in the case of species richness (Table 4.2). Fine-scale environmental information did not appear to help model the reciprocal Simpson index. The highest abundance and species richness was found in areas of vertical walls covered by the cold-water coral *Lophelia pertusa* (Figure 4.10). In these 50m sections, average abundances of 734 individuals and 15 different morphospecies were recorded.

Only 33.4km^2 (representing 0.52% of the 3D area of canyon surveyed) of seabed in the Whittard Canyon was likely to have vertical walls and only a total of 9.4km^2 of seabed or 0.15% had vertical walls shallower than 2,000m in depth where colonial cold-water corals were likely to occur. However, even fewer vertical areas were identified as likely harbouring high abundances,

species richness or diversity (1/D), and even more rarely to be suitable for coral growth (Figure 4.11). These respectively represent 11.2km^2 or 0.17% (4.3km^2 or 0.07% shallower than 2,000m depth) and 1.5km^2 or 0.02%.
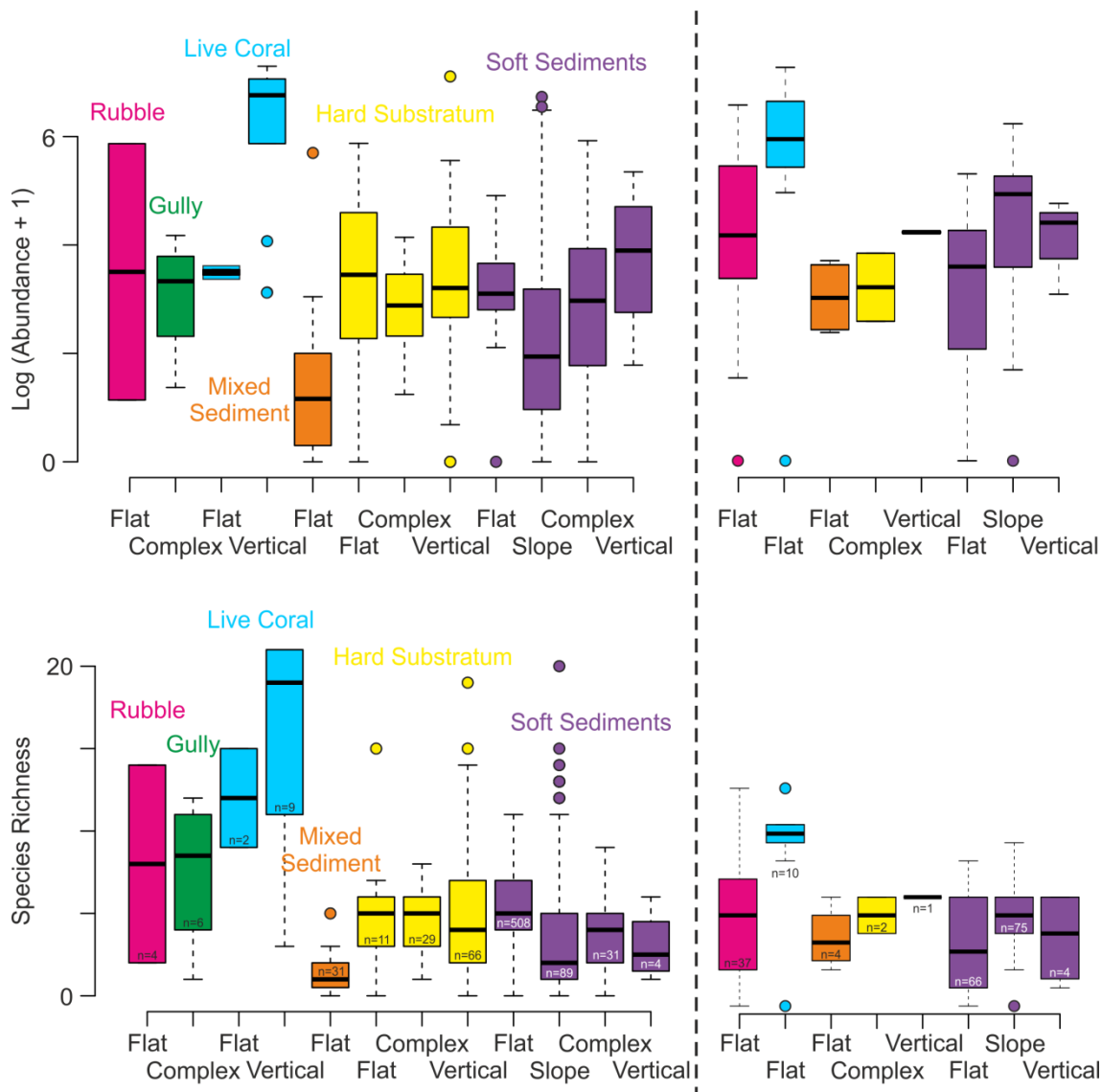
Figure 4.10: Boxplot of abundance and species richness as observed for different substratum types and terrain gradients , ROV *ISIS* transects (left) and ROV *Genesis* (right). The number (n) of 50m sections from each substratum type is shown.
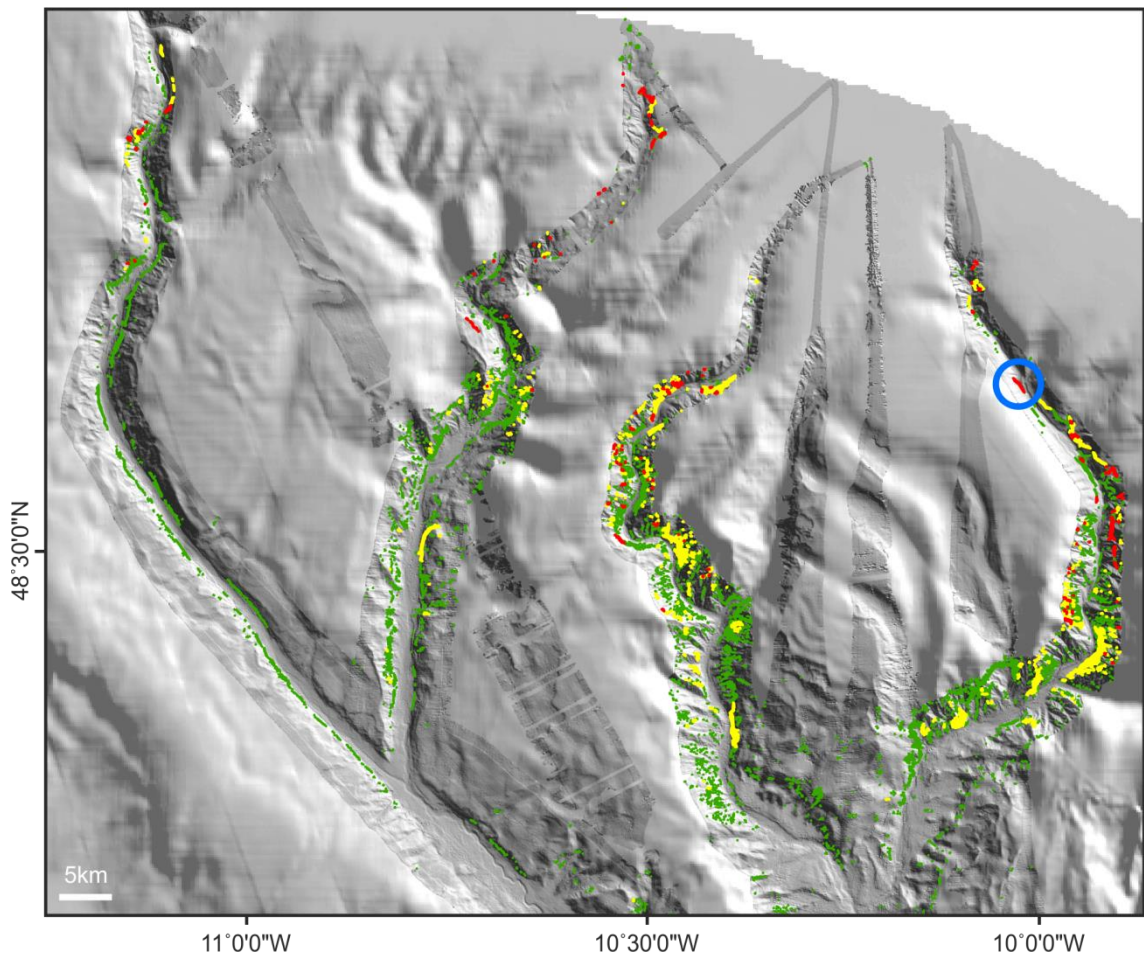
Figure 4.11: Predictive maps of steep slope areas (slopes of >35°, representing near vertical walls, green) for which GAMs and RF models indicated a high combined potential for abundance, species richness and/or diversity (1/D) (yellow) or cold-water corals (red). The blue circle represents the location of the vertical wall colonised by *Lophelia pertusa* and imaged during the *JC* -035-036 cruises. Background bathymetry (201m resolution) collected by the Geological Survey of Ireland (GSI Dublin).

4.5   Discussion

Although submarine canyons are numerous (Harris and Whiteway (2011) identified over 5,800 large canyons worldwide), very few have been mapped to high resolutions and even fewer attempts have been made to build full coverage biological maps. Both GAMs and RF, although predictions differed, were able to produce useful predictive maps for abundance, species richness, reciprocal Simpson index and cold-water coral presence. Such differences in predictions between different models are not unexpected (Araújo & New 2007, Palialexis et al. 2011), and consideration of both model outputs is necessary when building ensemble predictions that can strengthen conclusions.

4.5.1   Whittard Canyon Processes and Ecology

The environmental parameters selected as most useful in modelling abundance, species richness and diversity (1/D) support what is currently known of the ecology of submarine canyons and cold-water corals. The spatial patterns observed can be explained in part by the morphology of Whittard Canyon as well as by examining what has been reported in the literature regarding water mass properties and dominant currents within the Bay of Biscay and for other submarine canyons in this region.

Deep-sea organisms are mostly reliant on the flux of organic material from the sunlit surface waters, which decreases exponentially with depth (Lutz *et al.* 2007). Although the complexity of canyons and their role in channelling material from the continental shelf may create significant variations in the spatial trends observed, decreases in abundance or species richness from a canyon's mid-slope towards the abyssal plain have been observed (Currie & Sorokin 2014, Duffy et al. 2014, Frutos & Sorbe 2014). In this study, depth (or other unmeasured co-varying environmental factors) appeared as a particularly strong predictor for both megafaunal abundance and species richness. Within Whittard Canyon a general decrease in organic content with depth has previously been found, but with the canyon floor showing local enrichment as compared to the nearby slope (Duineveld *et al.* 2001).

The highest species richness in Whittard Canyon was found between 1,200-1,300m at a large coral wall dominated by *Lophelia pertusa*. These observations coincided with bottom nepheloid layers (1,200-2,000m in depth) and increased suspended particulate organic matter (Huvenne *et al.* 2011). These nepheloid layers are related to the interface between the Mediterranean Outflow Water (MOW), a slightly warmer and more saline water mass that originates as a density driven overflow and flows northward along the continental slope past the Celtic Margin and Porcupine Bank, and the lower salinity Labrador Sea Water (van Aken 2000). The interface between these layers may result in locally focussed hydrodynamic processes such as internal waves, which may help keep organic matter in suspension, creating an enhanced food supply for the coral colonies (Mienis *et al.* 2007). This increased food supply may also propagate downslope and help explain why the highest density of corals found in Whittard Canyon (Huvenne *et al.* 2011) was

located deeper than the potential density envelope suggested to be optimal for coral growth in the NE Atlantic (Dullo *et al.* 2008). Increased flow velocities were also associated with an increase in the density of filter feeders along the slope of the nearby Goban Spur at 1,000-1,500m in depth (Flach *et al.* 1998). Similarly, dense assemblages of oysters, *Neopycnodonte zibrowii*, were found in association with the shallower (~500-800m) interface of the Eastern North Atlantic Water and the MOW in a different branch of Whittard Canyon (Johnson et al., 2013), and in another canyon of the Bay of Biscay (Van Rooij *et al.* 2010a).

In addition to keeping food in suspension, it is suggested that the hydrodynamic regime of the Bay of Biscay resulted in erosional features characterised by 'step-like banks' with walls of hard substratum suitable for attachment (Van Rooij *et al.* 2010a). The MOW flowing from east to west affects the morphology of the seabed with the eastern flank receiving higher sediment deposition while the western flank acts as an obstacle which intensifies bottom current and leads to increased erosion (Van Rooij *et al.* 2010b). Comparable 'step-like' features were visible in portions of the east facing coral wall found along dive *JC*-116 on the western flank of the eastern canyon branch. Similar patterns have been observed for cold-water coral occurrences in Penmarc'h, Guilvinec (De Mol *et al.* 2011), Dangeard and Explorer canyons (Stewart et al. 2014) of the Bay of Biscay while *Lophelia pertusa* frameworks were also associated with vertical cliffs in the Lacaze-Duthiers canyon, north western Mediterranean (Gori *et al.* 2013). *Solenosmilia variabilis* was also observed in the *Lophelia pertusa* dominated wall, but was also found to occur deeper (up to 1,850m).

All biological characteristics measured in this study decreased towards the thalweg. As indicated by the relationships with BPI, standard deviation of slope and curvature, this decrease towards the thalweg may result from a reduction in habitat complexity. As opposed to the more morphologically diverse canyon walls, where overhangs and gullies were observed, the thalweg was, for the most part, characterised by flat areas of soft sediment. High numbers of cenrianthids, ophiuroids sp1 and *Acanella* sp. were sometimes observed in flat sediment areas, but diversity remained low. Of course, this analysis focused on epibenthic megafauna, and patterns in the infauna may

show very different trends. However macrofaunal abundance and biomass in Whittard Canyon decreases with depth, with their composition changing with respect to the relative abundance of different feeding type (Duineveld *et al.* 2001). Positive associations with topographic variability, as measured by fine-scale BPI, standard deviation of slope or rugosity, have been reported for cold-water coral occurrences (Henry et al. 2010, Rengstorf et al. 2013), while these variables have also been found to help explain benthic community composition (Jones & Brewer 2012, Henry et al. 2013). More frequent disturbances within the canyon's thalweg owing to density flows or turbidity currents could also explain the reduced diversity. Similarly, they could also explain the reduced species richness found at shallower depths towards the head of the canyon where more physical activity could be occurring. Patterns of reduced diversity, although mediated by productivity, have been recorded for polycheates in areas of canyons characterized by frequent disturbances (Paterson *et al.* 2011). Lowered infaunal species diversities have also been found at the head of the active La Jolla canyon (Vetter & Dayton 1998).

Overall the morphological complexity of the canyon led to a rich habitat diversity with high species turnover. Although no comparisons were done with transects on the continental slope, this high local variation appeared to provide multiple niches which would promote species coexistence and positively influence regional diversity. Although a few species appeared in multiple habitats (*Acanella* sp., *Anthomasthus* sp. and cerianthids, although species level identification could not be achieved and may lead to further differentiations) many were restricted to a specific set of environmental conditions and very few transects were dominated by a single species. These most common taxa appeared in relatively wide depth bands. Similarly wide depth bands (800-2,100m) had been found for *Anthoptilum* sp., *Anthomastus* sp., and *Acanella* sp. on the Mid-Atlantic Ridge (Mortensen *et al.* 2008).

At two distinct depth bands (700-1,000m and 3,700-4,000m), cerianthids appeared to dominate the observations, suggesting that, although morphologically similar, at least two separate species likely occurred within Whittard Canyon. At the deeper depths, cerianthids were observed buried in soft sediment, while along the shallower transects, they were observed along a semi-lithified sediment wall. Along this transect, burrowing ophiuroids were

also repeatedly observed, forming an assemblage similar to the one described for canyons of the South West Approaches, NE Atlantic (Davies et al. 2014). The authors also describe *Kophobelemnon stelliferum*-cerianthids assemblages where high abundances of *Pentametrocrinus* sp. and xenophyophores are also found. *Kophobelemnon sp.* and cerianthids were observed along the shallower transects carried out in the inner branches. *Syringammina fragilissima* was the dominant species along transect *JC*-114, at depths (1,200m) similar to where high densities of xenophyophores (likely *Syringammina* sp.) were also reported for the Gully canyon, NW Atlantic (Kenchington et al. 2014). This species is also reported for the upper Nazaré canyon (1,500m), NE Atlantic, where they occur on steep sediment-covered slopes in areas of enhanced food fluxes (Gooday *et al.* 2011).

Within particular habitats the presence of certain species, such as cold-water corals, appeared to have a further influence on biological characteristics. Although the diversity of octocorals had been found to decrease in *Lophelia pertusa* reef environments (Morris *et al.* 2013), when all taxa were considered, vertical walls dominated by cold-water corals were found to have the highest diversity of all habitats observed. This may be the result of ecological facilitation whereby the three dimensional coral structure could provide attachment for a variety of sessile species, positively affect hydrodynamic patterns for filter feeders or provide a complex habitat affording protection against predation (Buhl-Mortensen & Mortensen 2004, Clark et al. 2008). By comparison with observations on the Mid-Atlantic ridge and in many other cold-water coral areas, *Acesta excavata* and crinoids were found to associate with *Lophelia pertusa* (Mortensen *et al.* 2008). In addition, *Actinauge* sp., *Desmophyllum* sp., *Echinus* spp., *Gorgonocephalus* sp., *Geodia* spp., and various species of soft corals were commonly observed. These particularly rich and populated habitats may also help maintain the regional diversity by acting as sources to help colonize less suitable (sink) habitats (De Leo et al. 2014).

4.5.2   Assessment and Limitations

Since variations in bottom-water temperature, salinity and oxygen concentration within Whittard canyon tend to covary with depth (Duros *et al.*

2011), the lack of fine-scale layers describing hydrographic conditions is not expected to be a major limitation of the present study. However, adding information on average current velocities or temperature and salinity across different water masses has been found to help explain species spatial patterns in other studies (Henry *et al.* 2013). The bathymetry-derived environmental descriptors were able to explain a good percentage of the spatial variation observed in biological characteristics; still, a significant percentage remained unexplained and finer-scale information on current speeds in areas of internal waves or local variation in food supply may prove useful. Unfortunately, the information was not available at the appropriate fine scale for the entire canyon. Similarly, a detailed seafloor sediment interpretation map was not available for the Whittard Canyon. The increase in model performance obtained when imagery-derived environmental descriptors were included indicated that fine-scale substratum information would be valuable. Although sidescan sonar backscatter can be used as a proxy for sediment hardness, the complex morphology of the canyon seafloor affects the angle of incidence, and therefore backscatter strength. Even though 'true slant range' correction was applied during processing, the backscatter was still visibly affected by seabed morphology, only providing a useful proxy for sediment hardness in flatter areas.

Presently, the spatial predictions presented represent only a snapshot in time, where spatial patterns may continue to change over time. While temporal variability in submarine canyons can range from diurnal (Matabos *et al.* 2014) to seasonal (Juniper *et al.* 2013) to short- and long-term geological time-scales, the major influence on abundance and species diversity observed in this study was associated with long-lived reef-building coral colonies. For the scale of this survey, no temporal differences were observed between the *ISIS* ROV surveys of 2007 and 2009. As such, the additional one year time difference with the externally collected dataset (ROV *Genesis*) is unlikely to be the reason for the disparity in estimates. As mentioned before, the more plausible causes are the differences in depth and average substratum type.

With the current environmental information available, the statistical approach which produced the best predictions depended on the biological characteristic of interest. RF have been found to outperform many other

statistical techniques including GAMs when predicting distributions of tree species (Marmion *et al.* 2009b), but comparable performances were obtained for butterflies (Marmion *et al.* 2009a). However, GAMs were found to be more successful in predicting fish abundance as opposed to RF which performed better for species richness and diversity (Knudby *et al.* 2010). The authors attribute this difference to the binary splits of tree based methods having more difficulties in predicting extreme values. As the results of RF do not lead to clear relationships between the response and the environmental variables, a more effective approach might be to use GAMs first for data exploration and variable selection while subsequently using RF for predicting (Baccini *et al.* 2004). In our study, although the external dataset came from the two inner canyon branches for which no previous data were available, it did not suggest that one model consistently outperformed the other. However, the assessment clearly indicated that great care must be taken when considering prediction outputs, particularly if only a split-sample assessment was possible and if extrapolation outside of the originally sampled area was carried out.

4.6   Conclusion

Within Whittard Canyon, the highest abundance, species richness and diversity (1/D) were found on vertical structures. However, it is clear from our study that areas predicted to have such potential are very spatially limited. Out of the 100 hours of benthic imagery collected very few such walls were encountered, and our analysis suggests that less than 0.1% of the canyon's surface area may harbour a structure similarly colonized.

This study provides an example of how important ecological areas can be identified using remotely acquired acoustic seabed mapping techniques and a limited amount of biological information. Such an approach provides a cost effective strategy to facilitate the location of rare biological communities and help ensure that appropriate monitoring can be implemented. Without the continued production of adequately-assessed high-resolution biological maps to inform sampling designs, the spatial complexity of diversity patterns within large geomorphological features may be underestimated and rare diversity hotspots may remain difficult to find.

4.7   Acknowledgments

4.8   Supporting Information

4.8.1   Autocorrelation

A)

B)
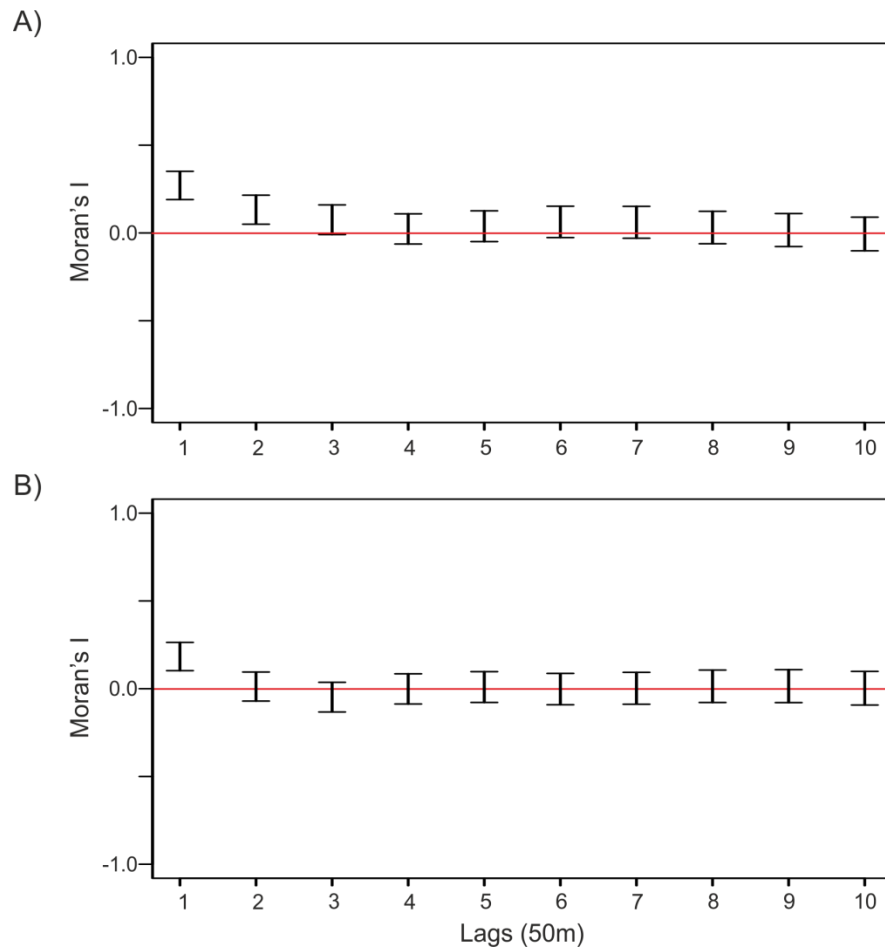
Figure S4.1: Moran's I at various lags based on all continuous 50m sections for A) species richness and B) cold-water coral presence

Figure 4.11

4.8.2   Beta Diversity
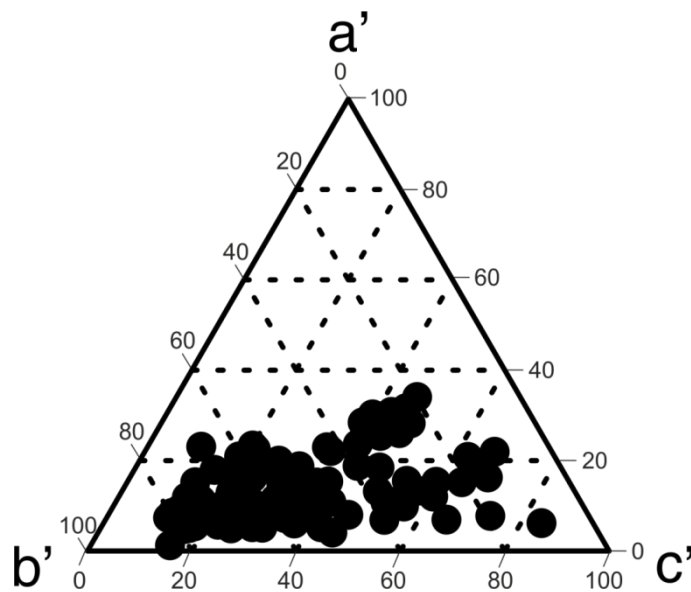
Figure S4.2: Between dive scatterplot showing components of Beta diversity : a' (percentage of species occurring in both samples), b' (percentage of species occurring in neighbouring, but not focal sample) and c' (percentage of species occurring in focal, but not neighbouring samples) as described by Koleff et al. (2003).

4.8.3   Substratum and Species Composition by Transects
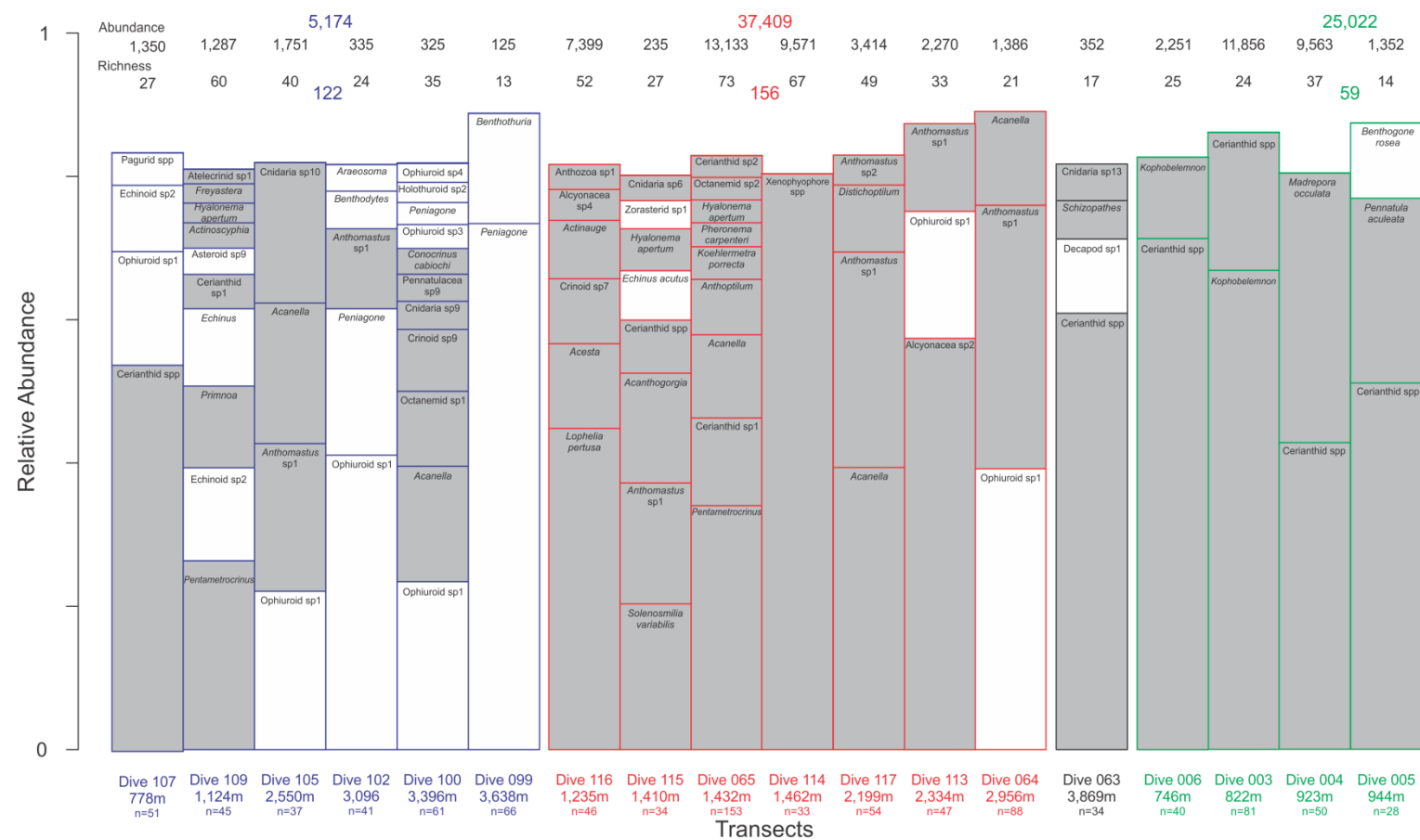
Figure S4.3: Relative abundance of the most common morphospecies for each dive (in grey filter/suspension feeders) whose cumulative sum represents at least 90% of the individual observed. Dives from the eastern and western branches of the canyons are represented in red and blue respectively, while the lower channel is represented in black and the two inner branches visited during the *Belgica* cruise are shown in green. Abundances and species richness are listed above the bars, while corresponding numbers of video sections and average depths are listed below.
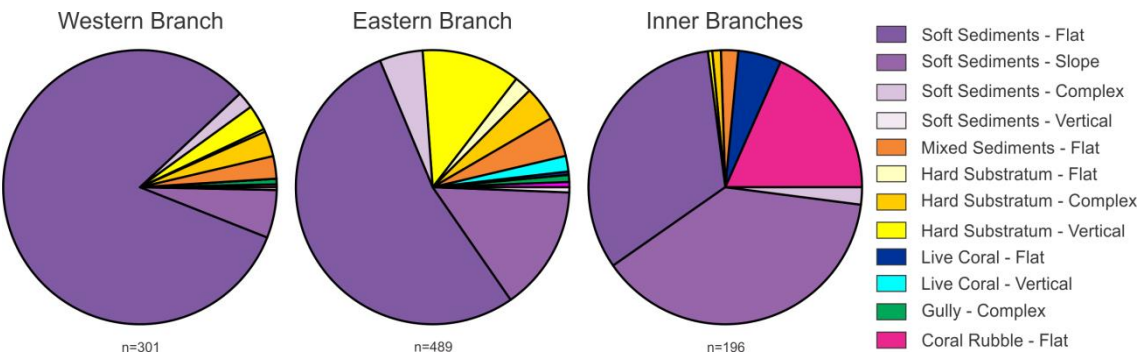
Figure S4.4: Relative abundance of the substratum type encountered in each of the canyon branches . Corresponding numbers of video sections are listed below.

4.8.4 Species Composition by Substratum Type
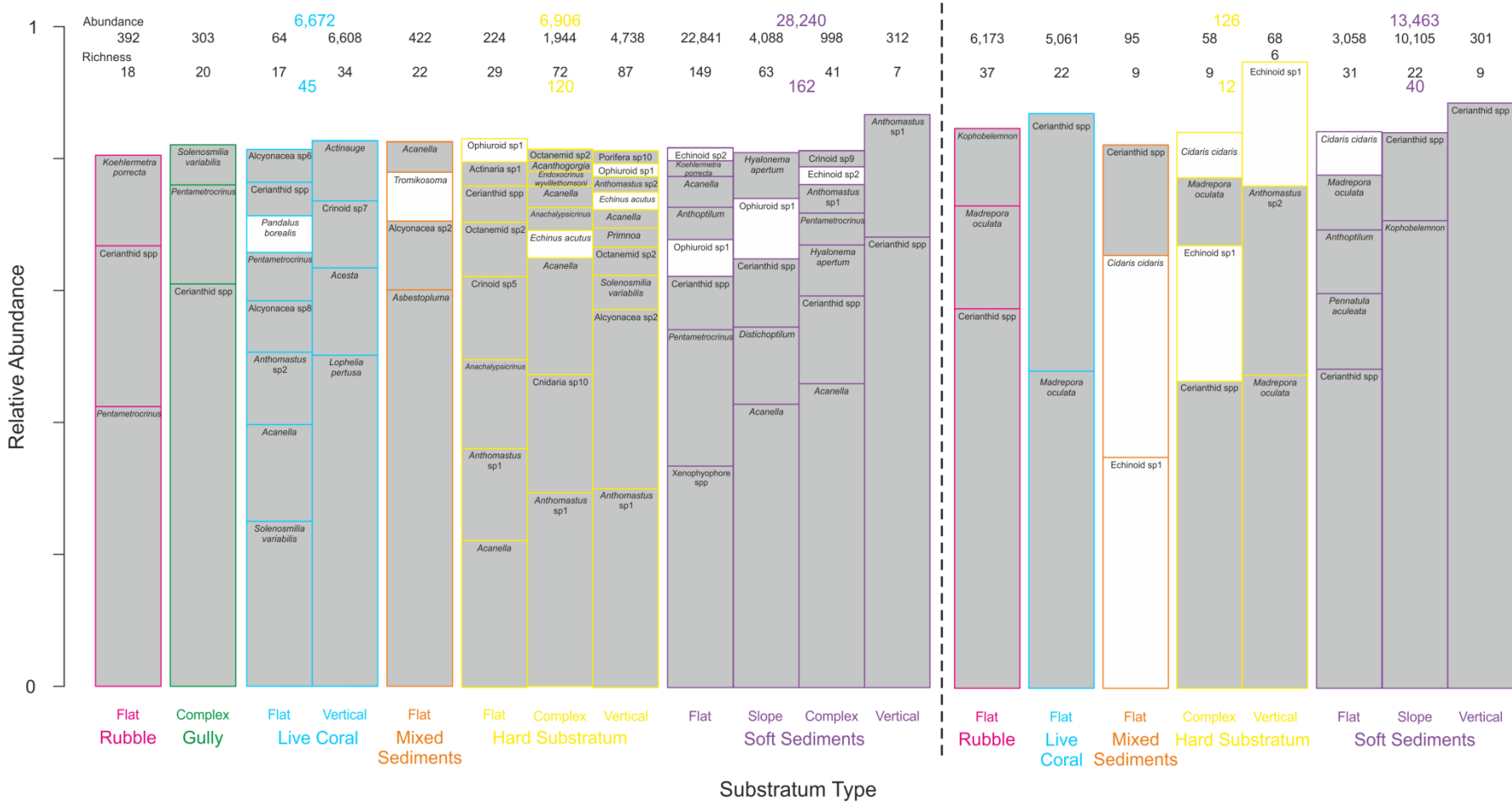
Figure S4.5: Relative abundance of the most common morphospecies for each substratum type (in grey filter/suspension feeders) whose cumulative sum represents at least 90% of the individual observed. Images collected during the RV *Belgica* cruise are shown on the right. Abundances and species richness are listed above the bars. Corresponding numbers of video sections are listed in Figure 10.

4.8.5   Species Composition by Depth Bands
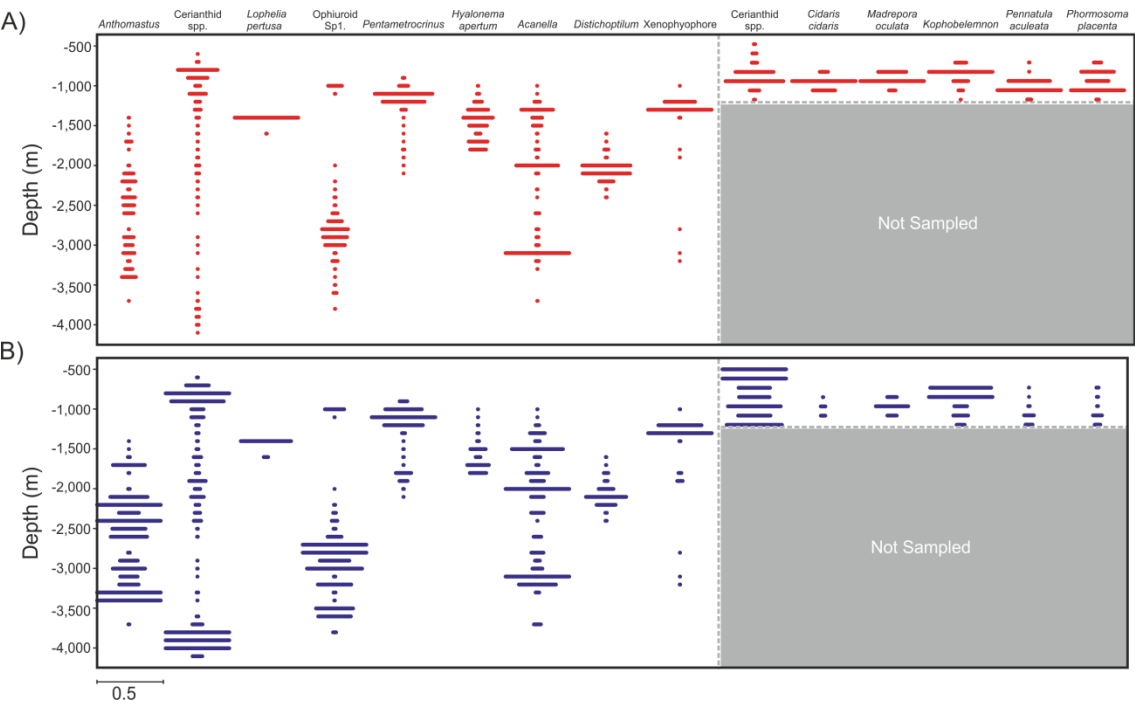
Figure S4.6: Relative abundance of the most commonly observed taxa for each 100m depth bands : A) across all depths for a given taxa and B) across all taxa for a given depth band.

Chapter 5:
Multiple scale analysis of information content: Quantifying information lost for informed mapping resolution choice

Katleen Robert[1], Daniel O.B. Jones[2] and Veerle A.I. Huvenne[2]

[1] School of Ocean and Earth Science, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, UK

[2] National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

5.1　Abstract

Choices of appropriate scale and pixel size are intrinsic to the collection of both ecological and remotely sensed data. Although there is no single appropriate scale at which to study a system and each scale can provide information regarding different processes, in many cases, particularly in deeper marine environments, a single resolution is often chosen arbitrarily owing to data acquisition, time or cost limitations without the knowledge of how this choice affects the information acquired. As part of the CODEMAP project, three areas each of Whittard Canyon and Rockall Bank, NE Atlantic, were mapped at different resolutions using an ROV or AUV, while the surroundings were also mapped at lower resolutions using ship-borne multibeam. ROV video transects were further collected in each area to examine spatial patterns in megabenthic invertebrate density and diversity. Using information theory metrics, we examined how the Shannon entropy and Kullback-Leibler divergence changed across scale for both geomorphological (bathymetry, backscatter and associated derived descriptors) and biological characteristics. A different suite of environmental descriptors was selected for each of the areas considered, but spatial patterns in predictions remained consistent for resolutions up to 20-50m for Whittard Canyon and 2-5m for Rockall Bank. A comparison between raster grids and triangular irregular network (TINs) representation of topography did not show strong differences between the two, but TINs minimized computation time. The information theory approach allowed estimates of substratum cover type to be compared across sampling method (estimates based on benthic imagery and acoustic mapping), and showed that rarer substratum types tended to be underestimated by the latter technique. We suggest that using this approach, fine-scale examination of a small area can provide information upon which broader-scale surveys can be designed as well as help tease out the relative importance of processes occurring over various scales.

5.2   Introduction

Questions of scale are forever recurring in ecology and pixel size is one of the basic considerations of remote sensing (Turner 1989, Levin 1992). At too coarse a resolution, important features driving ecological processes might be missed, while at too fine a resolution, biological interactions may obscure species-environment relationships (Turner et al. 1989a, Wiens 1989). Although there is no single appropriate scale at which to study a system and each scale can provide information regarding different processes, understanding how these change across scale is an important step in determining whether scaling relationships exists and what would be the most appropriate way to handle information collected over different scales (Levin 1992).

Acoustic techniques, usually involving multibeam bathymetry or sidescan sonar backscatter, are the most commonly employed methods to map remotely the seabed (Brown et al. 2011). However, in many cases, particularly in deeper environments, a single resolution is often chosen arbitrarily owing to data acquisition, time or cost limitations without the knowledge of how this choice affects the information acquired. In highly heterogeneous deep-water environments, such as those dominated by iceberg ploughmarks or those in canyon systems, biological communities can change over ranges of 20-50m, while at bathyal depths (200-3,000m) ship-borne bathymetry is often processed at >20m pixel size (Robert et al. 2014). Without *a priori* knowledge of the complexity of these habitats, traditional survey methods would be likely to miss some of the major processes explaining species composition and diversity. On the other hand, too fine a scale might not be economically feasible and results in the trade-off that only very limited spatial extents can be covered. The resulting small sample size will cause decreased predictive potential (Hernandez et al. 2006) and if only a limited range of environmental descriptors was encountered, subsequent attempts at extrapolating species-environment relationships to a broader area will be problematic (Guisan & Thuiller 2005, Menke et al. 2009).

Biological information is often extracted from benthic imagery data. To maximize data acquisition efficiency, long ROV imagery transects are usually acquired and separated into shorter sections for statistical analysis (Jones & Brewer 2012, Robert et al. *submitted*). Although sound reasoning is often

given to justify choice of section length, robust quantitative methods to assess the influence of this decision are limited, which can affect the ability of a study to reveal successfully spatial trends. Optimal section length will vary based on characteristics of the biological community (Juhász-Nagy & Podani 1983). If the sampled area is too small many samples will have no biological records or only contain the most common species, resulting in low diversity and insufficient information to represent community composition and structure (Bartha et al. 1995). When too large a sample size is considered, multiple species assemblage may be included in the sample and spatial differences between samples may become obscured, affecting the ability of a study to establish spatial differences.

As biological information extracted from benthic imagery rarely covers the same spatial extent as the acoustic maps acquired, predictive habitat mapping can be used to identify species-environment relationships in order to build full-coverage biological maps to help inform management decisions (Vierod et al. 2014). Many bathymetry (e.g slope, orientation of steepest slope (aspect), rugosity, curvature) and backscatter (e.g. entropy, standard deviation, skewness) derived environmental descriptors or patch-landscape metrics (e.g. average patch size, patch cohesion) can be considered (Huvenne et al. 2002, Wilson et al. 2007, Robert et al. 2014). These different mapping approaches often lead to environmental variables being quantified over different scales, but the relative importance of each scale is usually difficult to assess. As multiple scales are likely to provide valuable information, their integration would be profitable.

In the majority of habitat mapping studies, the terrain is represented using raster grids with square pixels of a given size. However, other representations such as 'triangulated irregular networks' (TINs) exist and, in certain circumstances, have been shown to represent better terrain variability, while reducing computation time (Vivoni et al. 2004). TINs are based on Delaunay triangulation whereby a set of points are spatially placed to form triangular facets with maximized minimum angles (to limit long thin triangles) so as to avoid any points being within the circle formed by any other triangle. Although a raster grid is required to set up the TINs, their main advantage is that more triangles can be used over highly variable terrain, while fewer are

used in flat areas. Hence, resolution is not static as in raster grids, but varies across the terrain, minimizing the number of nodes required and speeding up computations (Pedrini 2008). By depicting nodal terrain features, TINs also have a better ability to represent three dimensional structures as opposed to two dimension raster grids (Lee 1991), and as such may be of particular interest for topographically complex areas.

It is not surprising that predictions made for the same area using maps of different resolutions have been shown to produce different results, not all of which are suitable for their intended use (Rengstorf et al. 2012). Although much importance as been given to acquiring high-resolution datasets, depending on habitat variability and heterogeneity, this might not always be required. If information on terrain or biological variability and optimal pixel size could be acquired rapidly using a brief preliminary survey, time and money investments on high-resolution surveys could be spent only when likely to provide significant improvements.

In this study, we examine how much information is lost by acquiring or processing data to increasingly coarser resolutions in order to determine the most useful scale at which habitat mapping should be carried out within two complex deep-sea environments: a morphologically variable submarine canyon and a bank with heterogeneous substratum composition characterized by iceberg ploughmarks. Such questions are particularly important in these environments as rare habitats tend to be overlooked at coarsened resolutions which affects the transfer of information across scales (Turner et al. 1989b). For examining bathymetric variation, we used multibeam echosounder maps collected at three different resolutions (represented using both raster grids and TINs) over three areas. Heterogeneity in sediment type percentage cover was explored using estimates obtained from benthic imagery, high-resolution sidescan sonar maps and coarser multibeam backscatter maps. In both habitats, we coupled this environmental information with remotely acquired imagery from which biological characteristics (density and diversity) were derived. We employed an information theory approach based on metrics such as Shannon entropy and Kullback-Leibler divergence (methods developed for the quantification and transmission of information, and signal processing) to link geomorphological and biological characteristics across scales (Brunsell et

al. 2008, Brunsell & Anderson 2011). We suggest that such an information theory approach based on a fine-scale examination of a small area can guide larger survey designs as well as help tease out the relative importance of processes occurring over various scales.

5.3   Methods

5.3.1   Habitat Surveys

5.3.1.1   Whittard Canyon

In 2009, three areas of Whittard Canyon, Bay of Biscay, NE Atlantic, were mapped at fine resolutions (1m) using a Simrad SM2000 multibeam echosounder on the working-class ROV *ISIS*, while the broader canyon was also mapped at a lower resolution (50m) using the RRS *James Cook*'s EM120 multibeam system (Figure 5.1) during cruise *JC*-035 -036. The area had also been mapped at a third resolution (111m) as part of the Irish National Seabed Survey (INFOMAR) whose data are available online (http://www.infomar.ie/). All *JC*-035 -036 bathymetric data were processed using the CARIS HIPS & SIPS software suite or IFREMER's CARAIBES software and all maps were projected to WGS1984, UTM Zone 29N grids.

The finest resolution grids (1m) were resampled to coarser resolutions (3, 5, 10, 20 and 50m) using the mean aggregate function in ArcGIS 10 to simulate the smoothing associated with lower resolution multibeam. The *JC* ship bathymetry (50m) was aggregated to 100m, while the INFOMAR bathymetry (111m, resampled to 100m) was aggregated to 200m. With this resampling scheme, a comparison between coarsened high-resolution data and originally acquired coarser resolution data could also be achieved. Bathymetry-derived terrain layers for slope, aspect (converted to northness and eastness), bathymetric position index (BPI, difference between a cell value and average of surrounding cells), terrain ruggedness index (TRI, average of the absolute differences between cell value and the values of the surrounding cells) and roughness (difference between maximum and minimum of the surrounding cells) were generated using 3x3 pixel size windows for all resolutions in the
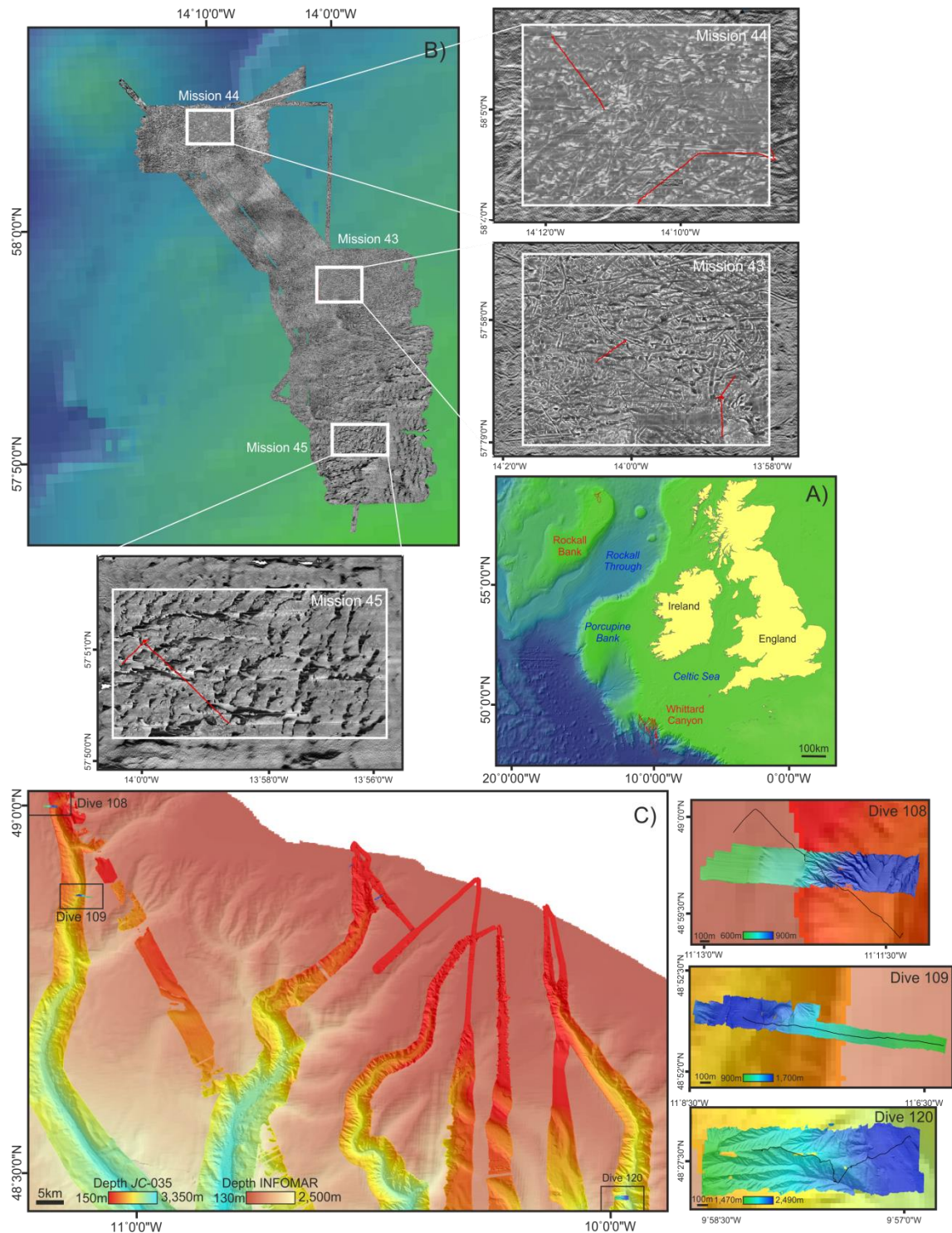
Figure 5.1: Multi-scale surveys of Rockall Bank and Whittard Canyon . A) Location within the northeast Atlantic, background bathymetry from GEBCO (http://www.gebco.net/). B) The Rockall bank backscatter from the RRS *James Cook* multibeam survey (2m resolution) with the 3 AUV sidescan sonar surveys (0.5m resolution) and ROV imagery transects. C) The Whittard Canyon RRS *James Cook* multibeam survey (50m resolution) overlaid on the INFOMAR bathymetry (100m resolution) and the 3 ROV bathymetric surveys (1m resolution) and ROV transects.

statistical software R (R Development Core Team 2014) using the package
'raster'.

TINs of comparable resolutions to each of the raster grids were derived
using ArcGIS 10 from each of the tree surveys. As the resolution of TINs is
variable, resolution was determined by taking the average area of the triangles
(Stoy et al. 2009b). When fewer nodes are used, the TIN model deviates
further from the raster grid model and coarser representations are obtained.
TINs were converted to a polygon shapefile and, for each triangle, terrain
layers were derived for slope, northness and eastness. In R, BPI, TRI and
roughness for each triangle were calculated based on the depth values of all
surrounding triangles.

Three ROV video transects were collected using a down-looking colour
camera (Pegasus, Insite Tritech Inc. with SeaArc2 400W, Deepsea Power&Light
illumination). All megabenthic invertebrates larger than 1cm were identified to
morphospecies and georeferenced using the ROV's ultra-short baseline
navigation system. Two parallel lasers spaced by 10cm were used for scaling
and to standardize transect widths to 2.5m. To examine spatial patterns in
megafaunal density and diversity (reciprocal of Simpson's index , 1/D) (Hill
1973) were computed using video transects subdivided into segments
representing lengths of 1, 3, 5, 10, 20, 50, 100 and 200m.

5.3.1.2 Rockall Bank

Instead of comparing TINs and rasters, for Rockall bank estimates of
substratum percent cover obtained from three different mapping techniques of
varying resolutions were compared. In 2011, during expedition *JC*-060, three
areas of seabed on Rockall Bank were mapped using an AUV mounted sidescan
sonar (EdgeTech FS2200, 410kHz) at very high resolutions (0.5m; Robert et al.,
2014). These areas were located within a larger area of a bathymetric survey
(Kongsberg EM710) on the northwestern side (Figure 5.1), for which
backscatter data was processed to 2m resolution. Both sidescan sonar and
multibeam backscatter data were processed using the in-house software PRISM
(Le Bas & Hühnerbach 1998). For the sidescan sonar maps, sediment
interpretation maps were created using an unsupervised K-means classification
based on mean backscatter, average grey level difference, and variance within

a 9x9 pixel moving window (Robert et al. 2014). Six seabed facies (soft and mixed sediments, hard substratum, exposed bedrock as well as coral stand and rubble) were created. For the multibeam backscatter map, the sediment interpretation map was created by first applying a 3x3 pixel median filter to remove noise and, as texture indices were too strongly influenced by the nadir line, by separating the area into classes representing soft and mixed sediments, hard substratum and bedrock based on backscatter intensity only. Five ROV benthic imagery surveys (Kongsberg OE14-208 digital stills camera), representing 1,222 images, were also collected. For each image, 100 randomly distributed points were selected and the substratum type identified to estimate percentage cover.

The AUV acoustic dataset (0.5m) was aggregated (mean function) to 2, 5, 10 and 20m, while the ship-based maps (2m) were aggregated to 5, 10 and 20m pixels. The sediment interpretation maps were coarsened using the majority aggregation routine in ArcGIS and percentage covers based on 3x3 pixels surrounding image locations were computed. Backscatter mean and variance were also computed for the same windows. However, as backscatter data is not standardized different systems will give different values, as such direct comparison between sidescan and multibeam backscatter values for mean and variance could not be achieved. Still, comparison between derived environmental descriptors (e.g. percentage cover of substratum type) was carried out.

The benthic imagery was treated as for Whittard Canyon, all organisms larger than 1cm where identified to morphospecies and georeferenced. All images located within transect sections with lengths of 1, 2, 5, 10, 20 and 50m, were used to derive estimates of density and diversity (Shannon's diversity index, H') (Shannon 1948). Diversity indices were selected to be consistent with previous habitat mapping studies of these environments.

5.3.2   Statistical Analysis

Information content is quantified using the probability density function (PDF) of a variable of interest (Shannon 1948, Kullback 1959). By examining how it changes across the resolutions considered, we can determine how much information is lost and select the one which maximizes the information

retained, while remaining cost-effective to collect and computationally
efficient. Two metrics were derived for each of the variables (biological and
environmental) considered; the 'normalized Shannon Entropy' ($E_{S,n}$) and the
'Kullback-Leibler divergence' (D_{K-L}).

The Shannon Entropy (H) measures the amount of information needed to
encode a signal, and if the natural logarithm is used, the value is returned in
'natural units (nats)'. In order to obtain a relative measure which can be
compared across scale, the absolute value of H was divided by the natural log
of the number of histogram bins (N) (Stoy et al. 2009b).

$$E_{S,n} = \left| \frac{\sum_{i=1}^{N} p(i) \ln p(i)}{\ln(N)} \right|$$

where $p(i)$ is the discrete probability of a variable being in bin i. $E_{S,n}$ can vary
from 1 when the variable is uniformly distributed (high entropy, low order) to 0
when the variable follows a Dirac Delta function (low entropy, high order) and
less information is needed to encode the signal (Brunsell & Young 2008). A
PDF with a very narrow peak will show a small $E_{S,n}$, while a flat and wide PDF
will show much higher entropy. 20 bins were used to build all PDFs.

The Kullback-Leibler divergence (D_{K-L}) (also known as information
divergence or relative entropy) is a measure of distance between the PDFs of
two variables, x and y (Stoy et al. 2009b, Brunsell 2010). In other words, it
shows the amount of information lost when coarser y is used instead of x.
Hence, a small value indicates more similar PDFs, and similarly to entropy, it is
measured in nats when the natural logarithm is used.

$$D_{K-L} = \sum_{i=1}^{N} p(i) \ln \frac{p(i)}{q(i)}$$

where $p(i)$ and $q(i)$ are the PDFs for variables x and y respectively. However, to
obtain a measure that can be more easily compared across sites and
techniques, the values were normalized using a non-linear transformation
$1 - \exp(-D_{K-L})$, where 0 indicates that the two PDFs are the same and values
closer to 1 indicates strong differences. D_{K-L} was first calculated by
comparing each variable to its finest resolution to quantify the difference in

information following coarsening and then the biological variables were compared to environmental descriptors (e.g. bathymetry-derived such as slope, aspect and rugosity measures as well as substratum type percentage cover derived from the sediment interpretation maps) across all scales. This allows the information in spatial structure of a variable, in this case faunal density or diversity, to be examined with respect to variability in another, in this case environmental descriptors (Brunsell & Anderson 2011). As positive and negative associations between biological and environmental variables may occur, D_{K-L} was also calculated when biological histograms were considered to range from high to low values.

At each scale, prediction maps were generated using 'General Additive Models' (GAMs). As opposed to 'General Linear Models', GAMs allow the building of complex non-linear relationships using non-parametric smoothers such as regression splines (Guisan et al. 2002). Density was log(x+1) transformed prior to modelling and forward selection of environmental descriptors was carried out. Variables resulting in the highest deviance explained were added one step at a time until no more statistically significant (P-Value <0.05) variables could be added. Full coverage maps for density and diversity were built by predicting the expected value for the given section length in each pixel (raster grid) or triangle (TINs). Information theory metrics were computed using the R library 'entropy', while GAM modelling was carried out using library 'mgcv'.

5.4   Results

5.4.1   Information Content

The shape of the PDFs changed differently across scale based on the variables considered, with less change occurring between coarsest resolutions. PDFs for fine-scale resolution bathymetry showed strong peaks that dampened as the resolution was coarsened (Figure 5.2A). The ship-based bathymetry, both *JC* and INFOMAR, showed more similarities to each other than to the coarsened ROV bathymetry. For depth, orientation of steepest slope and rugosity measures (BPI, TRI and roughness) similar PDF curves were obtained

for both raster grids and TINs. However, steep slopes were underestimated by TINs and resolutions coarser than 50m. PDFs derived from backscatter data showed peaks at both low and high percentage covers for common substratum types, such as soft, mixed and hard sediment, which tended to migrate towards mid-range percentage covers as the resolution coarsened (Figure 5.2B). Percentages covered derived from the benthic imagery illustrated that percentage covers derived from the acoustics maps tended to be underestimated, particularly for rarer substratum types, such as rubble, bedrock or live corals. PDFs for biological responses showed strong peaks around zero when fine-scale resolutions were considered (Figure 5.3)
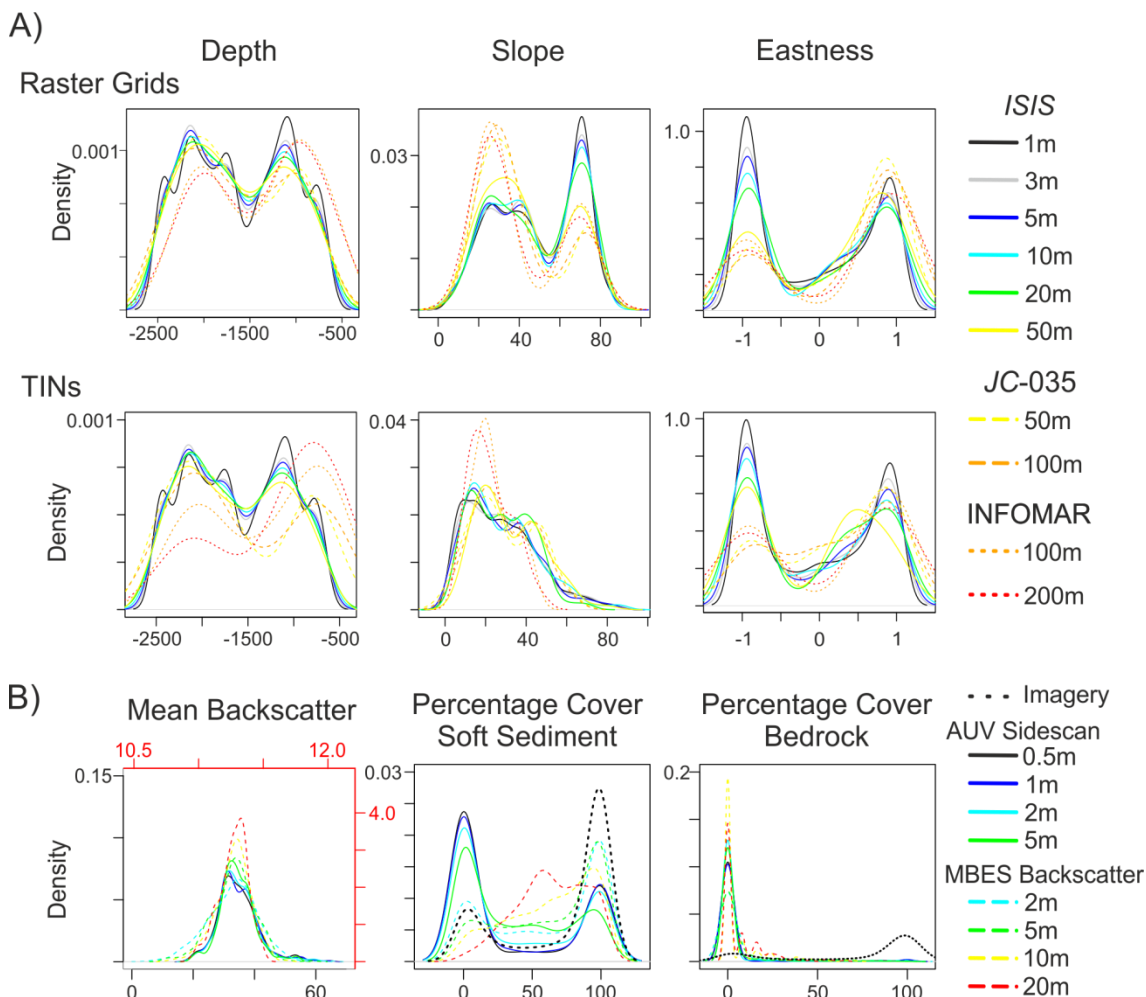

Figure 5.2: Probability density functions for environmental variables A) bathymetry-derived for raster grids (top) and TINs (bottom) and B) sidescan sonar and multibeam (MBES) backscatter (red axis for mean backscatter) derived. PDFs based on locations along the ROV transects
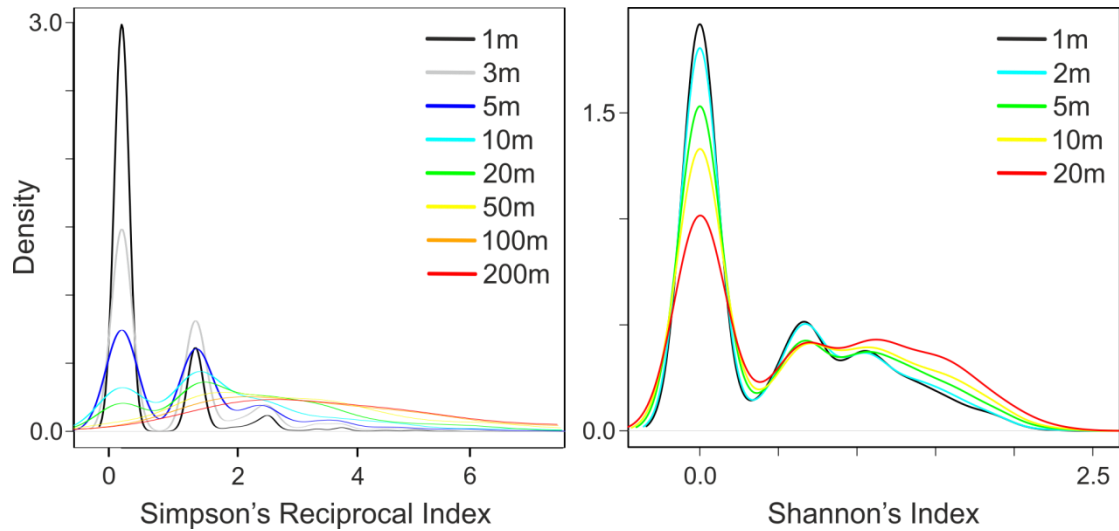
Figure 5.3: Probability density functions for diversity estimates for Whittard Canyon (right) and Rockall Bank (left).

For both raster grids and TINs, the $E_{S,n}$ of morphological characters such as depth, slope and aspect (both northness and eastness) decreased slightly with increasing pixel size, while bathymetric BPI, roughness, TRI and all biological characteristics increased (Figure 5.4A and 5.5A). In the case of the first four environmental variables, the broad PDFs became more concentrated towards mid-range values at coarser resolutions (less extreme values) when the fine-scale topographic variability was lost. On the other hand, for the other variables, which compare pixels with their surroundings, the finest resolutions led to localized calculations which failed to capture the surrounding change in topography and resulted in very narrow, peaked PDFs. However, too coarse a resolution resulted in the topographic variation being lost, implying that a mid-range resolution may be most useful. For environmental and biological descriptors, a transition in the rate of change in $E_{S,n}$ occurred around 20-50m pixel sizes, with values changing quickly until then, but more gradually thereafter. D_{K-L} values for bathymetry, slope and aspect remained nearly identical until 50m, thereafter increasing slightly. PDFs for TRI, BPI and roughness were much more strongly affected by changes in resolution, showing much higher D_{K-L} values than any other variable considered. These descriptors also showed more variable trends when computed using TINs. Comparing $E_{S,n}$ and D_{K-L} of resampled rasters to rasters acquired at coarser resolutions did not show any consistent trend. Similarly shaped $E_{S,n}$ curves
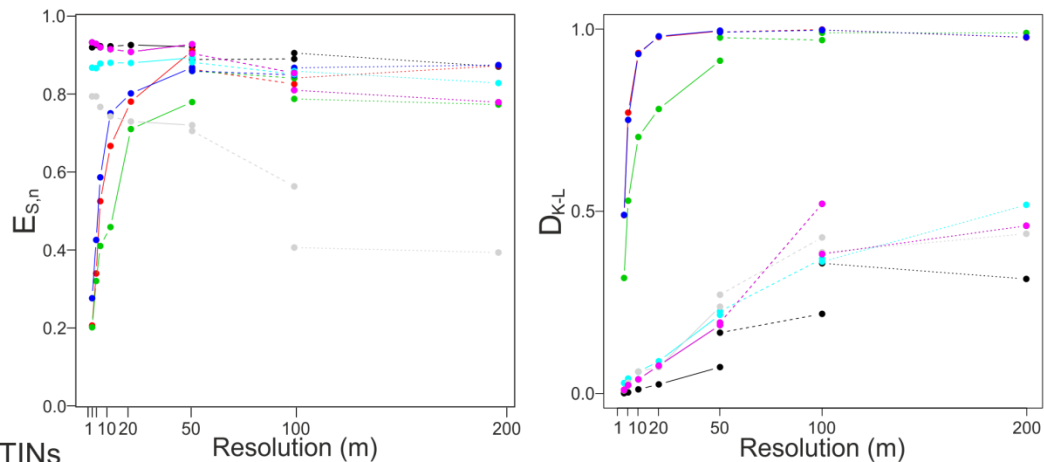
163

A)
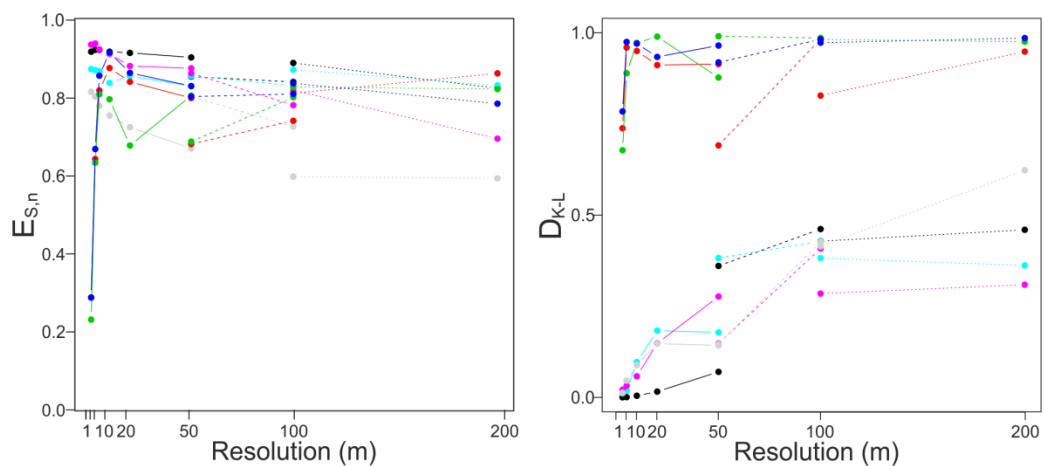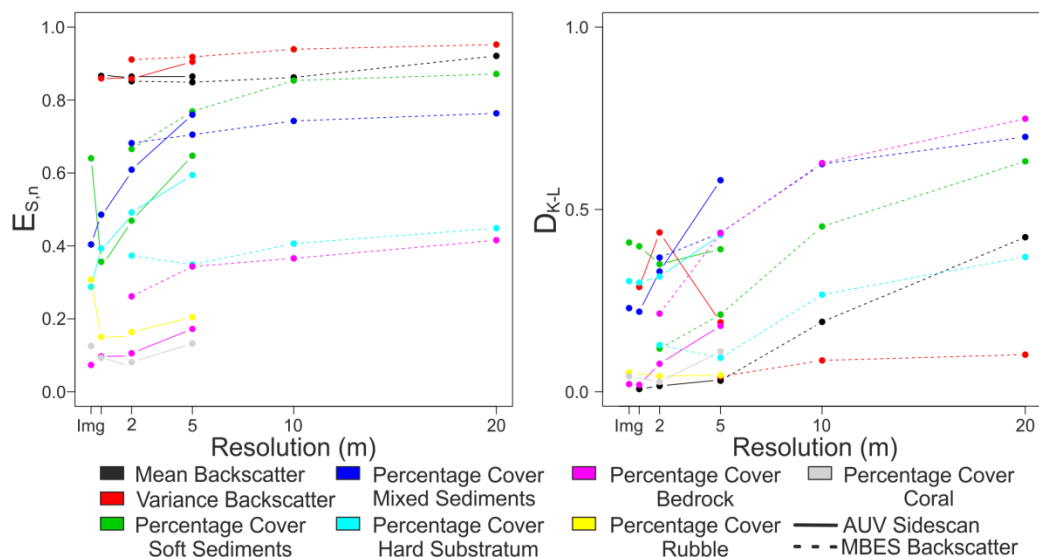
Figure 5.4: Normalized Shannon entropy and normalized Kullback-Leibler divergence across scale for environmental descriptors calculated A) for Whittard Canyon using raster grids (top) and TINs (bottom) as well as B) for Rockall Bank. PDFs built using environmental data at locations along the imagery transect.

A)

B)
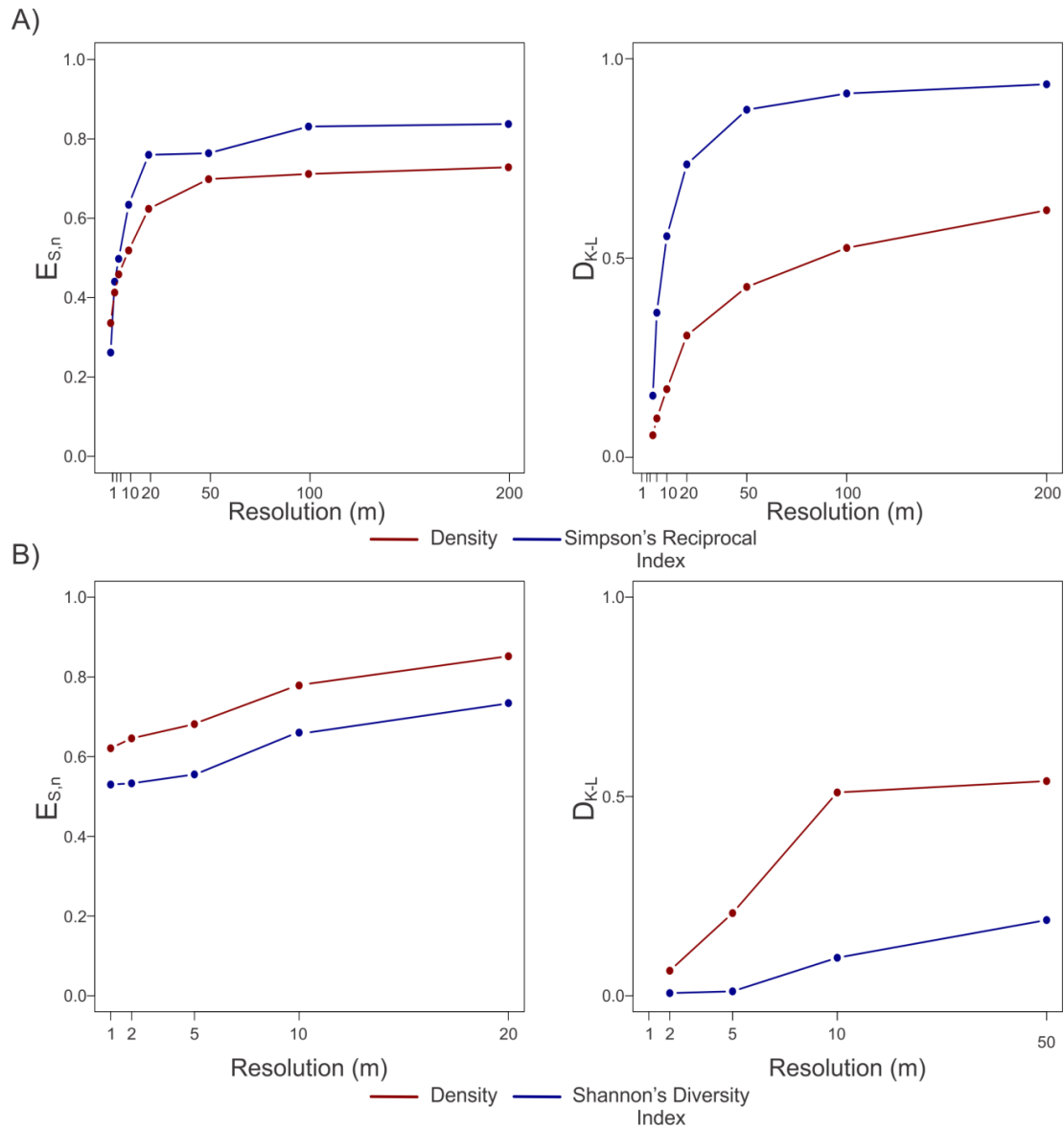
Figure 5.5: Normalized Shannon entropy and normalized Kullback-Leibler divergence across scale for response variables A) Whittard Canyon (Simpson's reciprocal index of diversity and density) and B) Rockall Bank (Shannon's diversity index and density) as calculated from imagery transect sections of varying lengths.

were observed for Rockall Bank, but the transition in the rate of change was observed around 2-5m pixel sizes (Figure 5.4B). As a result of frequent zeroes, rare substratum types tended to have lower $E_{S,n}$. D_{K-L} for backscatter mean and variance showed little change in PDFs until 5m, but differences were apparent for all other derived environmental descriptors computed with resolution >2m (Figure 5.4B). When small section lengths of Rockall Bank imagery transects were considered, $E_{S,n}$ values for diversity and density were

higher than those observed for Whittard Canyon (Figure 5.5B), indicating a more variable environment at fine-scales. However, D_{K-L} for Rockall Bank showed little change at the finest scales and were generally smaller than for Whittard Canyon, suggesting that the finest biological scales were able to capture the biological patterns that required longer sections lengths to be considered on Whittard Canyon.

The D_{K-L} indicated that bathymetry-derived environmental descriptors in Whittard Canyon provided information on the spatial structure of biological characteristics over a range of scales, but similar trends were observed for both density and diversity as well as between raster grids and TINs (Figure 5.6A). At the finer-scale environmental variables such as roughness and TRI provided the most information, while slope, depth, eastness and northness were more useful at fine to medium scales. BPI appeared influential at medium to broad scales. Based on these results, is clear that coarser scale biological patterns are associated with fine-scale information on depth and slope (smaller divergence in red occurs in top-left corner), but that these environmental variables considered at coarser scales do not provide much information on fine-scale biological patterns (larger divergence in blue occurs in bottom-right corner). For Rockall Bank, PDFs for mean and variance backscatter showed more similarities to biological PDFs at coarser scales, while the percentage cover of different substratum types acted at fine to medium scales (Figure 5.6B). When negative associations were investigated, smaller D_{K-L} values appeared at broader scale than for positive associations (supporting information 5.8.1). This shift in scale may be explained by the association of low densities and diversities with flatter regions in Whittard Canyon and to wide patches of soft sediments in Rockall Bank.

5.4.2   Predictive Models

For resolutions less than 50m in Whittard Canyon, depth always appeared as the environmental factor having the strongest influence on both diversity and density, while at coarser resolutions BPI gained in importance, particularly for density (Table 5.1). At finer resolutions, northness had a stronger influence on density, while diversity was influenced more by eastness. TRI was mostly selected for models based on the finest resolutions, while roughness

Figure 5.6: Normalized Kullback-Leibler divergence between the response variables (density and diversity) and selected environmental descriptors across scales for A) Whittard Canyon (raster grids on the right and TINs on the left) and B) Rockall Bank. For Whittard Canyon the 1, 3, 5, 10 and 20m are derived from the ROV bathymetry, while the 50m resolution is derived from the *JC*-035 bathymetry and the 100 and 200m resolutions are from the INFOMAR dataset. For Rockall Bank, 0.5 and 2m resolutions are from the AUV sidescan sonar, while the 5, 10 and 20m are from the multibeam backscatter. Percentage cover of substratum types estimated from the imagery (Img) are included. Small divergences are show by warm colours and indicate more similar PDFs.

and slope were selected across a wide range of resolutions. Although environmental descriptors based on either raster grids or TINs were able to explain very similar percentages of biological variation, models based on TINs had a slightly higher R^2 at mid resolutions. Up to 50m, predictions were relatively consistent across scales and similar trends appeared regardless of whether raster grids or TINs were used.

For the area covered by ROV dive 120, both diversity (1/D) and density were predicted to be highest in the region of highest complexity, where gullies were numerous, and in the steeper section along the canyon walls next to the thalweg (Figure 5.7). For dive 108, highest density was predicted in the shallowest areas, while increased biodiversity was expected on the canyon walls towards the thalweg. Both biological characteristics were predicted to be highest in the shallower portion of dive 109, in the transition slope between the steeper canyon walls and the flatter continental slope (predictions for dive 108 and 109 provided in Supporting Information 5.8.2).

Table 5.1: For each scale, most parsimonious models for density and diversity as obtained by general additive models for Whittard Canyon using raster grids (white rows) and TINs (grey rows). Bold R^2 indicate which of rasters or TINs provided the best performing model for a given scale. Models based on environmental data derived from bathymetry collected by the ROV *ISIS*, using the *JC*-035 shipborne multibeam and as part of the INFOMAR project are shown separately.

Section Length	n	log1p(Density)	R^2	Reciprocal Simpson Index	R^2
1m - ROV *ISIS*	4973	s(depth) + s(northness) + s(tri) + s(eastness) + s(tpi)	**0.063**	s(depth) + s(roughness) + s(eastness)	**0.072**
	4960	s(depth) + s(northness) + s(slope) + s(eastness)	0.058	s(depth) + s(tri) + s(eastness)	0.071
3m - ROV ISIS	1817	s(depth) + s(northness) + s(eastness) + s(tri)	0.117	s(depth) + s(eastness) + s(roughness)	**0.263**
	1810	s(depth) + s(tpi) + s(northness) + s(slope) + s(roughness) + s(eastness)	**0.137**	s(depth) + s(eastness) + s(slope)	0.255
5m - ROV ISIS	1086	s(depth) + s(roughness) + s(tpi) + s(eastness) + s(slope)	**0.189**	s(depth) + s(roughness) + s(tpi) + s(eastness) + s(slope)	**0.383**
	1088	s(depth) + s(tri) + s(slope) + s(northness)	0.166	s(depth) + s(slope) + s(eastness)	0.373
10m - ROV ISIS	548	s(depth) + s(slope) + s(eastness) + s(tpi)	0.247	s(depth) + s(roughness) + s(eastness)	0.463
	544	s(depth) + s(tpi) + s(slope) + s(northness)	**0.267**	s(depth) + s(eastness) + s(northness) + s(tri)	**0.465**
20m - ROV ISIS	275	s(depth) + s(roughness) + s(tri)	0.304	s(depth) + s(slope) + s(eastness)	**0.525**
	273	s(depth) + s(tpi) + s(roughness) + s(tri) + s(eastness)	**0.362**	s(depth) + s(roughness) + s(eastness)	0.515
50m - ROV ISIS	112	s(depth) + s(tpi) + s(northness)	**0.509**	s(depth) + s(northness)	0.522
	108	s(tpi) + s(eastness) + s(roughness) + s(northness) + s(depth) + s(slope)	0.450	s(depth) + s(slope)	**0.570**
50m - *JC*-035	109	s(depth) + s(tpi) + s(northness)	0.324	s(depth) + s(roughness) + s(eastness)	0.418
	87	s(roughness) + s(northness)	**0.383**	s(depth) + s(northness) + s(slope)	**0.482**
100m - *JC*-035	55	s(tpi) + s(northness) + s(depth) + s(eastness)	**0.571**	s(depth) + s(eastness) + s(roughness) + s(slope)	**0.709**
	43	s(roughness)	0.241	s(northness)	0.453
100m - INFOMAR	76	s(tpi) + s(slope) + s(roughness)	**0.521**	s(depth) + s(tpi) + s(northness)	**0.553**
	55	s(tpi)	0.280	s(depth) + s(roughness)	0.434
200m - INFOMAR	39	s(tpi) + s(tri)	**0.440**	s(slope) + s(roughness) + s(northness)	0.455
	28	s(eastness)	0.085	s(eastness) + s(slope) + s(roughness)	**0.828**

1m - ROV *ISIS*

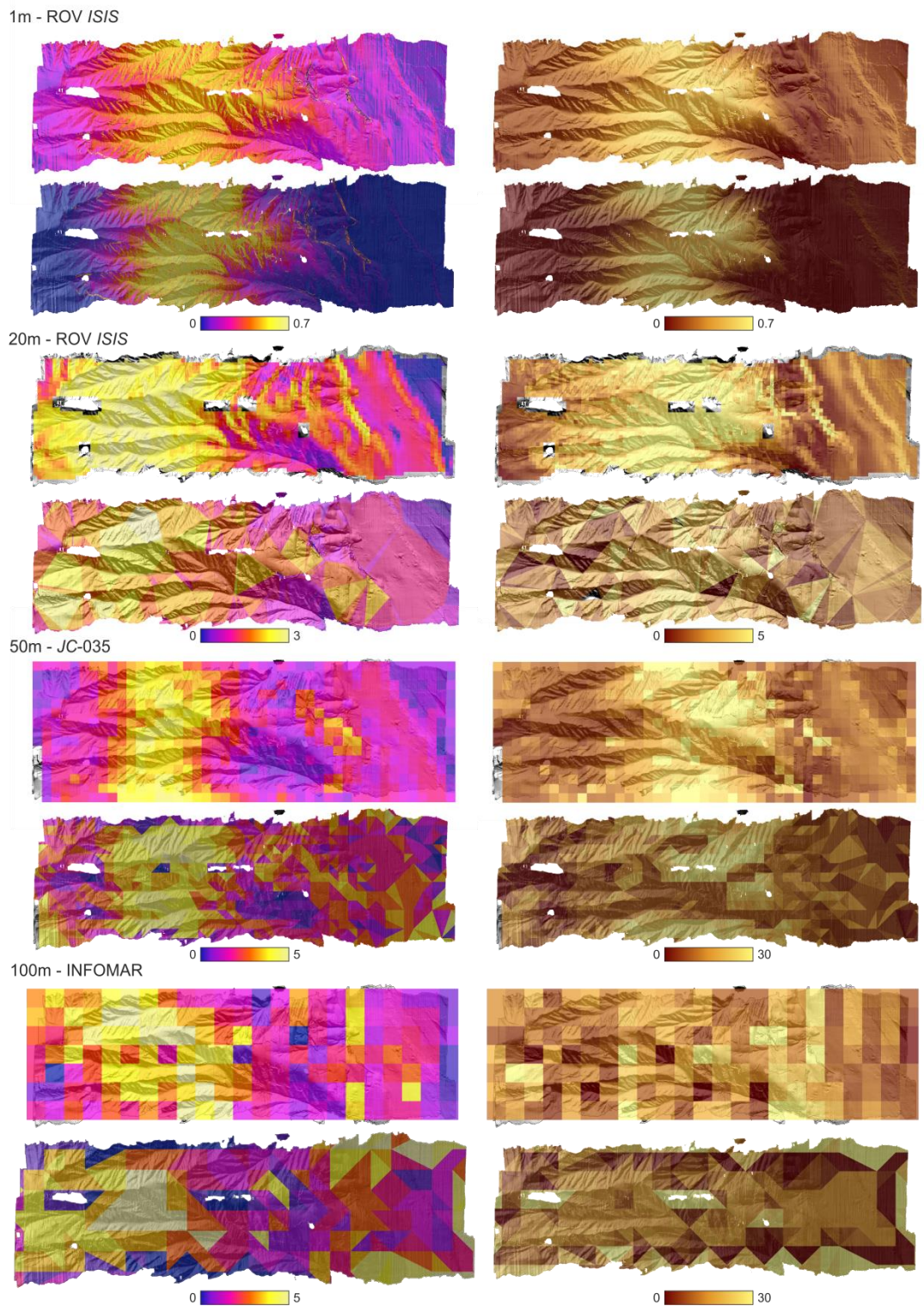20m - ROV *ISIS*

50m - *JC*-035

100m - INFOMAR

Figure 5.7: Prediction comparison for Whittard Canyon's ROV dive 120 based on varying resolutions for raster grids (top) and TINs (bottom) for Simpson's reciprocal index of diversity (left) and density (right) overlain on hillshade built from the 1m bathymetry. Predictions for dives 108 and 109 are presented in supporting information 5.8.1.

At fine resolutions on Rockall Bank, percentage cover of bedrock was most often found as the most significant environmental predictor, with a general increase in both densities and diversity (H') to be expected with increased bedrock cover (Table 5.2). For these scales, presence of soft sediments was also a strong predictor, but the opposite effect was observed. At coarser scales, backscatter mean and variance became more important in predicting both densities and biodiversity. Presence of live corals also positively influenced both biological variables, while rubble had a negative one at higher percentage covers. Estimates of percentage covers derived from the imagery were able to explain a much higher percentage of the variation observed, but cannot be employed for predictive mapping. To characterise common substratum types with large patch sizes, multibeam backscatter remained effective and percentages cover of soft sediments and hard substratum were selected as significant predictors for both diversity and density. However, the multibeam backscatter was unable to capture finer-scale features such as rubble patches or cold-water corals which remained even when the high-resolution AUV sidescan sonar data was coarsened. As a result for the same resolutions, predictive models based on multibeam backscatter were generally only able to explain half of the variation explained by models based on AUV sidescan data for density; diversity was not similarly affected.

Table 5.2: For each scale, most parsimonious models for density and diversity as obtained by the general additive models for Rockall Bank. Models based on environmental data derived from photographic imagery, AUV sidescan sonar and shipborne multibeam backscatter (MBES) are shown separately.

| Section Length	n	log1p(Density)	R^2	Shannon's Diversity Index	R^2
0.5m - Imagery	1222	s(hard) + s(soft) + s(mixed) + s(variance)	0.698	s(soft) + s(hard) + s(mixed) + s(rubble)	0.600
0.5m - AUV Sidescan	1222	s(bedrock) + s(soft) + s(coral) + s(hard) + s(mixed)	0.186	s(soft) + s(bedrock) + s(variance) + s(rubble)	0.159
1m - AUV Sidescan	1222	s(bedrock) + s(mean) + s(coral) + s(mixed) + s(rubble)	0.200	s(bedrock) + s(mixed) + s(hard) + s(coral)	0.164
2m - AUV Sidescan	1052	s(bedrock) + s(mean) + s(coral) + s(mixed) + s(rubble)	0.259	s(bedrock) + s(mixed) + s(hard) + s(coral)	0.190
5m - AUV Sidescan	891	s(mean) + s(bedrock) + s(rubble) + s(variance) + s(coral) + s(hard) + s(mixed)	0.290	s(bedrock) + s(mixed) + s(variance) + s(hard) + s(coral) + s(rubble)	0.222
10m - AUV Sidescan	561	s(mean) + s(bedrock) + s(coral) + s(variance) + s(hard) + s(soft)	0.327	s(mean) + s(bedrock) + s(variance) + s(coral) + s(soft)	0.236
2m - MBES Backscatter	1052	s(hard) + s(variance) + s(mean) + s(soft)	0.141	s(mean) + s(hard) + s(variance) + s(soft)	0.151
5m - MBES Backscatter	891	s(soft) + s(variance) + s(hard) + s(bedrock)	0.173	s(variance) + s(soft) + s(bedrock) + s(mixed) + s(mean)	0.216
10m - MBES Backscatter	561	s(hard) + s(mixed) + s(mean) + s(variance) + s(soft)	0.169	s(mean) + s(hard) + s(mixed) + s(variance)	0.179
20m - MBES Backscatter	377	s(mean) + s(soft)	0.160	s(mean) + s(bedrock) + s(soft)	0.199
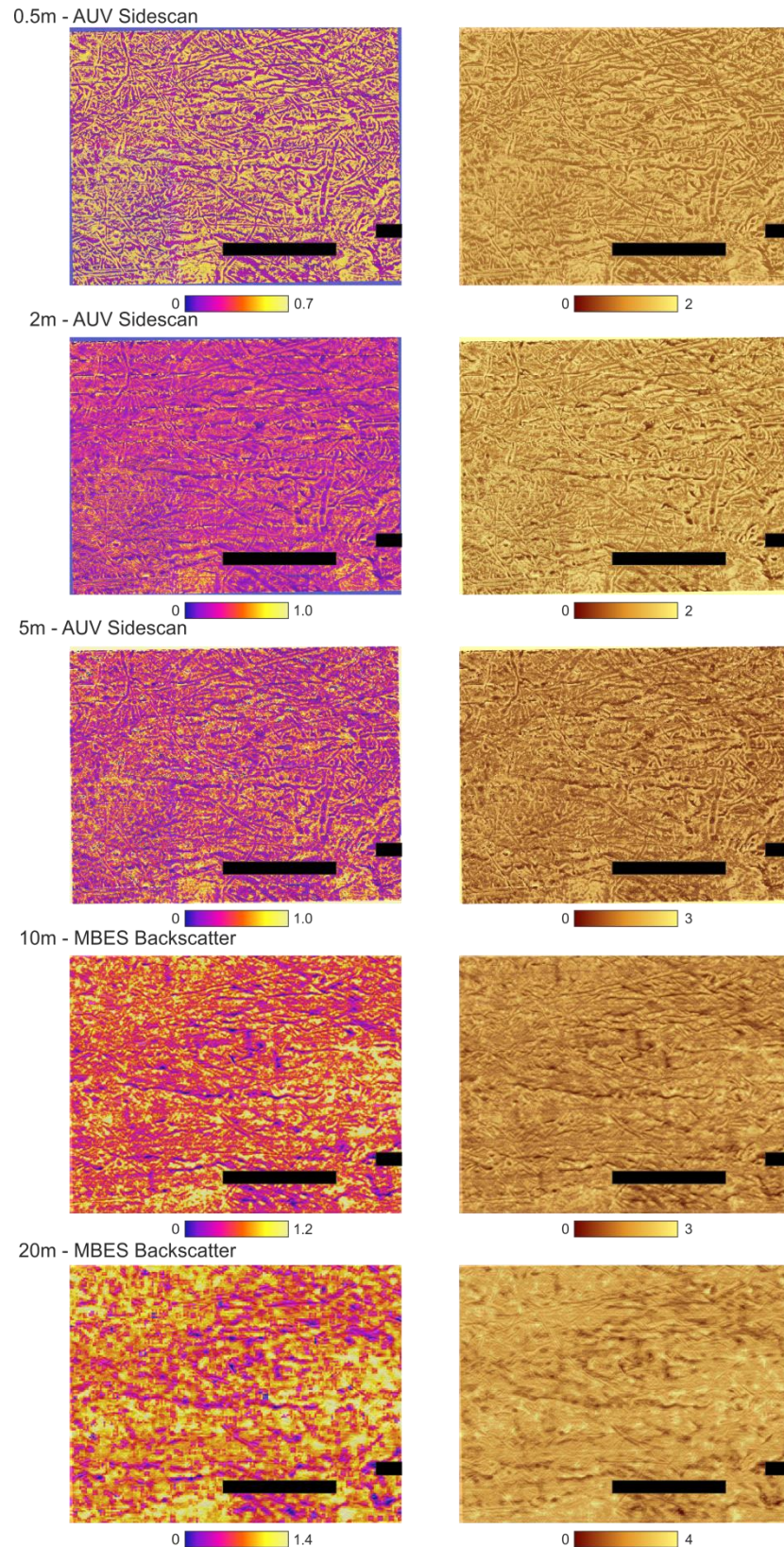
Figure 5.8: Prediction comparison for Rockall Banks AUV mission 43 based on varying resolutions for Shannon's diversity index (left) and density (right) overlain on 0.5m sidescan sonar layer. Black rectangles show areas for which no data are available. Predictions for missions 44 and 45 are presented in supporting information 5.8.3.

A strong influence of substratum composition was observed at fine resolutions, with abundances and diversity clearly following iceberg ploughmarks, low values were observed in the soft sediment infilled centres, while higher values were found on the hard substratum dominated edges (Figure 5.8). Overall lower biodiversity and abundances were predicted for area M44, while both biological variables were higher in area M45 (Supporting Information 5.8.3). For area M45, a clear influence of the presence of bedrock and cold-water corals could only be observed at the finest resolution.

5.5   Discussion

This study is a rare opportunity to examine deep-sea sites over multiple resolutions from fine-scale AUV or ROV derived information (both acoustic and imagery) to ship-based bathymetry and backscatter. As opposed to only collecting data at a single predefined resolution, we were able to use a multiple scale approach to examine the information content of the data collected and determine the most appropriate scale at which data processing or further sampling should be carried out for particular sites. Although such an approach was first developed for selecting plot sizes for vegetation community studies (Juhász-Nagy & Podani 1983), it has been extended to vegetation studies based on remotely sensed data (Brunsell et al. 2008, Stoy et al. 2009b) and our study shows its potential for informing predictive habitat mapping of marine environments using acoustic techniques.

5.5.1   Scale Choice and Spatial Variability

Based on these results, we suggest that for Whittard Canyon, 20-50m is an appropriate scale at which to carry out analyses given the extent and variables (both biological and environmental, based on bathymetry) considered. On the other hand, the Rockall Bank example clearly illustrates that a single scale cannot be applied indiscriminately to all deep-sea areas. Where heterogeneity occurs at much finer scales, resolutions of less than 5m may be required.

Change in entropy across scale was not linear, and this has ramifications when trying to transfer ecological information from one scale to another (Stoy

et al. 2009a). At very short benthic imagery section lengths, too many sections were devoid of organisms or only contained the most common species resulting in consistently low diversities and low $E_{S,n}$. Although a similar effect was observed for density, the tail of the distribution tended to be longer as small high abundance patches affected the distribution. If, instead, longer transect sections were considered, variations in diversity became obscured as local patches disappeared, leading away from zero and towards a more regional diversity measure. Although different diversity indices were used in each of the two habitats, density was overall much higher on Rockall Bank, resulting in few empty samples at finer spatial scales and correspondingly higher $E_{S,n}$. In the case of environmental variables, those with broad PDFs at fine scales (e.g. depth, aspect, mean backscatter or percentage cover of common sediment types) tended to become more centralized with increasingly coarser resolutions and showed smaller changes in $E_{S,n}$. Variables which showed strong narrow peaks and lower $E_{S,n}$ at fine resolutions (e.g. BPI, TRI, rugosity and percentage cover of rarer sediment types such as coral or bedrock) widened with decreasing resolution until the resolution became too poor, resulting in flat terrain or dominance by a single common substratum type.

Differences in rates of species turnover between sites may help account for the different spatial patterns observed for density and diversity on Rockall Bank, compared to Whittard Canyon. High turnover in species composition between transects on Whittard Canyon has been noted previously (Robert et al. *accepted*). However, in part likely the result of the smaller distances between transects, such a difference was not observed on Rockall Bank (Robert et al. *submitted*). As such, even at relatively small section lengths on Rockall Bank, the most commonly occurring species would have been sampled, and changes in diversity were instead more visible at larger spatial scales between survey areas. On the other hand, density could be more strongly influenced by landscape structure, responding to finer-scale variations and requiring higher resolution environmental descriptors. Similar differences in the scales at which habitat heterogeneity needed to be quantified in order to explain patterns of diversity (broader scale) and abundance (finer scale) have also been reported for intertidal zones (Archambault & Bourget 1996). In the current study, both

coral and rubble substrata were frequently selected as explanatory variables for density, but these substratum types were not available when the seabed was mapped using only multibeam backscatter. On the other hand, mixed sediments were repeatedly selected as important environmental predictors for diversity and those remained identifiable even at coarser resolutions.

As indicated by the scale-wise D_{K-L} between biological characteristics and topographic variables, terrain variability captured at medium spatial scales consistently provided information on species spatial structure at medium to broad scales. As sediment characterisation extracted from benthic imagery did not significantly improve modelling of diversity in Whittard Canyon (Robert et al. *accepted*), biological interactions such as competition, predation or facilitation (Menge & Olson 1990, Etter & Mullineaux 2001, Jackson et al. 2001) may be more useful in explaining fine-scale biodiversity patterns. On the other hand, abundance within Whittard Canyon as well as biological variables on Rockall Bank showed responses to substratum percentage cover over finer spatial scales. However, the information available on percentage cover of different substratum types changed across scales, highlighting the importance of taking into account multiple scales.

5.5.2   Implication for Predictive Mapping

As different environmental variables were found to be most useful over different scales and because the scales at which biological patterns are better described are not always the same as those best suited to capture the environmental variables, it is clear that incorporating a multi-scale approach is likely to improve model performances. For this study, broader-scale environmental variables were derived from coarsened bathymetry in order to investigate the implications of resolution choice in data acquisition and a single window size was applied in the predictive models to focus on specific scales and limit the number of variables considered. This led to the drawback that environmental variables representing different scales had different number of pixels and could not all be included in the same predictive model. Deriving environmental descriptors using windows of varying sizes would allow the variability in environmental features to be captured over multiple scales and included in the predictive model (Wilson et al. 2007, Ismail et al. *under*

revision). Descriptors based on both bathymetry and backscatter could be derived and additional metrics such as landscape indices could also be included. These metrics are calculated based on sediment interpretation maps and also allow multiple scales to be taken into account by considering the spatial patterns and complexity of the surrounding habitat patches (McGarigal et al. 2012). As long as the resolutions remain such that smaller patches of important substratum types are still captured within the acoustic maps, these can be characterised and included in the modelling process, while also considering the dominant substratum type in surrounding areas of increasing sizes. All these approaches would allow the broader surroundings of a sample site to be characterised over multiple scales, while retaining the high resolution needed for accurate environmental characterization.

The use of data collected at coarser resolution clearly highlighted problems in the transfer of information across scale, as the presence of rubble or cold-water corals could not be captured by the multibeam backscatter. Comparison between environmental variables derived from different mapping tools (benthic imagery, sidescan sonar and multibeam backscatter) indicated that high-resolution data is required to explain successfully biological characteristics, particularly density, on Rockall Bank. However, the much larger amount of variation explained when imagery-derived percentage cover of substratum type was available suggests that acoustic mapping even at high resolution may underestimate the coverage of certain rarer but important substratum types.

TINs had been suggested as a potential approach to preserve topographic features at coarser resolutions (Stoy et al. 2009b), but our results did not suggest strong differences in their ability to represent terrain variability or improve the creation of predictive maps. However, TINs led to a decrease in the number of computations, from ~2.5 million pixels to ~1.9 million triangles at 1m, and from 1,907 pixels to 1,433 triangles at 50m.

5.6   Conclusion

As a trade-off tends to exist between resolution and extent (high resolution surveys can only cover smaller extents than coarser surveys for the

same amount of time), it is highly valuable to determine the coarsest resolution at which a survey could be carried out and still answer a question of interest. Once achieved, time can be invested in obtaining greater coverage, with high resolution surveys carried out only in regions where significant improvements would be obtained. This approach also provides a tool to help inform management decisions by reducing the probability that a scale mismatch will occur between resolution of environmental mapping surveys and biological data collection. It also allows a quantitative comparison of the information obtained for a given environmental variable measured using different sampling approaches. These techniques were successfully applied to two distinct habitats using differing sampling schemes (photographs, videos, bathymetry and backscatter data were employed), demonstrating their generality and potential for implementation in other environments.

Both Whittard Canyon and Rockall Bank are complex deep-sea systems exhibiting high spatial heterogeneity over different scales and as such provided excellent areas to examine the usefulness of an information theory approach to address questions of resolution choice using benthic imagery and two acoustic mapping techniques. The information theory approach applied here on the multiple scale datasets indicates that ship-based acoustic mapping of Whittard Canyon at resolutions of 20-50m was adequate to capture the trends in the spatial distribution of biological characteristics, but that much finer resolutions (<2-5m) are needed for a sediment heterogeneous area such as Rockall Bank. However, more benthic imagery would be required before predictive maps of species assemblages could be built for Whittard Canyon and for these, higher resolution bathymetry may be highly valuable. The highest resolution datasets may not always be needed to produce effective predictive maps and in the interest of covering large regions of conservation importance, smaller nested surveys providing *a priori* information can be invaluable in determining the most appropriate mapping scale.

5.7 Acknowledgments

5.8 Supporting Information

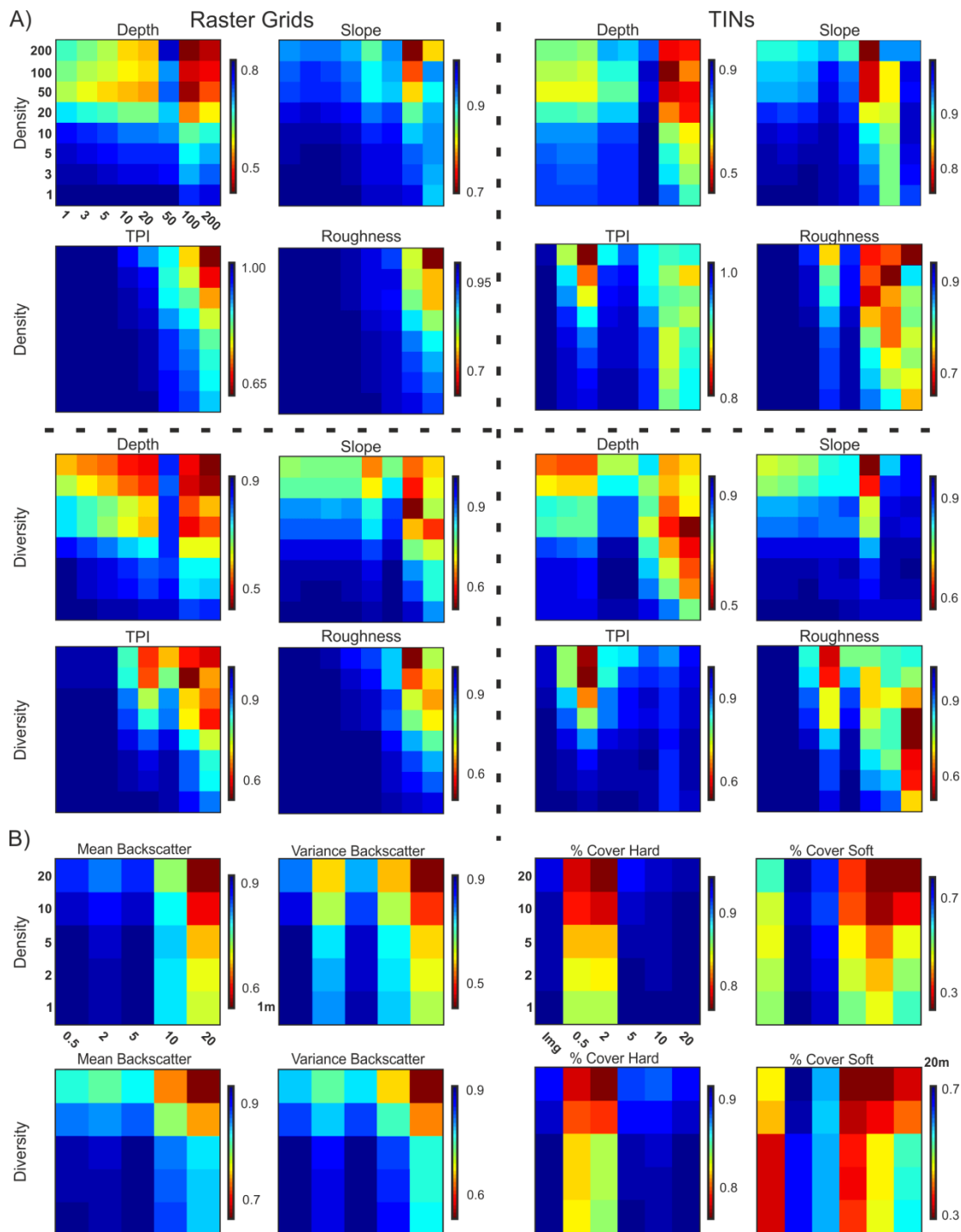### 5.8.1 Divergence under Negative Associations

Figure S 5.1: Normalized Kullback-Leibler divergence between the response variables (density and diversity, histogram bins from highest to lowest values) and selected environmental descriptors across scales for A) Whittard Canyon (raster grids on the right and TINs on the left) and B) Rockall Bank. The same resolutions as for Figure 5.6 are employed. Small divergences are show by warm colours and indicate more similar PDFs.

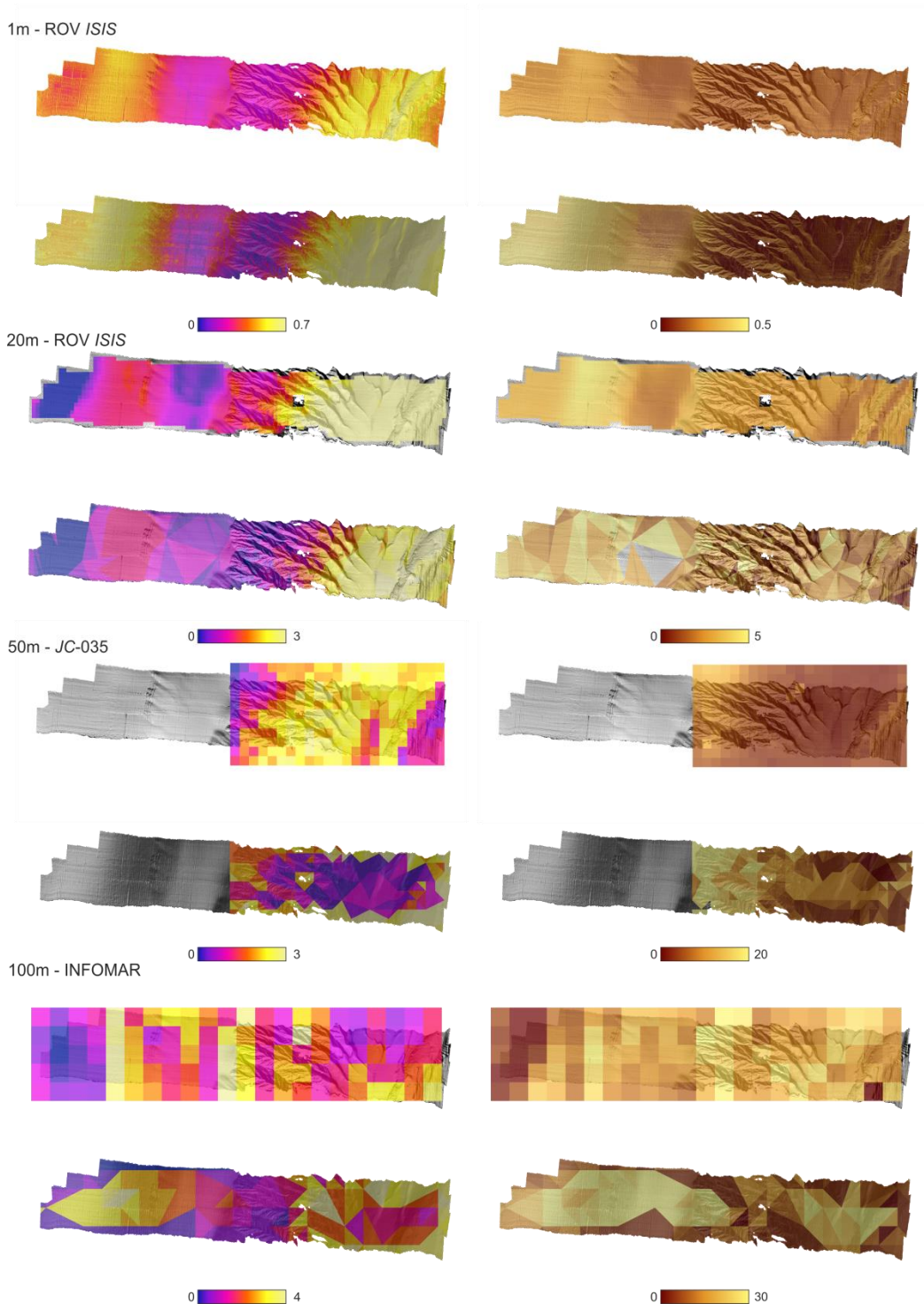### 5.8.2   Prediction Comparison for Whittard Canyon

Figure S5.2: Prediction comparison for dive 108 based on varying resolutions for raster grids (top) and TINs (bottom) for Simpson's reciprocal index of diversity (left) and density (right) overlain on hillshade built from the 1m bathymetry.
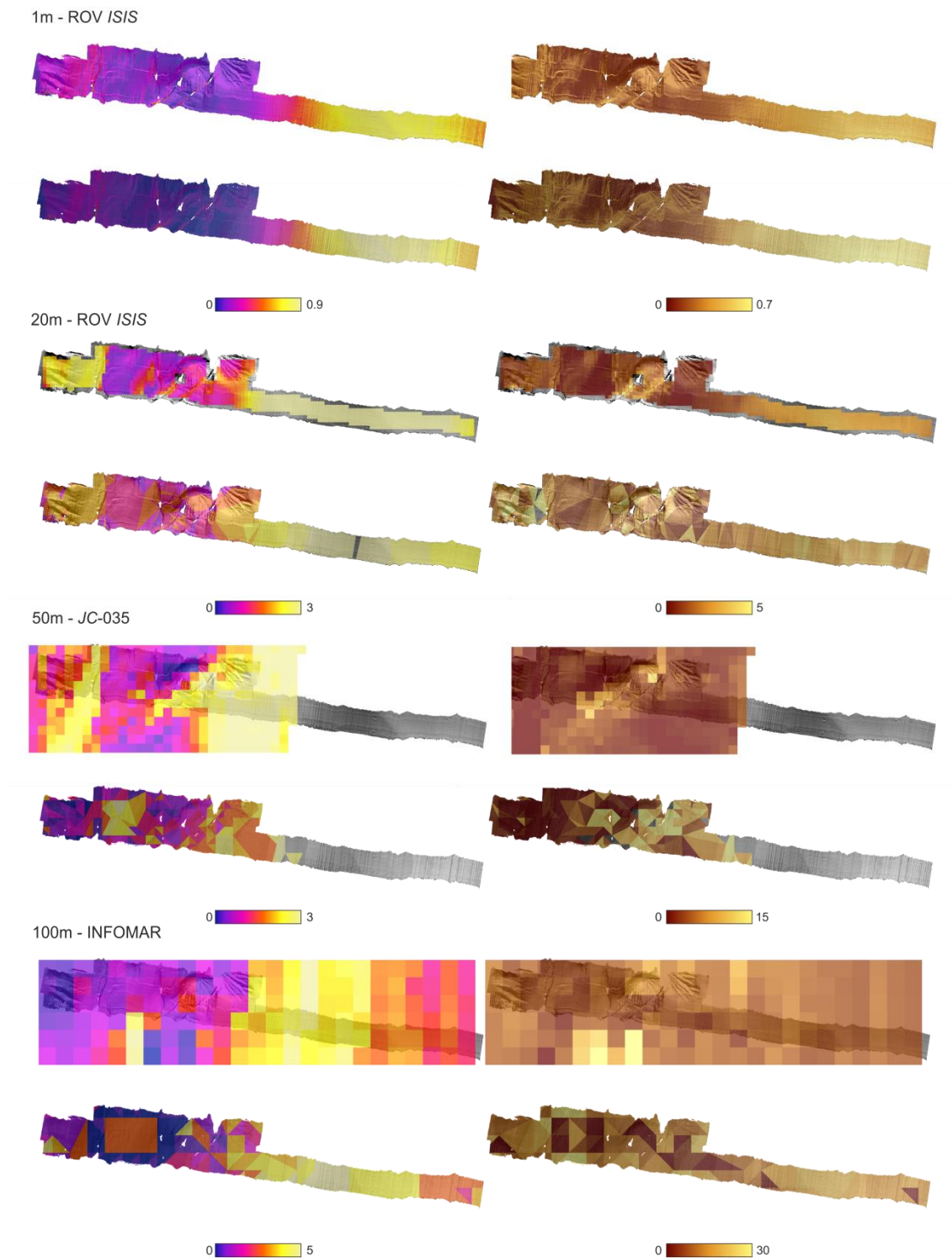
Figure S5.3: Prediction comparison for dive 109 based on varying resolutions for raster grids (top) and TINs (bottom) for Simpson's reciprocal index of diversity (left) and density (right) overlain on hillshade built from the 1m bathymetry.

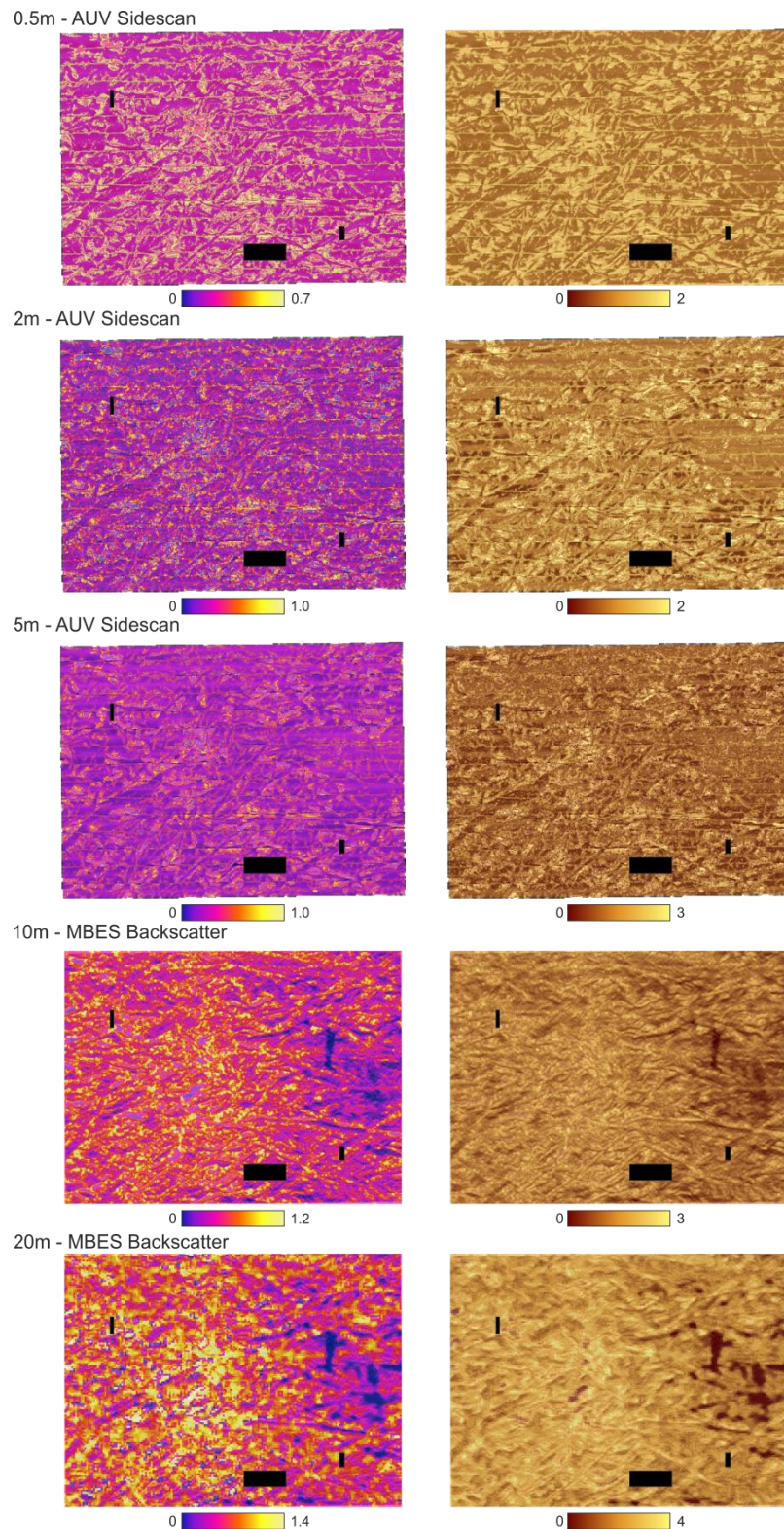### 5.8.3   Prediction Comparison for Rockall Bank

Figure S5.4: Prediction comparison for Rockall Banks AUV mission 44 based on varying resolutions for Shannon's diversity index (left) and density (right) overlain on 0.5m sidescan sonar layer. Black rectangles show areas for which no data are available.
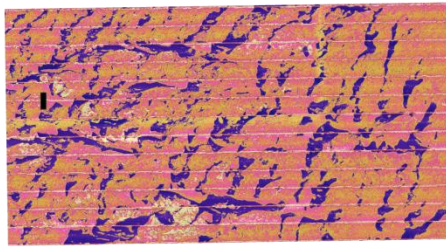
181

Figure S5.5: Prediction comparison for Rockall Banks AUV mission 45 based on varying resolutions for Shannon's diversity index (left) and density (right) overlain on 0.5m sidescan sonar layer. Black rectangles show areas for which no data are available.
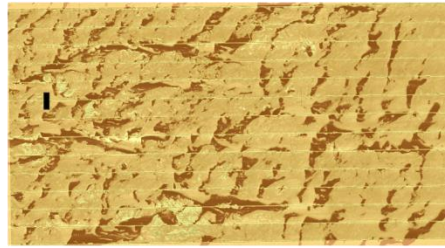
Chapter 6:
Summary and Conclusions

By taking advantage of the links between marine biodiversity and habitat heterogeneity, abiotic proxies can provide direct applications for the management of natural resources by establishing adequate representations of biotic components via efficient survey techniques such as acoustic mapping. With increasing anthropogenic excursions into deeper marine environments, through activities such as deep-sea fishing or seabed mining, there is an increasing need to improve further our ability to produce seafloor habitat maps at multiple scales. However, there often exists a disconnect between the scales at which ground-truthing occurs and the resolution and extent of the acoustic data. Increasingly high resolutions are now available, but as the areas covered often remain limited, tools are needed to integrate better the multi-disciplinary datasets acquired over multiple scales in order to support marine spatial planning. This thesis examined and compared a wide range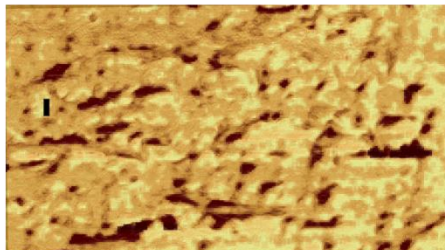 of statistical approaches to habitat mapping in complex deep-sea settings in order to increase the amount of information that can be extracted from acoustic seafloor maps and better capture spatial variability in species distributions so as to improve the creation of predictive biological maps.

6.1   Scientific Contributions

This thesis provided five main important contributions: 1) demonstrated the usefulness of landscape indices in improving statistical model performances and 2) highlighted the importance of topographic complexity in structuring benthic spatial patterns, 3) established that more robust predictions could be obtained by applying ensemble modelling approaches, 4) quantified information lost as a result of scale choice and 5) applied these findings to current conservation needs. When viewed in consideration of their implication for wider marine spatial planning, the significance of these results can be understood in the novel techniques put forth and their overall improvement over commonly employed approaches.

Chapter 6 -
Summary and Conclusions

6.1.1   Thesis Objectives

1) Evaluate the potential of landscape indices in increasing the amount of biological variation which can be explained by sediment interpretation maps

The inclusions of class and landscape metrics had previously been employed in terrestrial environments, but their application to the marine environment had remained limited (Wedding et al. 2011). However, the work carried out on Rockall Bank clearly showed their usefulness for predictive habitat mapping in deeper settings. By allowing the spatial patterns in habitat patches of the surrounding areas to be taken into account, variation explained in both biodiversity and species composition could be improved by ~40% when compared to the sole use of bathymetry-derived environmental descriptors. The increased explanatory power provided by the inclusion of landscape metrics in explaining species distribution patterns was also successfully applied to the extrapolation of these species-environment relationships to build full-coverage biological maps for the extent of the acoustic surveys. This was able to help mitigate the loss of variation explained when very fine-scale environmental information derived from the imagery was unavailable over the entire extent.

2) Examine the importance of vertical structures and habitat complexity measures in structuring spatial patterns in benthic communities

In Whittard Canyon, topographic variability had a strong effect on species compositions and distributions. Beta-diversity across the canyon was generally high, with most transects only sharing less than 40% of observed species with other transects. On the other hand, areas of high species richness, abundance or diversity were not common. Steep slopes usually harboured higher diversity, while bathymetric position index (BPI) indicated lower abundance and species richness towards the canyon's thalweg in areas dominated by soft sediments. Detailed biological analysis of the cold-water coral dominated wall found in the eastern branch showed the highest abundance and number of species as compared to any other habitat types encountered anywhere else within the canyon.

3) Find the most appropriate modelling approach to extrapolate localised biological information to the extent covered by the acoustic map, using externally acquired datasets to assess classification and prediction accuracy

On Rockall Bank, three different statistical approaches (redundancy analysis, random forest and maximum entropy), with different strengths, weaknesses and areas of uncertainties were used to build full-coverage predictive maps for both species assemblages and biodiversity. Even though all three techniques gave different outputs, this allowed the resulting maps to be combined, based on prediction variability and model performances, in order to build more robust ensemble prediction and uncertainty maps. When ensemble mapping techniques were implemented in Whittard Canyon (based on predictions from general additive models and random forest), it was possible to pinpoint potential areas of high biological interest and make habitat suitability predictions for cold-water corals.

The usefulness of ensemble maps was clearly illustrated in their improved performances when assessed using datasets independently collected during subsequent cruises. In addition, all statistical techniques presented used well-established criteria for parameter selection, reducing subjectivity and ensuring robustness and repeatability.

4) Quantify the change in information contained in datasets of varying resolutions and its impact on model performances and resulting predictive maps

To improve efficiency in habitat mapping surveys, there remained a strong need to assess how much information was lost or gained as datasets were collected at specific resolutions. Such information was shown to be quantifiable using an information theory approach (using Shannon entropy and Kullback–Leibler divergence). For both Rockall Bank and Whittard Canyon, spatial variation observed in species distributions between the two habitats occurred over different spatial scales. When variable selection was carried out over a range of resolutions, different environmental descriptors were found to be significant predictors of biological characteristics across the scales considered.

On Rockall Bank, past iceberg activity led to a high sediment heterogeneity and higher diversity was usually found associated with mixed sediments, exposed bedrock and live cold-water coral colonies which required fine resolution maps (<2-5m) in order to capture the complex biological spatial patterns. However, even at the highest resolution (0.5m), probability density functions of estimated sediment composition based on sidescan sonar-derived information as opposed to imagery-derived, showed a tendency of the former to underestimate rarer substratum types (such as bedrock or live coral cover). This difference in the quantification of environmental variability resulted in a 20-40% increase in biological variation explained when models included fine-scale imagery-derived information. On the other hand, in Whittard Canyon, sediment composition was more consistent and variables such as depth, slope and BPI appeared as better descriptors of biological variables. Spatial patterns in both density and diversity could adequately be captured up to resolutions of 20-50m.

Time limitations are significant constraints on many marine surveys and there is usually a trade-off between the resolution obtained and the extent covered. The approach developed is a simple and elegant way to examine quantitatively this lost in information as well as to determine the most useful scales at which to further examine the influence of particular environmental descriptors on species distributions in a given system.

5) *Address these findings with respect to current conservation needs and future information requirements*

Through the use of landscape indices and ensemble mapping methods, it was possible to establish broader-scale trends useful in informing management measures for Rockall Bank. The M45 area showed the highest suitability for cold-water coral and predicted diversity, but would only be protected through the implementation of the cSAC. In Whittard Canyon, ensemble mapping was able to identify vertical walls with a high potential to harbour rare, but ecological important species and habitats. As these communities only represent a very small fraction of the canyon's seafloor (<1%), haphazard sampling may be ineffective for localising them. The methods developed can provide a useful guide for future sampling and help in the monitoring of these important habitats. Similarly the use of an information

theory approach illustrated that analysis of a small survey area can help determine how much information is lost when a particular resolution is chosen and provide a quantitative method to select an appropriate sampling resolution for a given system. The application of these techniques to two distinct habitats with differing sampling schemes (photographs, videos, bathymetry and backscatter data were employed) demonstrated their generality and potential for implementation in other environments.

The developed methods are much more involved than other techniques commonly employed in seabed mapping. However, the <u>additional work requires no further costly sampling or access to specialized equipment.</u> The increase in prediction performances obtained demonstrated a clear usefulness for marine spatial planning. Comparison of the statistical approaches clearly showed differences in predictions, but a single approach did not consistently outperform the others when multiple species assemblages or biological variables were considered. As such, continued reliance on a particular statistical method may lead to decisions being biased by the chosen method as opposed to reflecting real characteristics of the dataset. Our results also suggest that comparison between statistical methods showing one method to outperform the others may not always be extendable to other habitats, species or assemblages. Such risk can easily be <u>mitigated by considering multiple statistical methods and building weighted ensemble maps taking into account prediction variability and individual model performances.</u>

Although difficulties are associated with using different imagery collecting platforms, limited access to sampling equipment and cruise opportunities result in the need to develop assessment methods combining different datasets in order to make efficient use of all the information available. <u>Considering multiple resampling schemes when carrying out the model assessment step helps ensure that undue confidence is not given and those areas of particularly high variability or uncertainty are clearly identified.</u> In turn, such areas may be seen as areas for which future sampling would be highly valuable and whose characterisation would greatly improve current knowledge. In any case, predictive maps only represent the current knowledge of a system based on the (often limited) data available and proper depictions of

spatial uncertainty, in the form of confidence maps, are needed to inform the conservation and management process.

6.1.2   Ecological Hypotheses

Hypothesis 1: Areas of higher heterogeneity are expected to provide a higher number of niches and result in higher diversity and species turnover, but which environmental variables may be responsible for the heterogeneity observed at any given scale is not always know a priori. Finer-resolution habitat characterisation over a small extent will cause different environmental variables (sediment composition) to be identified as driving the patterns observed as opposed to coarser-resolution broader-extent surveys (topographic variability).

Regions of higher sediment or topographic heterogeneity tended to harbour higher biodiversity, and as such could be employed as proxies for biodiversity and for the presence of certain species of interest. However, the environmental variables responsible for creating this heterogeneity changed between habitats. On Rockall Bank, combinations of small intermingled patches of different sediment types, providing both soft sediment and hard substratum had a positive effect on biodiversity, particularly when combined with the presence of cold-water corals. On the other hand in Whittard Canyon, depth and topographic heterogeneity resulted in high species turnover, while sedimentary properties did not significantly improve the modelling of biodiversity. At a finer-scale, cold-water coral colonies were often found to harbour specific species assemblages or show higher species richness than other habitats. This may in part be due to the enhanced complexity created by their framework (altering local current dynamics as well as providing additional attachment possibilities), increasing heterogeneity at an even finer scale.

Hypothesis 2: Environmental variables will act over a range of specific scales to drive biological patterns even when derived from the same original raster (bathymetry or backscatter). The influence on species distributions of an environmental variable measured at one scale may show a different response curve when examined at another scale.

Even though sets of environmental variables were derived from the same rasters (bathymetry and backscatter), different spatial scales were needed to capture the variables responsible for the heterogeneity observed and specific variables were found to have a stronger explanatory power at specific scales. Variables describing topographic variability or rarer substratum types needed fine-scale characterisation in order to be useful in modelling biological characteristics while depth was consistently selected regardless of resolution. Although a range of scales were examined, the species-environment response curves obtained tended to remain similar across scales. BPI was measured for a given resolution over multiple scales, but when selected, always showed higher abundance, species richness, diversity and likelihood of coral presence to be associated with topographic highs as opposed to flatter regions. However, certain species-environment response curves (such as percentage of rubble) did show significant unimodal relationships which could be misrepresented if only coarser scale environmental information, which tended to underestimate percentage cover of rarer substratum types, was available. Including other environmental variables (such as water mass properties or disturbance rates) would likely result in a broader range of scales needed to be considered.

Hypothesis 3: *Species, communities and macroecological properties (e.g. abundance, species richness and biodiversity) will vary over different scales even within the same geographical area. Species are likely to require the finest-scale habitat characterisation followed by communities and macroecological properties.*

Even within a single habitat, species, communities and macroecological properties varied over different scales and required environmental information of different resolution. However, not all macroecological properties required environmental information over the same scale. In both Rockall Bank and Whittard Canyon, abundance was found to require finer-scale environmental information than biodiversity in order to be adequately modelled. The inclusion of imagery-derived sediment composition descriptors in Whittard Canyon did not significantly improve the modelling of species diversity, but improved modelling of abundance. On Rockall Bank, the use of coarser shipborne backscatter maps caused models for abundance to exhibit low

explanatory power, but diversity was not similarly affected. Since abundance can be more easily affected by a single species, while diversity requires a change in multiple species, it is likely more sensitive to finer-scale structures. Indeed when examining specific taxa, such as cold-water corals, finer scale environmental variables were found to be important predictors. In turn, the presence of cold-water corals had an effect on various other species, particularly other suspension feeders. Modelling this species assemblage required fine-scale environmental characterisations, while soft sediment associated species assemblages, whose patch size was generally larger, would be expected to be less significantly affected by acquisition of data at coarser resolutions.

6.2   Future Directions

The work carried out illustrated the value of high-resolution maps to capture biological variation. However, as the extents covered by such maps remain limited, tools capturing the high variability of finer-scale datasets in such a way as to allow its inclusion into broader-scale maps, more appropriate for marine spatial planning, still need to be developed. Modeling approaches such as Bayesian hierarchical models show great promise as they allow for prior fine-scale information to be included into spatial models to create full coverage predictions at broader scales (Wilson et al. 2011). Combining top-down and bottom-up approaches to habitat mapping may also improve our ability to provide reliable predictions over multiple-scales. Similarly to how TINs were employed, it would be possible to use 'objects' (areas identified as being of similar characteristics and composition) from automated classification algorithms, as well as their associated descriptors, as the basic shape upon which to create predictions of biological characteristics using bottom-up approaches. Such a technique would combine the broader-scale habitat classification of the acoustic maps, with finer-scale predictors of biotic characteristics while also minimizing computational time.

Much of the current marine mapping being carried out considers seafloor heterogeneity as captured using two-dimensional rasters. In addition to underestimating the vertical structural complexity of the seabed (which was

an important driver of biodiversity in Whittard Canyon), such an approach also does not take into account water column properties, important in regulating bentho-pelagic linkages. As such, incorporating point-cloud data to capture the three dimensional structure and complexity of the ocean is an important future step. Metrics to describe vertical complexity in point-cloud data have started to be developed from LIDAR datasets, and their usefulness in representing forest canopy heterogeneity and predicting passerine species distributions established (Goetz et al. 2010, Zellweger et al. 2013, 2014). The extension of such metrics to the marine environment will be particularly valuable in representing complex vertical structures such as the cold-water coral dominated cliff found in Whittard Canyon (Huvenne et al. 2011).

6.3   Conclusion

Class and landscape metrics, describing the heterogeneity of habitat patches, were found to be useful proxies for biodiversity and species distribution for areas of highly variable sediment composition, while topographic variables, such as slope, rugosity or bathymetric position index, were shown to influence species richness, abundance and biodiversity in morphologically complex areas. However, the choice of which abiotic proxy to employ needs to be carefully considered with respect to scale (both extent studied and resolution mapped), and will also change between ecosystems or when different biological variables are considered. The manner in which these proxies are employed to represent the biological variation is crucial. Many bottom-up statistical approaches to predictive habitat mapping were examined for both univariate and multivariate variables (redundancy analysis, general linear and additive models, random forest and maximum entropy), but the more reliable results were obtained when the strengths of each approach were combined to form ensemble predictions taking into account prediction variability and model performances.

As must be clearly apparent throughout this thesis, various concurrent processes act to drive the spatial patterns observed over multiple scales and hierarchical studies will continue to be needed in order to quantify these complex species-environment relationships and help increase our

understanding of the roles of habitat heterogeneity in maintaining biodiversity. The implementation of the approaches developed was demonstrated in two very different habitats, which indicates that their applications are robust and that their extension to other ecological processes and environments should be feasible. With the current limitations associated with biological sampling, the successful identification of abiotic proxies for biodiversity and their implementation in robust statistical approaches for the creation of full-coverage biotic maps is a crucial step. This will help ensure that we can establish the baseline state of the deep-sea benthos and develop appropriate management and conservation measures before impacts of anthropogenic activities further alters the current state of this ecosystem.

Appendix A :
Image catalogue of morphospecies

Images Deposited Online as part of the SERPENT Media Archive

Site: Rockall Bank

- <u>Sites (2112)</u>
  - <u>Atlantic (1917)</u>
    - **Rockall Bank (79)**

Number of items: **79**.

<u>Actinaria (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 188

<u>Actinauge richardi (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 329

<u>Alcyonacea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 192
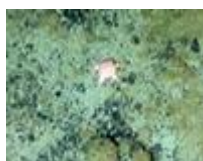

<u>Ascidiacea (1 file)</u>
Site: Atlantic >
Rockall Bank
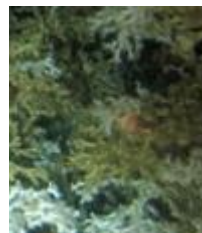Depth (m): 299

<u>Asterias rubens (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 290

<u>Asteroidea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 184

<u>Asteroidea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 192

<u>Asteroidea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 219

<u>Asteroidea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 317

<u>Asteroidea (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 292

<u>Bolocera tuediae (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 219

<u>Brachyura (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 218

<u>Bryozoa (1 file)</u>
Site: Atlantic >
Rockall Bank
Depth (m): 218

Bryozoa (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 295

Ceramaster (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 216

Ceriantharia (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 309

Buccinidae (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 323

Cidaris cidaris (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 220

Cnidaria (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 192

Caridea (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 314

Cnidaria (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 191

Caryophyllia smithii (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 293

Comatulida (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 190

Caryophyllia (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 295

Caryophyllia (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 295

Custacea (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 315

Cephalopoda (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 299

Cyclostomatida (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Demospongiae Orange Encrusting (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 219

Demospongiae Pink Encrusting (1 file)
Site: Atlantic > Rockall Bank
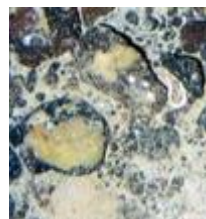Depth (m): 293

Demospongiae Red Encrusting (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 185

Demospongiae White Encrusting (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 187

Demospongiae Yellow Encrusting (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 187

Echinus (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 184

Echinus (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 217

Echnius (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 294

Henricia (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 219

Hippasteria (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 189

Holothuroidea (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 229

Holothuroidea (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 297

Hydrozoa (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 191

Hydrozoa (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 192

Hymedesmia (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 217

Kophobelemnon stelliferum (1 file)
Site: Atlantic > Rockall Bank
Depth (m): 292

Lophelia pertusa (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 218

Munida sarsi (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 217

Nephrops norvegicus (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 315

Ophiuroidae (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 187

Ophiuroidae (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Paguridae (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 293

Parastichopus tremulus (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 218

Paromola cuvieri (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 191

Pennatulacea (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 226

Phelliactis (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 293

Porania pulvillus (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Porifera (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 192

Porifera Brown Lamellate (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 226

Porifera Cup (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 191

Porifera Cup (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 187

Porifera Cup (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 226

Porifera Orange
Branching (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 227

Porifera Orange
Branching (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 227

Porifera Orange
Lobose (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 226

Porifera Orange
Lobose (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 229

Porifera Spherical (1
file)
Site: Atlantic >
Rockall Bank
Depth (m): 192

Porifera Tan
Columnar (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Porifera White
Lamellate (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 190

Porifera Yellow
Branching (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Porifera Yellow
Columnar (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 228

Porifera Yellow
Lobose (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 187

Porifera Yellow
Spherical (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 191

Portunidea (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 217

Reteporella (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 216

Sabellidae (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 217

Sipuncula (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 192

SERPENT

Scientific and environmental ROV partnership using existing industrial technologies
http://archive.serpentproject.com/view/sites/Rockall_Bank.html

Stylaster (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Unidentified (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 219

Unidentified (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 190

Unidentified (1 file)
Site: Atlantic >
Rockall Bank
Depth (m): 229

Scientific and environmental ROV partnership using existing industrial technologies
http://archive.serpentproject.com/view/sites/Rockall_Bank.html

Site: Whittard Canyon

- Sites (2112)
  - Atlantic (1917)
    - **Whittard Canyon (176)**

Number of items: **176**.

Acanthogorgia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1346

Actinaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2932

Actinauge (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1338

Actinernus
michaelsarsi (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1368

Actiniaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3142

Actiniaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1331

Actiniaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1288

Actinoscyphia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1139

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1359

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2746

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1334

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2338

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3118

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1600

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1599

Anthomastus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2831

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1479

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3189

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1344

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2767

Alcyonacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1337

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3245

Anachalypsicrinus (1
file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3118

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3286

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -989

Anthomastus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1710

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1420

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1332

SERPENT

200

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1157

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1341

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1155

Antipatharia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2912

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1065

Araeosoma (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2613

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -962

Ascidiacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2671

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1518

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3302

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2854

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1510

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -551

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1446

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -541

Anthozoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1427

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -544

SERPENT

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1705

Astropectinidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1497

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1274

Atelecrinidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1380

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1135

Benthodytes (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2939

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1109

Benthodytes typica (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2884

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -976

Benthogone rosea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -680

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1608

Benthothuria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3175

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1477

Bivalvia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1342

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1424

Brisingida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3063

Asteroidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1312

Brisingida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1503

Brisingida, 11 arms
(1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2553

Brisingida, 16 arms
(1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1744

Ceriantharia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -844

Chaceon (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -563

Cidaris cidaris (1
file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1083

Cnidaria (1 file)
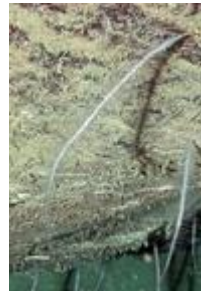Site: Atlantic >
Whittard Canyon
Depth (m): -3602

Cnidaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1368

Cnidaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1727

Cnidaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1006

Cnidaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3119

Cnidaria (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -978

Comatulida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2246

Comatulida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2687

Comatulida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3187

SERPENT

Comatulida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1501

Decapoda (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -731

Comatulida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2185

Deima (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3064

Corallimorpharia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1502

Desmophyllum (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2479

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2597

Echinidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1350

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3045

Echinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2913

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2670

Echinus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1022

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1356

Echinus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1540

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1516

Elpidiidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2395

Crinoidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1343

Endoxocrinus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2784

Enteropneusta (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2780

Epizoanthus paguriphilus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1049

Euplectella (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1576

Galatheidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2720

Galatheidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3286

Galatheidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -650

Galatheidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1315

Galatheidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2844

Gastropoda (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2592

Geodia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1309

Geodia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1305

Gorgonocephalidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1316

Hermit Crab (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -543

Histocidaris purpurata (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1253

Holothuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2912

Holothuroidea (1 file)
Site: Atlantic >
Whittard Canyon

Depth (m): -3436

Holothuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3163

Holothuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1419

Holothuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2814

Hyalonema apertum (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1419

Hymenaster (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2671

Isididae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1396

Isididae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1322

Koehlermetra (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1346

Leiopathes (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3119

Lophelia pertusa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1346

Madrepora oculata (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -989

Nemertea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1109

Nephrops (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1058

Nymphaster (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -891

Octacnemidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2725

Octopod (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1266

Octopodidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -992

Octopodidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1778

Octopodidae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2687

Ophiuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3469

Ophiuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1864

Ophiuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3010

Ophiuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1685

Ophiuroidea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2685

Parapagurus
pilosimanus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1050

Paroriza pallens (1
file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3064

Peniagone (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3082

Pennatula aculeata
(1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -917

Pennatulacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3245

Pennatulacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2685

Pennatulacea (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1457

Pentametrocrinus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1329

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3043

Phelliactis (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3175

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3175

Phormosoma
placenta (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1407

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1496

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1359

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1479

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1337

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1283

Porphyrocrinus (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1502

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1368

Primnoa (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1646

Porifera (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3051

Psychropotes
longicauda (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3584

Pycnogonida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1445

Scleractinia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2744

Scleractinia (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1325

Shrimp (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2566

Shrimp (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2687

Solenosmilia
variabilis (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1744

Spatangoida (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1061

Stylasteridae (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1388

Tromikosoma (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2559

Umbellula (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3286

Umbellula (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3286

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -215

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2449

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -2599

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1686

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1685

Unknown (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1344

SERPENT

Unknown star (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -3250

Xenophyophore (1 file)
Site: Atlantic >
Whittard Canyon
Depth (m): -1450

Appendix B :
Acoustics mapping of an abyssal hill

Contribution to:

Methods

Data processing

In July 2012 during the RRS *Discovery* cruise 377, as part of the AESA project (Autonomous Ecological Surveying of the Abyss, http://picturingthedeep.noc.ac.uk) a hill located in the Porcupine Abyssal Plain was mapped using the autonomous underwater vehicle (AUV) Autosub6000 (Figure B.1). A Kongsberg Simrad EM2000 echosounder (111 beams, operating at 200kHz) mounted on the AUV was used to collect bathymetry and backscatter data. Flying at 100m above the seabed, eight ~10km long swaths separated by 310m were collected with ~20-30% overlap for a total coverage area of ~28.8km^2. AUV position was recorded using range-only navigation algorithms and inertial navigation (McPhail 2009).

To create the bathymetric map, the raw data files were imported into CARIS HIPS and SIPS for visualization and editing. Large offsets in navigation between missions were observed and the following corrections were applied using a UTM projection, Zone 28N with datum WGS84: 105m (y-direction) and 40m (x-direction) for mission 052 and, 250m (y direction) and 30m (x-direction) for mission 056. The offshore tidal computation software POLPRED (Proudman Oceanographic Laboratory, NERC, UK) was used to predict and correct for tidal variations while sound velocity corrections were applied based on the profile obtained from a nearby CTD cast (49˚1.3981, 16˚34.2395). The AUV's time-stamped CTD depths were entered as a delta draft correction. The

final map was created using a BASE (Bathymetric Associated with Statistical Error) surface over a 10m grid.

The backscatter data were derived from the multibeam survey and processed using the PRISM (Processing of Remotely-sensed Imagery for Seafloor Mapping) software, version 5 (Le Bas & Hühnerbach 1998). Mosaicking was carried out in the image processing software ERDAS Imagine version 10.0 with lower backscatter values shown in darker shades. Raster layers (WGS 1984 UTM Zone 28N coordinate system) for use in ESRI ArcGIS were created for both bathymetry and backscatter (Figure B.1).

Habitat Maps

A top-down approach to define broad regions of topographical similarity was carried out following the method described in Verfaillie et al. (2009). Bathymetry-derived layers including depth, slope, standard deviation of slope, eastness, northness, general curvature (calculated based on a 3x3 pixel window), surface-area ratio and bathymetric position index (BPI, calculated at two scales, 100m and 500m) were created. A similar set of environmental layers (BPI was calculated at over 600m scale) was also derived from ship-borne multi-beam bathymetry acquired over the larger surrounding area (880km^2) at 60m resolutions during the *JC*-071cruise.

A principal component analysis (PCA) was carried out to reduce the number of variables: the first 4 axis were retained, representing 82.8% (D377) and 99.9% (JC071) of the variation. The first three axes were converted to RGB images to illustrate gradients in environmental conditions (Figure B.2). To provide a hard classification, fuzzy-clustering using the kmean algorithm and the Calinski- Harabasz criterion were carried out to determine the optimal number of classes and assign each pixel to its most likely class (Lucieer & Lucieer 2009) (Figure B.2).
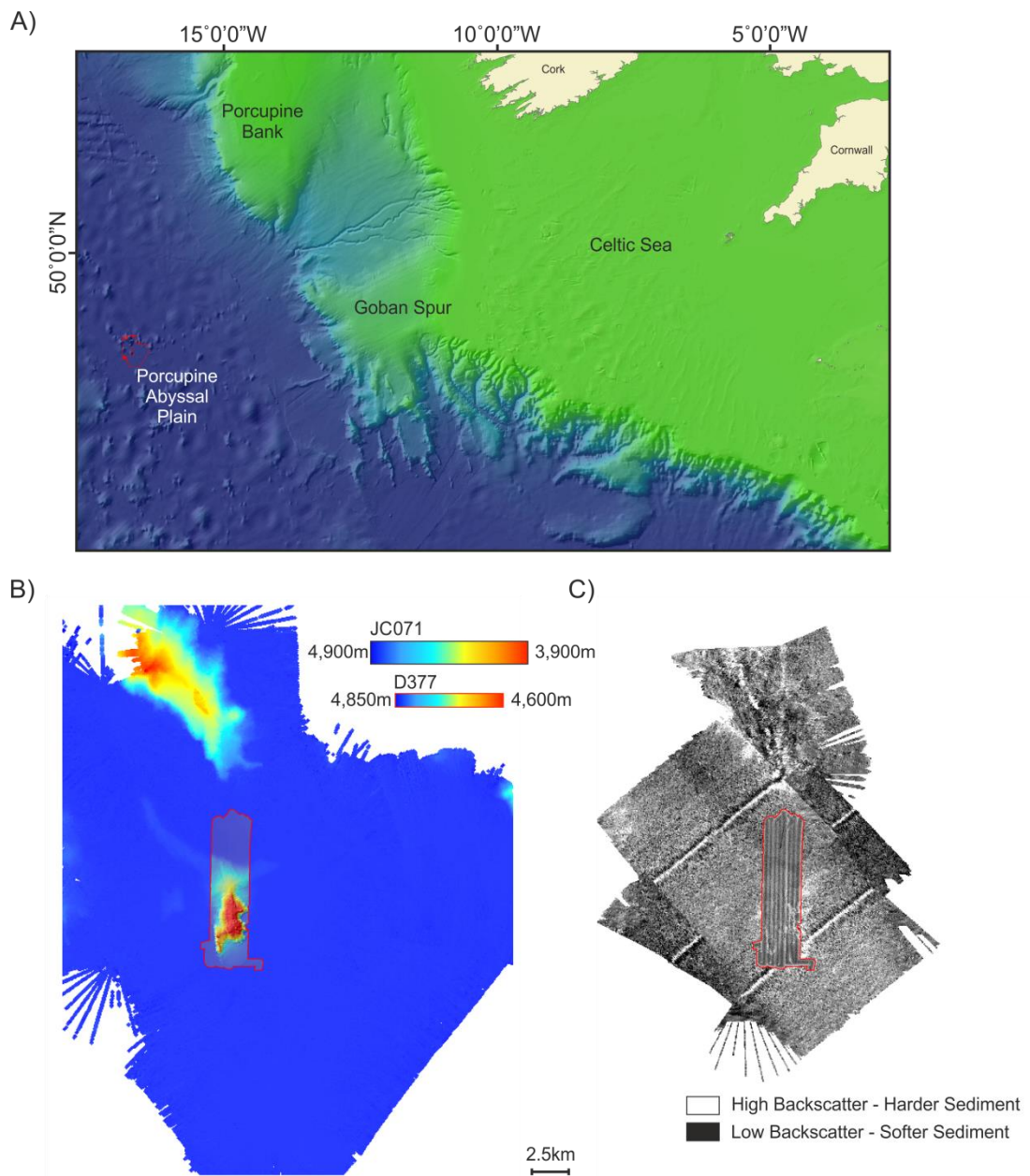
Figure B.1 Maps representing the AESA survey area A) Location in relation to Ireland and the UK, background bathymetry from GEBCO (http://www.gebco.net/). B) Multibeam bathymetry and C) backscatter of the area, larger extent representing the *JC*-071 shipboard data and small one the *D*-377 AUV-based data (outlined in red).

A)

B)
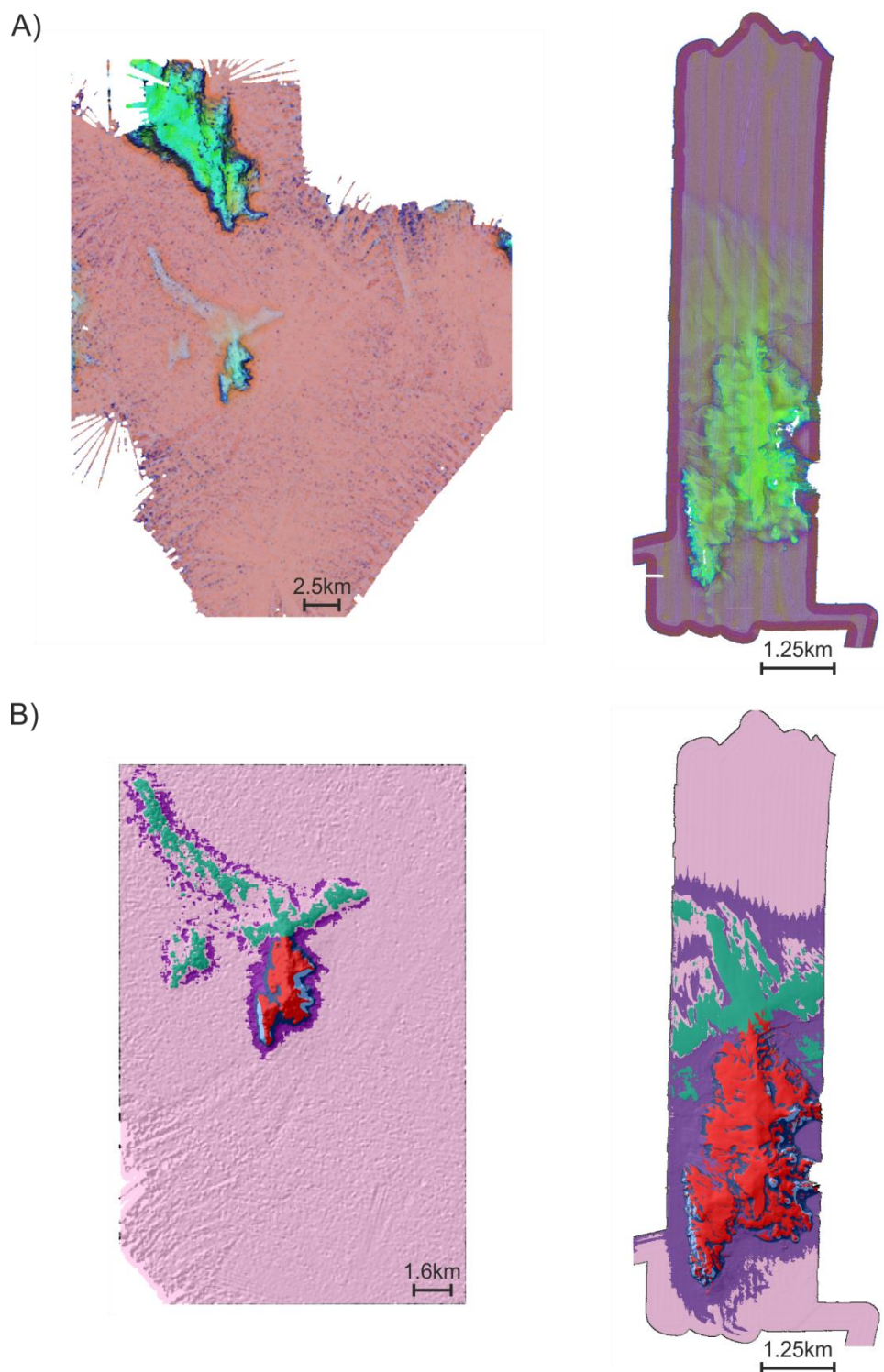
Figure B.2: Habitat classification of the abyssal hill A) RGB representation of the first three axis of the principal component analysis and B) results of the fuzzy clustering into six classes, *JC*-071 (left) and *D*-377 (right).

Appendix C :
Habitat suitability maps for cold-water corals in the Cap the Creus Canyon

Contribution to:

Lo Iacono C, **Robert K**, Gonzalez-Villanueva R, Gori A, Orejas C and Gili JM (*in prep*) Fine-scale ensembles for predictive mapping of CWC populations in the Cap de Creus Canyon (NW Mediterranean). *Remote Sensing and Environment*

Methods

Predictive Habitat Mapping

A 5m resolution bathymetric map of the Cap the Creus Canyon was collected in 2004 by Fugro N.V., AOA Geophysics, and the University of Barcelona, and in 2010 by the Instituto de Ciencias del Mar (ICM-CSIC) as part of the INDEMARES Project (Lo Iacono et al. 2012) (Figure C.1). It was employed to derive environmental layers: slope, rugosity, northness and eastness. Cold-water coral presence for *Madrepora oculata*, *Lophelia pertusa* and *Dendrophyllia cornigera* was determined through video analysis of 8 transects from the manned-submersible JAGO, carried out by A. Gori

A)

B)
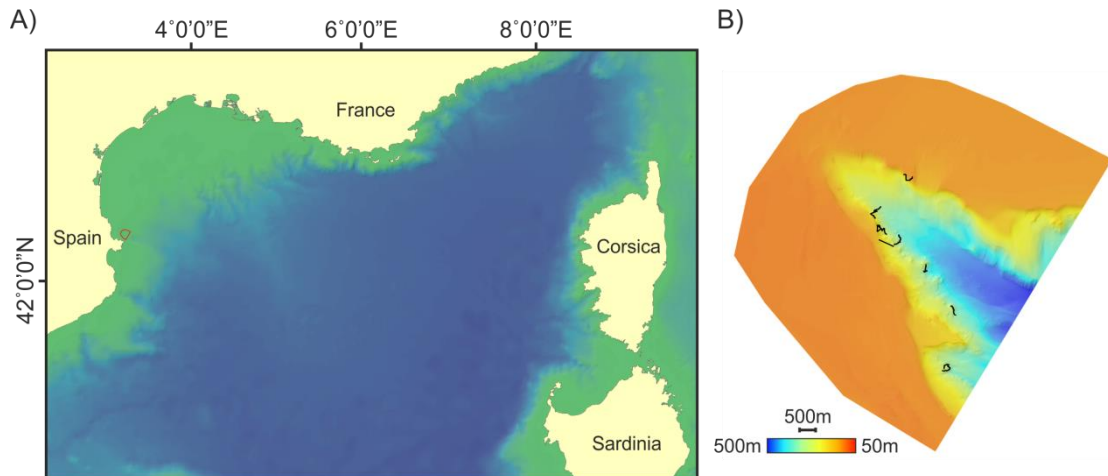
Figure C.1: Cap the Creus Canyon survey A) Location within the Mediterranean Sea, background bathymetry from GEBCO (http://www.gebco.net/).   B) Multibeam bathymetry and ROV transects.

In order to compare habitat suitability predictions with previously produced maps using MaxEnt (a presence-only modelling approach), two presence-absence techniques were additionally considered: 'General Additive Models' (GAMs) with binomial family and 'Random Forest' (RF). In addition to generating predictions at the original resolution (5m) (Figure C.2), maps were created using coarsened data (10m, 20m and 50m) in order to evaluate the influence of resolution on prediction outputs (Figure C.3).

Madrepora oculata
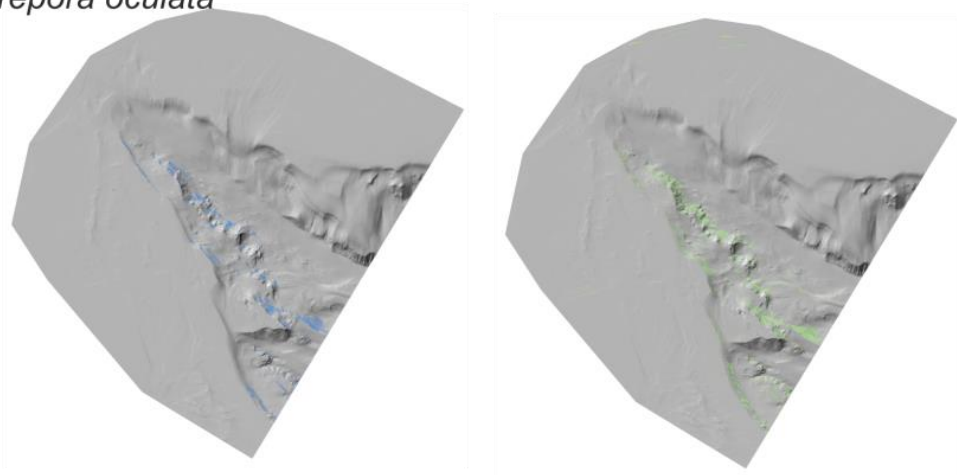
Lophelia pertusa
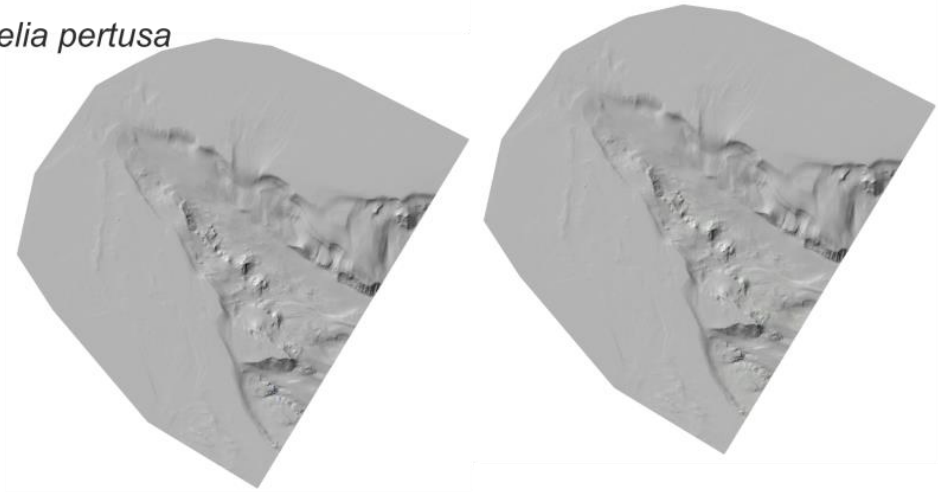
Dendrophyllia cornigera
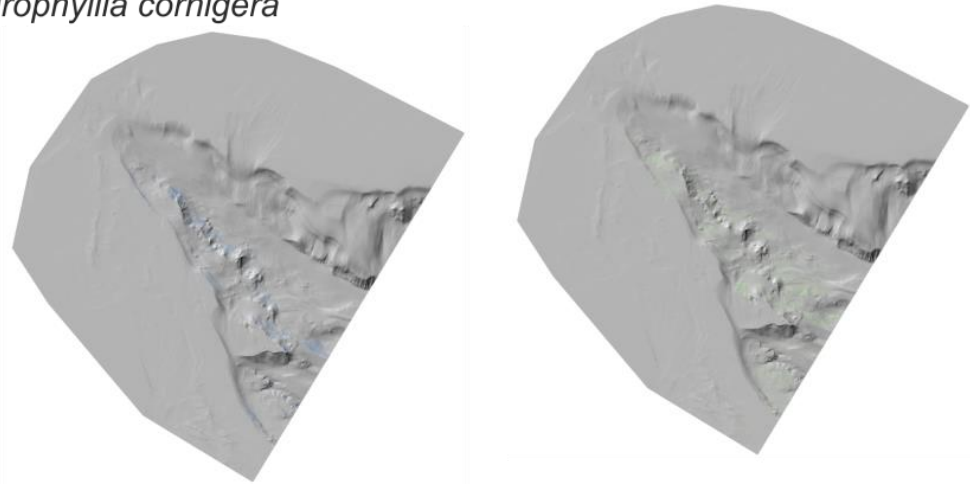
General Additive Model   Random Forest

Figure C.2: Comparison of habitat suitability predictions based on 5m resolution bathymetry between Random Forest and General Additive Models for three cold-water coral species

Madrepora oculata
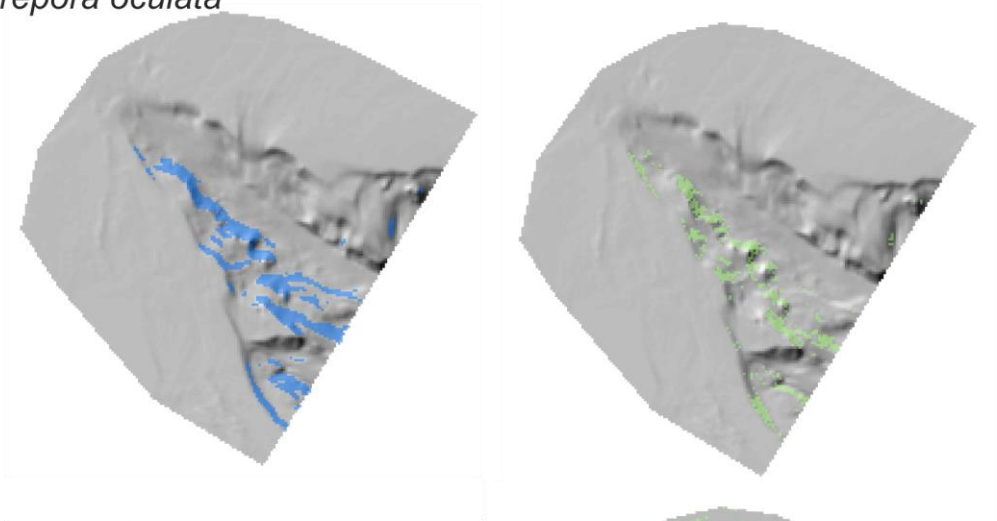
Lophelia pertusa
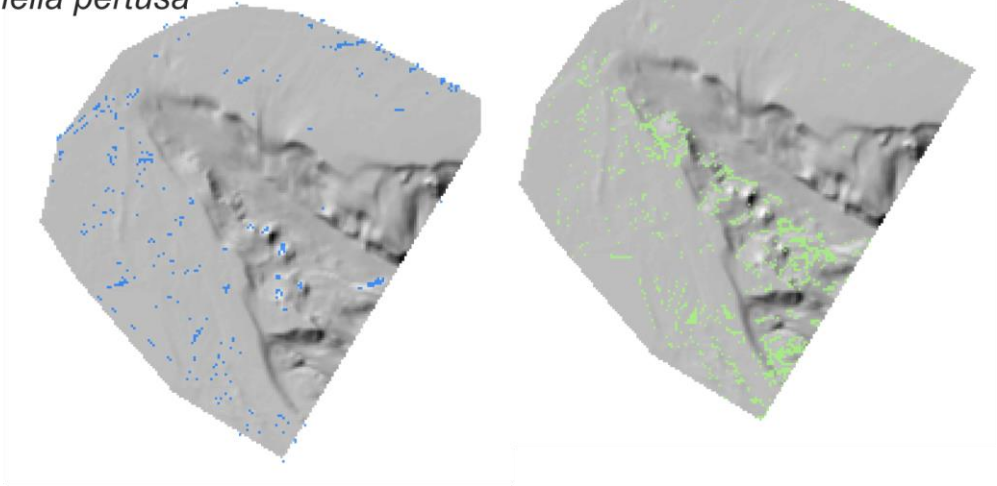
Dendrophyllia cornigera
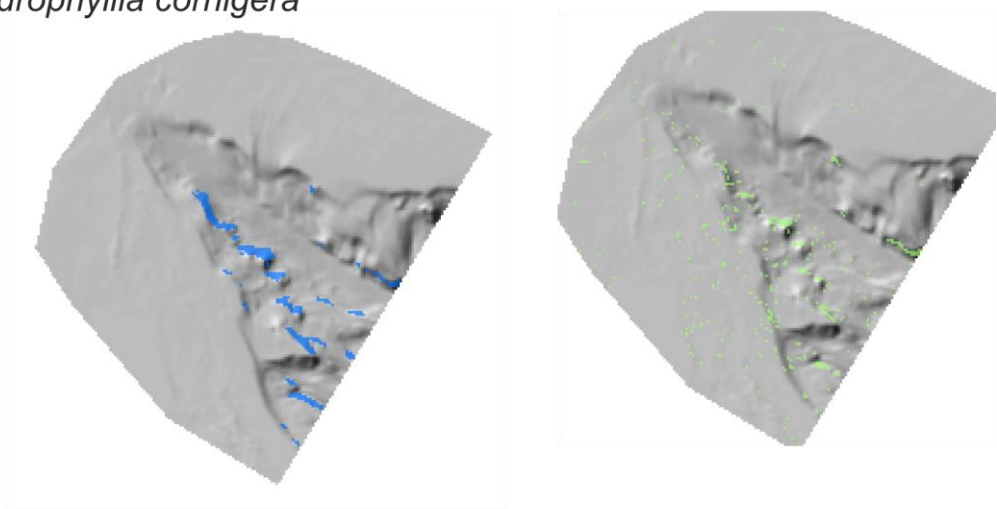
General Additive Model　　　　　Random Forest

Figure C.3: Comparison of habitat suitability predictions based on 50m resolution bathymetry between Random Forest and General Additive Models for three cold-water coral species

List of References

Able K, Twichell D, Grimes C and Jones R (1987) Sidescan sonar as a tool for detection of demersal fish habitats. Fishery Bulletin 84:725-737

Anderson MJ and Willis TJ (2003) Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. Ecology 84:511-525

Anderson TJ, Syms C, Roberts DA and Howard DF (2009) Multi-scale fish-habitat associations and the use of habitat surrogates to predict the organisation and abundance of deep-water fish assemblages. Journal of Experimental Marine Biology and Ecology 379:34-42

Araújo MB and New M (2007) Ensemble forecasting of species distributions. Trends in Ecology & Evolution 22:42-47

Araújo MB, Pearson RG, Thuiller W and Erhard M (2005) Validation of species-climate impact models under climate change. Global Change Biology 11:1504-1513

Archambault P and Bourget E (1996) Scales of coastal heterogeneity and benthic intertidal species richness, diversity and abundance. Marine Ecology Progress Series:111-121

Auster PJ, Lewis RS, Wahle LC, Babb IG and Malatesta RJ (1995) The use of side-scan sonar for landscape approaches to habitat mapping. In: O'Connell T, Wakefield W (eds) Applications of Side-scan Sonar and Laser-line Systems in Fisheries Research. Alaska Department of Fish and Game, Juneau, USA, p 1-7

Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecological Modelling 200:1-19

Baccini A, Friedl MA, Woodcock CE and Warbington R (2004) Forest biomass estimation over regional scales using multisource data. Geophysical Research Letters 31:L10501

Baker K, Wareham V, Snelgrove P, Haedrich R, Fifield D, Edinger E and Gilkinson K (2012) Distributional patterns of deep-sea coral assemblages in three submarine canyons off Newfoundland, Canada. Marine Ecology Progress Series 445:235-249

Bartha S, Collins S, Glenn S and Kertész M (1995) Fine-scale spatial organization of tallgrass prairie vegetation along a topographic gradient. Folia Geobotanica 30:169-184

References

Bellier E, Monestiez P, Durbec J-P and Candau J-N (2007) Identifying spatial relationships at multiple scales: Principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. Ecography 30:385-399

Bett BJ (2001) UK Atlantic margin environmental survey: Introduction and overview of bathyal benthic ecology. Continental Shelf Research 21:917-956

Billett DSM, Lampitt RS, Rice AL and Mantoura RFC (1983) Seasonal sedimentation of phytoplankton to the deep-sea benthos. Nature 302:520-522

Blondel P and Gómez Sichi O (2009) Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf. Applied Acoustics 70:1288-1297

Borcard D, Gillet F and Legendre P (2011) Numerical ecology with R. Springer, New York

Borcard D and Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51-68

Bourillet J-F, Reynaud J-Y, Baltzer A and Zaragosi S (2003) The 'Fleuve Manche': The submarine sedimentary features from the outer shelf to the deep-sea fans. Journal of Quaternary Science 18:261-282

Breiman L (2001) Random forests. Machine Learning 45:5-32

Brotons L, Thuiller W, Araújo MB and Hirzel AH (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27:437-448

Brown C, Hewer A, Meadows W, Limpenny D, Cooper K, Rees H and Vivian C (2001) Mapping of gravel biotopes and an examination of the factors controlling the distribution, type and diversity of their biological communities, Centre for Environment, Fisheries and Aquaculture Science, Lowestoft

Brown CJ, Sameoto JA and Smith SJ (2012) Multiple methods, maps, and management applications: Purpose made seafloor maps in support of ocean management. Journal of Sea Research 72:1-13

Brown CJ, Smith SJ, Lawton P and Anderson JT (2011) Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuarine, Coastal and Shelf Science 92:502-520

Brunsell NA (2010) A multiscale information theory approach to assess spatial-temporal variability of daily precipitation. Journal of Hydrology 385:165-172

Brunsell NA and Anderson MC (2011) Characterizing the multi-scale spatial structure of remotely sensed evapotranspiration with information theory. Biogeosciences 8:2269-2280

Brunsell NA, Ham JM and Owensby CE (2008) Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tall-grass prairie environment. Remote Sensing of Environment 112:2977-2987

Brunsell NA and Young CB (2008) Land surface response to precipitation events using MODIS and NEXRAD data. International Journal of Remote Sensing 29:1965-1982

Buhl-Mortensen L and Mortensen PB (2004) Symbiosis in deep-water corals. Symbiosis 37:33-61

Buhl-Mortensen L, Mortensen PB, Dolan M, Dannheim J, Bellec V and Holte B (2012) Habitat complexity and bottom fauna composition at different scales on the continental shelf and slope of northern Norway. Hydrobiologia 685:191-219

Buhl-Mortensen L, Vanreusel A, Gooday AJ, Levin LA, Priede IG, Mortensen PB, Gheerardyn H, King NJ and Raes M (2010) Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. Marine Ecology 31:21-50

Cain SA (1938) The species-area curve. American Midland Naturalist 19:573-581

Chave J (2013) The problem of pattern and scale in ecology: What have we learned in 20 years? Ecology Letters 16:4-16

Clark MR, Tittensor D, Rogers AD, Brewin P, Schlacher T, Rowden A, Stocks K and Consalvey M (2008) Seamounts, deep-sea corals and fisheries: Vulnerability of deep-sea corals to fishing on seamounts beyond areas of national jurisdiction. Census of Marine Life, UNEP-WCMC, Cambridge, UK

Clarke KR and Warwick RM (2001) Change in marine communities: An approach to statistical analysis and interpretation (2nd ed). PRIMER-6, Plymouth

Clements GR, Rayan DM, Aziz SA, Kawanishi K, Traeholt C, Magintan D, Yazi MFA and Tingley R (2012) Predicting the distribution of the Asian tapir in Peninsular Malaysia using maximum entropy modeling. Integrative Zoology 7:400-406

Colman J (1933) The Wood's Hole Oceanographic Institution: Sonic soundings, and an account of a winter cruise to the Somers Islands. Geographical Journal 82:326-336

Commito JA and Rusignuolo BR (2000) Structural complexity in mussel beds: The fractal geometry of surface topography. Journal of Experimental Marine Biology and Ecology 255:133-152

Compton TJ, Bowden DA, Pitcher CR, Hewitt JE and Ellis N (2013) Biophysical patterns in benthic assemblage composition across contrasting continental margins off New Zealand. Journal of Biogeography 40:75-89

References

Connor DW, Allen JH, Golding N, Howell KL, Lieberknecht LM, Northen KO and Reker JB (2004) Marine habitat classification for Britain and Ireland, Version 04.05. Joint Nature Conservation Committee, Peterborough, www.jncc.gov.uk/MarineHabitatClassification

Costello M, McCrea M, Freiwald A, Lundälv T, Jonsson L, Bett B, Weering TE, Haas H, Roberts JM and Allen D (2005) Role of cold-water *Lophelia pertusa* coral reefs as fish habitat in the NE Atlantic. In: Freiwald A, Roberts JM (eds) Cold-Water Corals and Ecosystems. Springer Berlin Heidelberg, p 771-805

Courtney R and Shaw J (2000) Multibeam bathymetry and backscatter imaging of the Canadian continental shelf. Geoscience Canada 27:31-42

Crawley M (2007) The R Book John Wiley & Sons Ltd, Chichester, UK

Crawley MJ (2005) Statistics: An introduction using R. John Wiley & Sons, Ltd, Chichester, UK

Currie DR and Sorokin SJ (2014) Megabenthic biodiversity in two contrasting submarine canyons on Australia's southern continental margin. Marine Biology Research 10:97-110

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J and Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783-2792

Davies AJ and Guinotte JM (2011) Global habitat suitability for framework-forming cold-water corals. PLoS ONE 6:e18483

Davies JS, Howell KL, Stewart HA, Guinan J and Golding N (2014) Defining biological assemblages (biotopes) of conservation interest in the submarine canyons of the South West Approaches (offshore United Kingdom) for use in marine habitat mapping. Deep Sea Research Part II: Topical Studies in Oceanography 104:208-229

De Leo FC, Smith CR, Rowden AA, Bowden DA and Clark MR (2010) Submarine canyons: Hotspots of benthic biomass and productivity in the deep sea. Proceedings of the Royal Society B: Biological Sciences 277:2783-2792

De Leo FC, Vetter EW, Smith CR, Rowden AA and McGranaghan M (2014) Spatial scale-dependent habitat heterogeneity influences submarine canyon macrofaunal abundance and diversity off the Main and Northwest Hawaiian Islands. Deep Sea Research Part II: Topical Studies in Oceanography 104:267–290

De Mol L, Van Rooij D, Pirlet H, Greinert J, Frank N, Quemmerais F and Henriet J-P (2011) Cold-water coral habitats in the Penmarc'h and Guilvinec Canyons (Bay of Biscay): Deep-water versus shallow-water settings. Marine Geology 282:40-52

de Moustier C (1986) Beyond bathymetry: Mapping acoustic backscattering from the deep seafloor with Sea Beam. The Journal of the Acoustical Society of America 79:316-331

Diaz JVM (2000) Analysis of multibeam sonar data for the characterization of seafloor habitats. University of New Brunswick

Dimitriadou E, Dolničar S and Weingessel A (2002) An examination of indexes for determining the number of clusters in binary data sets. Psychometrika 67:137-159

Dolan MFJ, Grehan AJ, Guinan JC and Brown C (2008) Modelling the local distribution of cold-water corals in relation to bathymetric variables: Adding spatial context to deep-sea video data. Deep Sea Research Part I: Oceanographic Research Papers 55:1564-1579

Dray S, Legendre P and Peres-Neto PR (2006) Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196:483-493

Dray S, Pélissier R, Couteron P, Fortin MJ, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, De Cáceres M, Dufour AB, Heegaard E, Jombart T, Munoz F, Oksanen J, Thioulouse J and Wagner HH (2012) Community ecology in the age of multivariate multiscale spatial analysis. Ecological Monographs 82:257-275

Duffy GA, Lundsten L, Kuhnz LA and Paull CK (2014) A comparison of megafaunal communities in five submarine canyons off Southern California, USA. Deep Sea Research Part II: Topical Studies in Oceanography 104:259-266

Dufrêne M and Legendre P (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. Ecological Monographs 67:345-366

Duineveld G, Lavaleye M, Berghuis E and de Wilde P (2001) Activity and composition of the benthic fauna in the Whittard Canyon and the adjacent continental slope (NE Atlantic). Oceanologica Acta 24:69-83

Dullo W-C, Flögel S and Rüggeberg A (2008) Cold-water coral growth in relation to the hydrography of the Celtic and Nordic European continental margin. Marine Ecology Progress Series:165-176

Dungan JL, Perry JN, Dale MRT, Legendre P, Citron-Pousty S, Fortin MJ, Jakomulska A, Miriti M and Rosenberg MS (2002) A balanced view of scale in spatial statistical analysis. Ecography 25:626-640

Dunton KH, Schonberg SV and Schell DM (1983) Geophysical and biological reconnaissance of rock habitats in Western Camden Bay, Beaufort Sea, Alaska, Institute of Water Resources, University of Alaska, Fairbanks, USA

Duros P, Fontanier C, Metzger E, Pusceddu A, Cesbron F, de Stigter HC, Bianchelli S, Danovaro R and Jorissen FJ (2011) Live (stained) benthic foraminifera in the Whittard Canyon, Celtic margin (NE Atlantic). Deep Sea Research Part I: Oceanographic Research Papers 58:128-146

References

Edinger EN, Sherwood OA, Piper DJW, Wareham VE, Baker KD, Gilkinson KD and Scott DB (2011) Geological features supporting deep-sea coral habitat in Atlantic Canada. Continental Shelf Research 31:S69-S84

Edsall TA, Poe TP, Nester RT and Brown CL (1989) Side-scan sonar mapping of Lake Trout spawning habitat in Northern Lake Michigan. North American Journal of Fisheries Management 9:269-279

Elith J and Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677-697

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE and Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43-57

Etter RJ and Mullineaux L (2001) Chapter 14: Deep-Sea communities. In: Bertness M, Gaines S, Hay M (eds) Marine Community Ecology. Sinauer Associates, Sunderland, MA, p 367-393

Fader GB and King LH (1981) A reconnaissance study of the surficial geology of the Grand Banks of Newfoundland, Ottawa, Canada, p 45-56

Farr HK (1980) Multibeam bathymetric sonar: Sea Beam and hydro chart. Marine Geodesy 4:77-93

Ferrier S and Guisan A (2006) Spatial modelling of biodiversity at the community level. Journal of Applied Ecology 43:393-404

Fielding AH and Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38-49

Flach E, Lavaleye M, de Stigter H and Thomsen L (1998) Feeding types of the benthic community and particle transport across the slope of the N.W. European continental margin (Goban Spur). Progress in Oceanography 42:209-231

Fortin M-J, Drapeau P and Legendre P (1989) Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83:209-222

Fosså JH, Mortensen PB and Furevik DM (2002) The deep-water coral *Lophelia pertusa*; in Norwegian waters: Distribution and fishery impacts. Hydrobiologia 471:1-12

Freiwald A, Helge Fosså J, Grehan A, Koslow T and Roberts JM (2004) Cold-water coral reefs, UNEP/WCMC, Cambridge, UK

Freiwald A, Wilson JB and Henrich R (1999) Grounding Pleistocene icebergs shape recent deep-water coral reefs. Sedimentary Geology 125:1-8

Frutos I and Sorbe JC (2014) Bathyal suprabenthic assemblages from the southern margin of the Capbreton Canyon ("Kostarrenkala" area), SE Bay of Biscay. Deep Sea Research Part II: Topical Studies in Oceanography 104:291-309

Gage JD (1986) The benthic fauna of the Rockall Trough: Regional distribution and bathymetric zonation. Proceedings of the Royal Society of Edinburgh Section B: Biology 88:159-174

Galloway JL and Collins WT (1998) Dual frequency acoustic classification of seafloor habitat using the QTC VIEW. OCEANS 1998:1296-1300

Garrabou J, Riera J and Zabala M (1998) Landscape pattern indices applied to Mediterranean subtidal rocky benthic communities. Landscape Ecology 13:225-247

Gettleson D, Hammer R and Putt R (1982) Geophysical and biological seafloor mapping of four oil and gas lease blocks in the South Atlantic Georgia embayment. OCEANS 1982:803-808

Gonzalez-Mirelis G and Lindegarth M (2012) Predicting the distribution of out-of-reach biotopes with decision trees in a Swedish marine protected area. Ecological Applications 22:2248-2264

Gooday AJ, Aranda da Silva A and Pawlowski J (2011) Xenophyophores (Rhizaria, Foraminifera) from the Nazaré Canyon (Portuguese margin, NE Atlantic). Deep Sea Research Part II: Topical Studies in Oceanography 58:2401-2419

Gori A, Orejas C, Madurell T, Bramanti L, Martins M, Quintanilla E, Marti-Puig P, Lo Iacono C, Puig P, Requena S, Greenacre M and Gili JM (2013) Bathymetrical distribution and size structure of cold-water coral populations in the Cap de Creus and Lacaze-Duthiers canyons (northwestern Mediterranean). Biogeosciences 10:2049-2060

Grassle JF (1985) Hydrothermal vent animals: Distribution and biology. Science 229:713-717

Grassle JF (1991) Deep-sea benthic biodiversity. Bioscience 41: 464-469

Greene HG, Yoklavich MM, Starr RM, O'Connell VM, Wakefield WW, Sullivan DE, McRea Jr JE and Cailliet GM (1999) A classification scheme for deep seafloor habitats. Oceanologica Acta 22:663-678

Greene HG, Yoklavich MM, Sullivan D and Cailliet GM (1995) A geophysical approach to classifying marine benthic habitats: Monterey Bay as a model. In: O'Connell T, Wakefield W (eds) Applications of Side-scan Sonar and Laser-line Systems in Fisheries Research, Juneau, Alaska, p 15-30

Grober-Dunsmore R, Frazer T, Lindberg W and Beets J (2007) Reef fish and habitat relationships in a Caribbean seascape: The importance of reef context. Coral Reefs 26:201-216

Guisan A, Edwards Jr TC and Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: Setting the scene. Ecological Modelling 157:89-100

Guisan A and Thuiller W (2005) Predicting species distribution: Offering more than simple habitat models. Ecology Letters 8:993-1009

References

Guisan A and Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecological Modelling 135:147-186

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S and Peterson AT (2007) What matters for predicting the occurences of trees: Techniques, data, or species' characteristics? Ecological Monographs 77:615-630

Gutt J and Piepenburg D (2003) Scale-dependent impact on diversity of Antarctic benthos caused by grounding of icebergs. Marine Ecology Progress Series 253:77-83

Gutt J, Starmans A and Dieckmann G (1996) Impact of iceberg scouring on polar benthic habitats. Marine Ecology Progress Series 137:311-316

Haedrich RL, Rowe GT and Polloni PT (1980) The megabenthic fauna in the deep sea south of New England, USA. Marine Biology 57:165-179

Hall-Spencer J, Allain V and Fosså JH (2002) Trawling damage to Northeast Atlantic ancient coral reefs. Proceedings of the Royal Society B: Biological Sciences 269:507-511

Halpern BS (2003) The impact of marine reserves: Do reserves work and does reserve size matter? Ecological Applications 13:117-137

Harris PT and Whiteway T (2011) Global distribution of large submarine canyons: Geomorphic differences between active and passive continental margins. Marine Geology 285:69-86

Hartigan JA and Wong MA (1979) Algorithm AS 136: A K-means clustering algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics) 28:100-108

Hastie T, Tibshirani R and Friedman J (2009) The elements of statistical learning: Data mining, inference, and prediction. Springer, New York, USA

Hecker B (1982) Possible benthic fauna and slope instability relationships. In: Saxov S, Nieuwenhuis JK (eds) Marine Slides and Other Mass Movements, Vol 6. Springer US, p 335-347

Henry L-A, Davies A and Roberts JM (2010) Beta diversity of cold-water coral reef communities off western Scotland. Coral Reefs 29:427-436

Henry L-A and Roberts JM (2007) Biodiversity and ecological composition of macrobenthos on cold-water coral mounds and adjacent off-mound habitat in the bathyal Porcupine Seabight, NE Atlantic. Deep Sea Research Part I: Oceanographic Research Papers 54:654-672

Henry LA, Moreno Navas J and Roberts JM (2013) Multi-scale interactions between local hydrography, seabed topography, and community assembly on cold-water coral reefs. Biogeosciences 10:2737-2746

Hernandez PA, Graham CH, Master LL and Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773-785

Hewitt J, Thrush S, Legendre P, Funnell G, Ellis J and Morrison M (2004) Mapping of marine soft-sediment communities: Integrated sampling for ecological interpretation. Ecological Applications 14:1203-1216

Hewitt JE, Thrush SF, Halliday J and Duffy C (2005) The importance of small-scale habitat structure for maintaining beta diversity. Ecology 86:1619-1626

Hill MO (1973) Diversity and evenness: A unifying notation and its consequences. Ecology 54:427-432

Hirzel A and Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling 157:331-341

Howell KL (2010) A benthic classification system to aid in the implementation of marine protected area networks in the deep/high seas of the NE Atlantic. Biological Conservation 143:1041-1056

Howell KL and Davies JS (2010) Deep-sea species image catalogue. Marine Biology and Ecology Research Centre, Marine Institute at the University of Plymouth. On-line version http://www.marlin.ac.uk/deep-sea-species-image-catalogue/

Howell KL, Davies JS, Jacobs C and Narayanaswamy BE (2009) Broadscale survey of the habitats of Rockall Bank and mapping of Annex I 'reef' habitat, Joint Nature Conservation Committee, Peterborough, UK

Howell KL, Holt R, Endrino IP and 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

Hutchinson GE (1959) Homage to Santa Rosalia or why are there so many kinds of animals? American naturalist 93:145-159

Huvenne VAI (2011) European Research Council Starting Grant, no. 258482

Huvenne VAI, Blondel P and Henriet JP (2002) Textural analyses of sidescan sonar imagery from two mound provinces in the Porcupine Seabight. Marine Geology 189:323-341

Huvenne VAI, Huenerbach V, Blondel P and Le Bas TP (2007) Detailed mapping of shallow-water environments using image texture analysis on sidescan sonar and multibeam backscatter imagery Underwater Acoustic Measurements, Heraklion, Crete

Huvenne VAI, Tyler PA, Masson DG, Fisher EH, Hauton C, Hühnerbach V, Le Bas TP and Wolff GA (2011) A picture on the wall: Innovative mapping reveals cold-water coral refuge in submarine canyon. PLoS ONE 6:e28755

IOC IHO and BODC (2003) GEBCO digital atlas Centenary edition of the GEBCO digital atlas, published on CD-ROM on behalf of the Intergovernmental Oceanographic Commission and the International Hydrographic

References

  Organization as part of the General Bathymetric Chart of the Oceans, British Oceanographic Data Centre, Liverpool, U.K.

Isachenko A, Gubanova Y, Tzetlin A and Mokievsky V (*in press*) High-resolution habitat mapping on mud fields: New approach to quantitative mapping of Ocean quahog. Marine Environmental Research

Ismail K, Huvenne VAI and Masson DG (*under revision*) Automated classification approach to map potential habitats in submarine canyons. Marine Geology

Jackson DA, Peres-Neto PR and Olden JD (2001) What controls who is where in freshwater fish communities: The roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences 58:157-170

Jackson EL, Attrill MJ and Jones MB (2006) Habitat characteristics and spatial arrangement affecting the diversity of fish and decapod assemblages of seagrass (*Zostera marina*) beds around the coast of Jersey (English Channel). Estuarine, Coastal and Shelf Science 68:421-432

Jenness J (2012a) DEM Surface Tools v. 2.1.305. Jenness Enterprises

Jenness J (2012b) Land Facet Corridor Designer, v. 1.2.848. Jenness Enterprises

Jensen A and Frederiksen R (1992) The fauna associated with the bank-forming deepwater coral Lophelia pertusa (Scleractinaria) on the Faroe shelf. Sarsia 77:53-69

JNCC (2010a) Conservation objectives and advice on operations, Joint Nature Conservation Committee, Peterborough, UK

JNCC (2010b) Offshore special area of conservation: Wyville Thomson Ridge, Joint Nature Conservation Committee, Peterborough, UK

JNCC (2012) Offshore special area of conservation: Hatton Bank, Joint Nature Conservation Committee, Peterborough, UK

Johnson MP, White M, Wilson A, Würzberg L, Schwabe E, Folch H and Allcock AL (2013) A vertical wall dominated by *Acesta excavata* and *Neopycnodonte zibrowii*, part of an undersampled group of deep-sea habitats. PLoS ONE 8:e79917

Jones D, Bett B and Tyler PA (2007a) Depth-related changes to density, diversity and structure of benthic megafaunal assemblages in the Fimbul ice shelf region, Weddell Sea, Antarctica. Polar Biology 30:1579-1592

Jones DOB, Bett BJ and Tyler PA (2007b) Depth-related changes in the arctic epibenthic megafaunal assemblages of Kangerdlugssuaq, East Greenland. Marine Biology Research 3:191-204

Jones DOB and Brewer ME (2012) Response of megabenthic assemblages to different scales of habitat heterogeneity on the Mauritanian slope. Deep Sea Research Part I: Oceanographic Research Papers 67:98-110

Juhász-Nagy P and Podani J (1983) Information theory methods for the study of spatial processes and succession. Vegetatio 51:129-140

Jumars PA (1976) Deep-sea species diversity : Does it have a characteristic scale? Journal of Marine Research 34:217-246

Juniper SK, Matabos M, Mihály S, Ajayamohan RS, Gervais F and Bui AOV (2013) A year in Barkley Canyon: A time-series observatory study of mid-slope benthos and habitat dynamics using the NEPTUNE Canada network. Deep Sea Research Part II: Topical Studies in Oceanography 92:114-123

Kaandorp JA (1999) Morphological analysis of growth forms of branching marine sessile organisms along environmental gradients. Marine Biology 134:295-306

Kenchington EL, Cogswell AT, MacIsaac KG, Beazley L, Law BA and Kenchington TJ (2014) Limited depth zonation among bathyal epibenthic megafauna of the Gully submarine canyon, northwest Atlantic. Deep Sea Research Part II: Topical Studies in Oceanography 104:67-82

Kenny AJ, Cato I, Desprez M, Fader G, Schuttenhelm RTE and Side J (2003) An overview of seabed-mapping technologies in the context of marine habitat classification. ICES Journal of Marine Science 60:411-418

Klitgaard AB (1995) The fauna associated with outer shelf and upper slope sponges (Porifera, Demospongiae) at the Faroe Islands, northeastern Atlantic. Sarsia 80:1-22

Knudby A, Brenning A and LeDrew E (2010) New approaches to modelling fish-habitat relationships. Ecological Modelling 221:503-511

Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence

Koleff P, Gaston KJ and Lennon JJ (2003) Measuring beta diversity for presence-absence data. Journal of Animal Ecology 72:367-382

Kostylev VE (2002) A framework for the conservation of benthic communities of the Scotian-Funday area of the Maritimes region, Fisheries and Oceans Canada, Dartmouth, Canada

Kostylev VE, Todd BJ, Fader GBJ, Courtney RC, Cameron GDM and Pickrill RA (2001) Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. Marine Ecology Progress Series 219:121-137

Kullback S (1959) Information theory and statistics. Wiley, Chapman & Hall

Last PR, Lyne VD, Williams A, Davies CR, Butler AJ and Yearsley GK (2010) A hierarchical framework for classifying seabed biodiversity with application to planning and managing Australia's marine biological resources. Biological Conservation 143:1675-1686

References

Laughton AS, Hill MN and Allan TD (1960) Geophysical investigations of a seamount 150 miles north of Madeira. Deep Sea Research 7:117-141

Le Bas TP and Hühnerbach V (1998) PRISM: Processing of remotely-sensed imagery for seafloor mapping handbook. Southampton Oceanography Centre, UK

Le Bas TP and Huvenne VAI (2009) Acquisition and processing of backscatter data for habitat mapping - Comparison of multibeam and sidescan systems. Applied Acoustics 70:1248-1257

Leduc A, Drapeau P, Bergeron Y and Legendre P (1992) Study of spatial components of forest cover using partial Mantel tests and path analysis. Journal of Vegetation Science 3:69-78

Lee JAY (1991) Comparison of existing methods for building triangular irregular network, models of terrain from grid digital elevation models. International Journal of Geographical Information Systems 5:267-285

Legendre P (1993) Spatial autocorrelation: Trouble or new paradigm? Ecology 74:1659-1673

Legendre P, Dale MRT, Fortin M-J, Gurevitch J, Hohn M and Myers D (2002) The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25:601-615

Legendre P and Fortin M (1989) Spatial pattern and ecological analysis. Vegetatio 80:107-138

Legendre P and Legendre L (1998) Numerical ecology. Elsevier Science, Amsterdam

Leslie HM (2005) A synthesis of marine conservation planning approaches. Conservation Biology 19:1701-1713

Levin L, Etter R, Rex M, Gooday A, Smith C, Pineda J, Stuart C, Hessler R and Pawson D (2001) Environmental influences on regional deep-sea species diversity. Annual Review of Ecology and Systematics 32:51-93

Levin LA and Sibuet M (2012) Understanding continental margin biodiversity: A new imperative. Annual Review of Marine Science 4:79-112

Levin SA (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. Ecology 73:1943-1967

Levins R and Culver D (1971) Regional coexistence of species and competition between rare species. Proceedings of the National Academy of Sciences 68:1246-1248

Liu C, Berry PM, Dawson TP and Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385-393

Lo Iacono C, Gràcia E, Diez S, Bozzano G, Moreno X, Dañobeitia J and Alonso B (2008) Seafloor characterization and backscatter variability of the

Almería Margin (Alboran Sea, SW Mediterranean) based on high-resolution acoustic data. Marine Geology 250:1-18

Lo Iacono C, Orejas C, Gori A, Gili JM, Requena S, Puig P and Ribó M (2012) Habitats of the Cap de Creus continental shelf and Cap de Creus Canyon, Northwestern Mediterranean. In: Harris PT, Baker EK (eds) Seafloor Geomorphology as Benthic Habitat: GeoHab Atlas of Seafloor Geomorphic Features and Benthic Habitats. Elsevier Inc, London, UK, p 457-469

Loots C, Vaz S, Planque B and Koubbi P (2011) Understanding what controls the spawning distribution of North Sea whiting (*Merlangius merlangus*) using a multi-model approach. Fisheries Oceanography 20:18-31

Loreau M (2000) Biodiversity and ecosystem functioning: Recent theoretical advances. Oikos 91:3-17

Lucieer V and Lucieer A (2009) Fuzzy clustering for seafloor classification. Marine Geology 264:230-241

Lucieer VL (2008) Object oriented classification of sidescan sonar data for mapping benthic marine habitats. International Journal of Remote Sensing 29:905-921

Lutz MJ, Caldeira K, Dunbar RB and Behrenfeld MJ (2007) Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. Journal of Geophysical Research: Oceans 112:C10011

MacArthur R (1955) Fluctuations of animal populations and a measure of community stability. Ecology 36:533-536

MacArthur RH and MacArthur JW (1961) On bird species diversity. Ecology 42:594-598

MacArthur RH and Wilson EO (1963) An equilibrium theory of insular zoogeography. Evolution 17:373-387

Magorrian BH, Service M and Clarke W (1995) An acoustic bottom classification rurvey of Strangford Lough, Northern Ireland. Journal of the Marine Biological Association of the United Kingdom 75:987-992

Makarenkov V and Legendre P (2002) Nonlinear redundancy analysis and canonical correspondence analysis based on polynomial regression. Ecology 83:1146-1161

Malatesta RJ and Auster PJ (1999) The importance of habitat features in low-relief continental shelf environments. Oceanologica Acta 22:623-626

Manel S, Dias J-M and Ormerod SJ (1999) Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: A case study with a Himalayan river bird. Ecological Modelling 120:337-347

References

Manel S, Williams HC and Ormerod SJ (2001) Evaluating presence-absence models in ecology: The need to account for prevalence. Journal of Applied Ecology 38:921-931

Marmion M, Luoto M, Heikkinen RK and Thuiller W (2009a) The performance of state-of-the-art modelling techniques depends on geographical distribution of species. Ecological Modelling 220:3512-3520

Marmion M, Parviainen M, Luoto M, Heikkinen RK and Thuiller W (2009b) Evaluation of consensus methods in predictive species distribution modelling. Diversity and Distributions 15:59-69

Masson DG, Bett BJ, Billett DSM, Jacobs CL, Wheeler AJ and Wynn RB (2003) The origin of deep-water, coral-topped mounds in the northern Rockall Trough, Northeast Atlantic. Marine Geology 194:159-180

Matabos M, Bui AOV, Mihály S, Aguzzi J, Juniper SK and Ajayamohan RS (2014) High-frequency study of epibenthic megafaunal community dynamics in Barkley Canyon: A multi-disciplinary approach using the NEPTUNE Canada network. Journal of Marine Systems

McArthur MA, Brooke BP, Przeslawski R, Ryan DA, Lucieer VL, Nichol S, McCallum AW, Mellin C, Cresswell ID and Radke LC (2010) On the use of abiotic surrogates to describe marine benthic biodiversity. Estuarine, Coastal and Shelf Science 88:21-32

McClain C and Barry J (2010) Habitat heterogeneity, disturbance, and productivity work in concert to regulate biodiversity in deep submarine canyons. Ecology 91:964-976

McGarigal K, Cushman SA and Ene E (2012) FRAGSTATS (4.x): Spatial pattern analysis program for categorical and continuous maps. University of Massachusetts, Amherst, USA

McPhail S (2009) Autosub6000: A deep diving long range AUV. Journal of Bionic Engineering 6:55-62

Mellin C, Parrott L, Andrefouet S, Bradshaw CJA, MacNeil MA and Caley MJ (2012) Multi-scale marine biodiversity patterns inferred efficiently from habitat image processing. Ecological Applications 22:792-803

Menard HW and Dietz RS (1951) Submarine geology of the Gulf of Alaska. Geological Society of America Bulletin 62:1263-1285

Menge BA and Olson AM (1990) Role of scale and environmental factors in regulation of community structure. Trends in Ecology & Evolution 5:52-57

Menke SB, Holway DA, Fisher RN and Jetz W (2009) Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. Global Ecology and Biogeography 18:50-63

Menzies RJ, George RY and Rowe GT (1973) Abyssal environment and ecology of the world oceans. Wiley-Interscience, New York, USA

Micallef A, Le Bas TP, Huvenne VAI, Blondel P, Hühnerbach V and Deidun A (2012) A multi-method approach for benthic habitat mapping of shallow coastal areas with high-resolution multibeam data. Continental Shelf Research 39–40:14-26

Mienis F, de Stigter HC, White M, Duineveld G, de Haas H and van Weering TCE (2007) Hydrodynamic controls on cold-water coral growth and carbonate-mound development at the SW and SE Rockall Trough Margin, NE Atlantic Ocean. Deep Sea Research Part I: Oceanographic Research Papers 54:1655-1674

Monk J, Ierodiaconou D, Bellgrove A, Harvey E and Laurenson L (2011) Remotely sensed hydroacoustics and observation data for predicting fish habitat suitability. Continental Shelf Research 31:S17-S27

Monk J, Ierodiaconou D, Versace VL, Bellgrove A, Harvey E, Rattray A, Laurenson L and Quinn GP (2010) Habitat suitability for marine fishes using presence-only modelling and multibeam sonar. Marine Ecology Progress Series 420:157-174

Moran PAP (1950) Notes on continuous stochastic phenomena. Biometrika 37:17-23

Morell V (2007) Into the deep: First glimpse of Bering Sea canyons heats up fisheries battle. Science 318:181-182

Morris KJ, Tyler PA, Masson DG, Huvenne VIA and Rogers AD (2013) Distribution of cold-water corals in the Whittard Canyon, NE Atlantic Ocean. Deep-Sea Research Part II: Topical Studies in Oceanography 92:136-144

Morrison MA, Thrush SF and Budd R (2001) Detection of acoustic class boundaries in soft sediment systems using the seafloor acoustic discrimination system QTC VIEW. Journal of Sea Research 46:233-243

Mortensen PB and Buhl-Mortensen L (2005) Deep-water corals and their habitats in The Gully, a submarine canyon off Atlantic Canada. In: Freiwald A, Roberts JM (eds) Cold-Water Corals and Ecosystems. Springer Berlin Heidelberg, p 247-277

Mortensen PB, Buhl-Mortensen L, Gebruk AV and Krylova EM (2008) Occurrence of deep-water corals on the Mid-Atlantic Ridge based on MAR-ECO data. Deep Sea Research Part II: Topical Studies in Oceanography 55:142-152

Mortensen PB, Dolan M and Buhl-Mortensen L (2009) Prediction of benthic biotopes on a Norwegian offshore bank using a combination of multivariate analysis and GIS classification. ICES Journal of Marine Science 66:2026-2032

Newton RS and Stefanon A (1975) Application of side-scan sonar in marine biology. Marine Biology 31:287-291

Noss RF (1990) Indicators for monitoring biodiversity: A hierarchical approach. Conservation Biology 4:355-364

References

Oldeland J, Dorigo W, Lieckfeld L, Lucieer A and Jürgens N (2010) Combining vegetation indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation units from hyperspectral imagery. Remote Sensing of Environment 114:1155-1166

Orejas C, Gori A, Lo Iacono C, Puig P, Gili JM and Dale MRT (2009) Cold-water corals in the Cap de Creus canyon, northwestern Mediterranean: Spatial distribution, density and anthropogenic impact. Marine Ecology Progress Series 51:37-51

Oshiro T, Perez P and Baranauskas J (2012) How many trees in a Random forest? In: Perner P (ed) Machine Learning and Data Mining in Pattern Recognition, Vol 7376. Springer Berlin Heidelberg, p 154-168

Palialexis A, Georgakarakos S, Karakassis I, Lika K and Valavanis VD (2011) Prediction of marine species distribution from presence-absence acoustic data: Comparing the fitting efficiency and the predictive capacity of conventional and novel distribution models. Hydrobiologia 670:241-266

Paterson GLJ, Glover AG, Cunha MR, Neal L, de Stigter HC, Kiriakoulakis K, Billett DSM, Wolff GA, Tiago A, Ravara A, Lamont P and Tyler PA (2011) Disturbance, productivity and diversity in deep-sea canyons: A worm's eye view. Deep Sea Research Part II: Topical Studies in Oceanography 58:2448-2460

Pearce JL and Boyce MS (2006) Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43:405-412

Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E, Brotons L, McClean C, Miles L, Segurado P, Dawson TP and Lees DC (2006) Model-based uncertainty in species range prediction. Journal of Biogeography 33:1704-1711

Pedrini H (2008) Multiresolution terrain modeling based on triangulated irregular networks. Brazilian Journal of Geology 31:117-122

Peng J, Wang Y, Zhang Y, Wu J, Li W and Li Y (2010) Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. Ecological Indicators 10:217-223

Peres-Neto PR, Legendre P, Dray S and Borcard D (2006) Variation partitioning of species data matrices: Estimation and comparison of fractions. Ecology 87:2614-2625

Phillips NW, Gettleson DA and Spring KD (1990) Benthic biological studies of the Southwest Florida Shelf. American Zoologist 30:65-75

Phillips SJ and Dudík M (2008) Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161-175

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J and Ferrier S (2009) Sample selection bias and presence-only distribution models:

Implications for background and pseudo-absence data. Ecological Applications 19:181-197

Pianka ER (1966) Convexity, desert lizards, and spatial heterogeneity. Ecology 47:1055-1059

Pickett STA and Cadenasso ML (1995) Landscape ecology: Spatial heterogeneity in ecological systems. Science 269:331-334

Pittman SJ, McAlpine CA and Pittman KM (2004) Linking fish and prawns to their environment: A hierarchical landscape approach. Marine Ecology Progress Series 283:233-254

Preston JM, Christney AC, Bloomer SF and Beaudet IL (2001) Seabed classification of multibeam sonar images. OCEANS 2001:2616-2623

Przeslawski R, Currie DR, Sorokin SJ, Ward TM, Althaus F and Williams A (2011) Utility of a spatial habitat classification system as a surrogate of marine benthic community structure for the Australian margin. ICES Journal of Marine Science 68:1954-1962

Puig P, Canals M, Company JB, Martin J, Amblas D, Lastras G, Palanques A and Calafat AM (2012) Ploughing the deep sea floor. Nature 489:286-289

R Development Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing

Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M and Guisan A (2006) Are niche-based species distribution models transferable in space? Journal of Biogeography 33:1689-1703

Reid GS and Hamilton D (1990) A reconnaissance survey of the Whittard Sea Fan, Southwestern Approaches, British Isles. Marine Geology 92:69-86

Rengstorf AM, Grehan A, Yesson C and Brown C (2012) Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: Resolving terrain attribute dependencies. Marine Geodesy 35:343-361

Rengstorf AM, Yesson C, Brown C and Grehan AJ (2013) High-resolution habitat suitability modelling can improve conservation of vulnerable marine ecosystems in the deep sea. Journal of Biogeography 40:1702-1714

Rex MA (1981) Community structure in the deep-sea benthos. Annual Review of Ecology and Systematics 12:331-353

Rex MA, Etter RJ, Morris JS, Crouse J, McClain CR, Johnson NA, Stuart CT, Deming JW, Thies R and Avery R (2006) Global bathymetric patterns of standing stock and body size in the deep-sea benthos. Marine Ecology Progress Series 317:8

Rhoads D and Boyer L (1982) The effects of marine benthos on physical properties of sediments. In: McCall PL, Tevesz MJS (eds) Animal-Sediment Relations, Vol 100. Springer US, p 3-52

References

Riegl B and Piller WE (2000) Mapping of benthic habitats in northern Safaga Bay (Red Sea, Egypt): A tool for proactive management. Aquatic Conservation: Marine and Freshwater Ecosystems 10:127-140

Ries L, Fletcher RJ, Battin J and Sisk TD (2004) Ecological responses to habitat edges: Mechanisms, models, and variability explained. Annual Review of Ecology, Evolution and Systematics 35:491-522

Robert K, Jones DOB and Huvenne VAI (2014) Megafaunal distribution and biodiversity in a heterogeneous landscape: The iceberg scoured Rockall Bank, NE Atlantic. Marine Ecology Progress Series 501:67-88

Robert K, Jones DOB, Paul PA, Van Rooij D and Huvenne VAI (*accepted*) Finding the hot-spots within a biodiversity hotspot: Fine-scale biological predictions within a submarine canyon using high-resolution acoustic mapping techniques. Marine Ecology

Robert K, Jones DOB, Roberts JM and Huvenne VAI (*submitted*) (In) consistency of predictive mapping approaches for deep-water habitats: Considering multiple model outputs.

Roberts DG (1971) New geophysical evidence on the origins of the Rockall Plateau and Trough. Deep Sea Research and Oceanographic Abstracts 18:353-360

Roberts DG (1975) Sediment distribution on the Rockall Bank, Rockall Plateau Marine Geology 19:239-257

Roberts JM, Henry L-A, Long D and Hartley JP (2008) Cold-water coral reef frameworks, megafaunal communities and evidence for coral carbonate mounds on the Hatton Bank, north east Atlantic. Facies 54:297-316

Roberts S and Hirshfield M (2004) Deep-sea corals: Out of sight, but no longer out of mind. Frontiers in Ecology and the Environment 2:123-130

Rogers AD (1999) The biology of *Lophelia pertusa* (Linnaeus 1758) and other deep-water reef-forming corals and impacts from human activities. International Review of Hydrobiology 84:315-406

Ross RE and Howell KL (2012) Use of predictive habitat modelling to assess the distribution and extent of the current protection of 'listed' deep-sea habitats. Diversity and Distributions 19:433–445

Roy K, Jablonski D and Valentine JW (1996) Higher taxa in biodiversity studies: Patterns from Eastern Pacific marine molluscs. Philosophical Transactions: Biological Sciences 351:1605-1613

Sacchetti F, Benetti S, Ó Cofaigh C and Georgiopoulou A (2012) Geophysical evidence of deep-keeled icebergs on the Rockall Bank, Northeast Atlantic Ocean. Geomorphology 159–160:63-72

Sanders HL (1968) Marine benthic diversity: A comparative study. American naturalist 102:243-282

Sandwell DT, Smith WHF, Gille S, Kappel E, Jayne S, Soofi K, Coakley B and Géli L (2006) Bathymetry from space: Rationale and requirements for a new, high-resolution altimetric mission. Comptes Rendus Geoscience 338:1049-1062

Schiagintweit GEO (1993) Real-time acoustic bottom classification for hydrography a field evaluation of RoxAnn. OCEANS 1993

Schlacher TA, Williams A, Althaus F and Schlacher-Hoenlinger MA (2010) High-resolution seabed imagery as a tool for biodiversity conservation planning on continental margins. Marine Ecology 31:200-221

Schmitz OJ (2007) Ecology and ecosystem conservation. Island Press Washington, DC, USA

Schneider DC, Gagnon JM and Gilkinson KD (1987) Patchiness of epibenthic megafauna on the outer Grand Banks of Newfoundland. Marine Ecology Progress Series 39:1-13

Schoening T, Bergmann M, Ontrup J, Taylor J, Dannheim J, Gutt J, Purser A and Nattkemper TW (2012) Semi-automated image analysis for the assessment of megafaunal densities at the Arctic deep-sea observatory HAUSGARTEN. PLoS ONE 7:e38179

Sellanes J, Neira C, Quiroga E and Teixido N (2010) Diversity patterns along and across the Chilean margin: A continental slope encompassing oxygen gradients and methane seep benthic habitats. Marine Ecology 31:111-124

Shannon CE (1948) A mathematical theory of communications. The Bell System Technical Journal 27:379-423

Shumchenia EJ and King JW (2010) Comparison of methods for integrating biological and physical data for marine habitat mapping and classification. Continental Shelf Research 30:1717-1729

Simpson EH (1949) Measurement of diversity. Nature Australia 163:688

Sink K, Holness S, Harris L, Majiedt P, Atkinson L, Robinson T, Kirkman S, Hutchings L, Leslie R, Lamberth S, Kerwath S, Heyden Svd, Lombard A, Attwood C, Branch G, Fairweather T, Taljaard S, Weerts S, Cowley P, Awad A, Halpern B, Grantham H and Wolf T (2011) National biodiversity assessment 2011: Marine and coastal component. South African National Biodiversity Institute, Pretoria, South Africa, p 325

Slee JA (1932) Reflection methods of measuring the depth of the sea Institution of Electrical Engineers - Proceedings of the Wireless Section of the Institution, p 20-28

Smith CR, Jumars PA and DeMaster DJ (1986) *In situ* studies of megafaunal mounds indicate rapid sediment turnover and community response at the deep-sea floor. Nature 323:251-253

Sokal RR and Oden NL (1978) Spatial autocorrelation in biology: 1. Methodology. Biological Journal of the Linnean Society 10:199-228

References

Soltwedel T, Jaeckisch N, Ritter N, Hasemann C, Bergmann M and Klages M (2009) Bathymetric patterns of megafaunal assemblages from the arctic deep-sea observatory HAUSGARTEN. Deep Sea Research Part I: Oceanographic Research Papers 56:1856-1872

Stewart HA, Davies JS, Guinan J and Howell KL (2014) The Dangeard and Explorer canyons, South Western Approaches UK: Geology, sedimentology and newly discovered cold-water coral mini-mounds. Deep Sea Research Part II: Topical Studies in Oceanography 104:230–244

Stine RA (1995) Graphical interpretation of variance inflation factors. The American Statistician 49:53-56

Stoy P, Williams M, Disney M, Prieto-Blanco A, Huntley B, Baxter R and Lewis P (2009a) Upscaling as ecological information transfer: a simple framework with application to Arctic ecosystem carbon exchange. Landscape Ecology 24:971-986

Stoy PC, Williams M, Spadavecchia L, Bell RA, Prieto-Blanco A, Evans JG and van Wijk MT (2009b) Using information theory to determine optimum pixel size and shape for ecological studies: Aggregating land surface characteristics in Arctic ecosystems. Ecosystems 12:574-589

Teixidó N, Garrabou J and Arntz WE (2002) Spatial pattern quantification of Antarctic benthic communities using landscape indices. Marine Ecology Progress Series 242:1-14

Teixidó N, Garrabou J, Gutt J and Arntz W (2007) Iceberg disturbance and successional spatial patterns: The case of the shelf Antarctic benthic communities. Ecosystems 10:143-158

ter Braak CJF (1994) Canonical community ordination. Part I: Basic theory and linear methods Ecoscience 1:127-140

Tews J, Brose U, Grimm V, Tielbörger K, Wichmann MC, Schwager M and Jeltsch F (2004) Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures. Journal of Biogeography 31:79-92

Thistle D (1983) The role of biologically produced habitat heterogeneity in deep-sea diversity maintenance. Deep Sea Research Part A: Oceanographic Research Papers 30:1235-1245

Thomson CW (1874) The depths of the sea. MacMillan & Co, London, UK

Thrush SF, Halliday J, Hewitt JE and Lohrer AM (2008) The effects of habitat loss, fragmentation, and community homogenization on resilience in estuaries. Ecological Applications 18:12-21

Tilman D (2001) Functional diversity. In: Levin S (ed) Encyclopedia of Biodiversity. Academic Press, p 109-120

Tittensor DP, Baco AR, Brewin PE, Clark MR, Consalvey M, Hall-Spencer J, Rowden AA, Schlacher T, Stocks KI and Rogers AD (2009) Predicting

global habitat suitability for stony corals on seamounts. Journal of Biogeography 36:1111-1128

Toucanne S, Zaragosi S, Bourillet JF, Naughton F, Cremer M, Eynaud F and Dennielou B (2008) Activity of the turbidite levees of the Celtic–Armorican margin (Bay of Biscay) during the last 30,000 years: Imprints of the last European deglaciation and Heinrich events. Marine Geology 247:84-103

Turner M, Dale V and Gardner R (1989a) Predicting across scales: Theory development and testing. Landscape Ecology 3:245-252

Turner M, O'Neill R, Gardner R and Milne B (1989b) Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3:153-162

Turner MG (1989) Landscape ecology: The effect of pattern on process. Annual Review of Ecology and Systematics 20:171 -197

Turner MG and Gardner RH (1991) Quantitative methods in landscape ecology: The analysis and interpretation of landscape heterogeneity. Springer, USA

Tyler PA, Amaro T, Arzola R, Cunha MR, de Stigter H, Gooday A, Huvenne V, Ingels J, Kiriakoulakis K, Lastras G, Masson D, Oliveira A, Pattenden A, Vanreusel A, Van Weering T, Vitorino J, Witte U and Wolff G (2009) Europe's grand canyon: Nazaré submarine canyon. Oceanography and Marine Biology: An Annual Review 22:46-57

Valentine PC and Schmuck EA (1995) Geological mapping of biological habitats on Georges Bank and Stellwagen Bank, Gulf of Main region. In: Applications of Side-scan Sonar and Laser-line Systems in Fisheries Research, Juneau, Alaska, p 31-40

van Aken HM (2000) The hydrography of the mid-latitude Northeast Atlantic Ocean: II: The intermediate water masses. Deep Sea Research Part I: Oceanographic Research Papers 47:789-824

van den Wollenberg A (1977) Redundancy analysis an alternative for canonical correlation analysis. Psychometrika 42:207-219

Van Rooij D, De Mol L, Le Guilloux E, Wisshak M, Huvenne VAI, Moeremans R and Henriet JP (2010a) Environmental setting of deep-water oysters in the Bay of Biscay. Deep Sea Research Part I: Oceanographic Research Papers 57:1561-1572

Van Rooij D, Iglesias J, Hernández-Molina FJ, Ercilla G, Gomez-Ballesteros M, Casas D, Llave E, De Hauwere A, Garcia-Gil S, Acosta J and Henriet JP (2010b) The Le Danois Contourite Depositional System: Interactions between the Mediterranean Outflow Water and the upper Cantabrian slope (North Iberian margin). Marine Geology 274:1-20

Vandermeer JH (1972) Niche theory. Annual Review of Ecology and Systematics 3:107-132

References

Verfaillie E, Degraer S, Schelfaut K, Willems W and Van Lancker V (2009) A protocol for classifying ecologically relevant marine zones, a statistical approach. Estuarine, Coastal and Shelf Science 83:175-185

Vetter EW and Dayton PK (1998) Macrofaunal communities within and adjacent to a detritus-rich submarine canyon system. Deep Sea Research Part II: Topical Studies in Oceanography 45:25-54

Vetter EW and Dayton PK (1999) Organic enrichment by macrophyte detritus, and abundance patterns of megafaunal populations in submarine canyons. Marine Ecology Progress Series 186:137-148

Vetter EW, Smith CR and De Leo FC (2010) Hawaiian hotspots: Enhanced megafaunal abundance and diversity in submarine canyons on the oceanic islands of Hawaii. Marine Ecology 31:183-199

Vierod ADT, Guinotte JM and Davies AJ (2014) Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. Deep Sea Research Part II: Topical Studies in Oceanography 99:6-18

Vivoni E, Ivanov V, Bras R and Entekhabi D (2004) Generation of triangulated irregular networks based on hydrological similarity. Journal of Hydrologic Engineering 9:288-302

Wang L, Sousa WP and Gong P (2004) Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. International Journal of Remote Sensing 25:5655-5668

Ward TJ, Vanderklift MA, Nicholls AO and Kenchington RA (1999) Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. Ecological Applications 9:691-698

Warren DL, Glor RE and Turelli M (2008) Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. Evolution 62:2868-2883

Wedding L, Lepczyk C, Pittman S, Friedlander A and Jorgensen S (2011) Quantifying seascape structure: Extending terrestrial spatial pattern metrics to the marine realm. Marine Ecology Progress Series 427:219-232

Wentworth CK (1922) A scale of grade and class terms for clastic sediments. The Journal of Geology 30:377-392

Wheeler A, Beyer A, Freiwald A, de Haas H, Huvenne V, Kozachenko M, Olu-Le Roy K and Opderbecke J (2007) Morphology and environment of cold-water coral carbonate mounds on the NW European margin. International Journal of Earth Sciences 96:37-56

White M, Mohn C, Stigter H and Mottram G (2005) Deep-water coral development as a function of hydrodynamics and surface productivity around the submarine banks of the Rockall Trough, NE Atlantic. In:

Freiwald A, Roberts JM (eds) Cold-Water Corals and Ecosystems. Springer Berlin Heidelberg, p 503-514

Whiteway T, Heap A, Lucieer V, Hinde A, Ruddick R and Harris P (2007) Seascape of the Australian margin and adjacent sea floor: Methodology and results. GeoScience Australia, Canberra, Australia

Wienberg C, Beuck L, Heidkamp S, Hebbeln D, Freiwald A, Pfannkuche O and Monteys X (2008) Franken Mound: Facies and biocoenoses on a newly-discovered "carbonate mound" on the western Rockall Bank, NE Atlantic. Facies 54:1-24

Wiens JA (1989) Spatial scaling in ecology. Functional ecology 3:385-397

Williams A, Althaus F, Dunstan PK, Poore GCB, Bax NJ, Kloser RJ and McEnnulty FR (2010) Scales of habitat heterogeneity and megabenthos biodiversity on an extensive Australian continental margin (100-1100 m depths). Marine Ecology 31:222-236

Williams CB (1964) Patterns in the balance of nature and related problems in quantitative ecology. Academic Press, New York, USA

Wilson JB (1979a) The distribution of the coral *Lophelia pertusa* in the north-east Atlantic. Journal of the Marine Biological Association of the United Kingdom 59:149-164

Wilson JB (1979b) 'Patch' development of the deep-water coral *Lophelia pertusa* on Rockall Bank. Journal of the Marine Biological Association of the United Kingdom 59:165-177

Wilson JB and Desmond JM (1986) Shelf break faunas on the eastern margin of the Rockall Trough and Faeroe-Shetland Channel. Proceedings of the Royal Society of Edinburgh Section B: Biology 88:315-316

Wilson Jr RR, Smith Jr KL and Rosenblatt RH (1985) Megafauna associated with bathyal seamounts in the central North Pacific Ocean. Deep Sea Research Part A: Oceanographic Research Papers 32:1243-1254

Wilson MFJ, O'Connell B, Brown C, Guinan JC and Grehan AJ (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Marine Geodesy 30:3-35

Wood SN and Augustin NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecological Modelling 157:157-177

Wynn RB, Huvenne VAI, Le Bas TP, Murton BJ, Connelly DP, Bett BJ, Ruhl HA, Morris KJ, Peakall J, Parsons DR, Sumner EJ, Darby SE, Dorrell RM and Hunt JE (2014) Autonomous underwater vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience. Marine Geology 352:451-468

Yachi S and Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. Proceedings of the

References

National Academy of Sciences of the United States of America 96:1463-1468

Yoklavich M, Cailliet G and Moreno G (1993) Rocks and fishes: Submersible observations in a submarine canyon

Yoklavich MM, Cailliet GM, Greene HG and Sullivan D (1995) Interpretation of side-scan sonar records for rockfish habitat analysis: Examples from Monterey Bay. In: O'Connell T, Wakefield W (eds) Applications of Side-scan Sonar and Laser-line Systems in Fisheries Research. Alaska Department of Fish and Game, Juneau, USA, p 11-14

Zajac R (2008) Challenges in marine, soft-sediment benthoscape ecology. Landscape Ecology 23:7-18

Zajac RN, Lewis RS, Poppe LJ, Twichell DC, Vozarik J and DiGiacomo-Cohen ML (2003) Responses of infaunal populations to benthoscape structure and the potential importance of transition zones. Limnology and Oceanography 48:829-842